

CIVIL CONFLICT AND EDUCATION:

**HOW DOES EXPOSURE TO CIVIL CONFLICT AFFECT HUMAN CAPITAL
ACCUMULATION? EVIDENCE FROM STANDARDIZED EXIT EXAMS IN
COLOMBIA**

© 2015 Silvia Consuelo Gómez Soler

ISBN 978 90 8666 388 0

Publisher: Boekenplan, Maastricht

www.boekenplan.nl

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form, or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission in writing, from the author.

Civil Conflict and Education:
How Does Exposure to Civil Conflict
Affect Human Capital Accumulation?
Evidence from Standardized Exit Exams
in Colombia

DISSERTATION

to obtain the degree of Doctor at the Maastricht University,
on the authority of the Rector Magnificus Prof. dr. L.L.G. Soete,
in accordance with the decision of the Board of Deans,
to be defended in public on Tuesday, 17 November 2015, at 14:00 hours

by

Silvia Consuelo Gómez Soler

Promoter:

Prof. dr. Jaap Dronkers

Co-supervisor:

Dr. Jorge Restrepo

Associate Professor - Pontificia Universidad Javeriana - Bogotá, Colombia

Director of the Resource Centre for Conflict Analysis (CERAC) - Bogotá, Colombia

Assessment Committee:

Prof. dr. Rolf v.d. Velden

Prof. dr. Wim Groot

Dr. Lutz Krebs

Prof. Louis Volante, Brock University, Canada

Dr. Andrea Ichino, EUI Florence, Italy

Acknowledgements

This long journey has come to an end and I would to thank a group of very special people for their encouragement, support and guidance.

First, I would like to express my greatest gratitude to my promoter and supervisor, Prof. Jaap Dronkers, for his guidance and encouragement during the process of writing my PhD thesis. I am extremely grateful for his time, interest, and helpful comments. During these years, I have learned that one of the most important factors for a successful PhD experience is to have a supervisor that is truly committed to the process. In that sense I was extremely lucky. Thank you Prof. Dronkers!

I would also like to express my gratitude to my co-supervisor, Dr. Jorge Restrepo, who has been my mentor ever since I was an undergraduate student at Universidad Javeriana. His guidance and support have been essential in writing this book, and most importantly, in developing an academic career. Jorge, muchas gracias por su apoyo desde el primer día.

I would also like to thank the assessment committee, Prof. dr. Rolf v.d. Velden, Prof. dr. Wim Groot, Dr. Lutz Krebs, Prof. Louis Volante and Dr. Andrea Ichino for reading the dissertation and providing very valuable comments and suggestions to improve the contents of this book.

I am very grateful for the support that I have received from my current employer, Universidad de La Sabana, and my previous employer, Pontificia Universidad Javeriana. Muchas gracias a Ignacio Penagos, Juanita Cifuentes, Alejandro Gaviria y Patricia Castillo por su apoyo, compañerismo y comprensión durante este proceso. Esta tesis doctoral no sería posible sin el apoyo de mis amigos del Octavo. I would also like to thank Prof. Paul Winters, Prof. Martha Misas, Prof. Álvaro Montenegro, Dr. Ana María Díaz, Dr. Gonzalo Hernandez, Dr. Andrés Rosas and Dr. Flavio Jacome for their academic support.

I would like to express my gratitude to all the people responsible for the GPAC program (specially Mindel and Eddy). Thank you Mindel for always being there to support me and the other fellows of the program. I would also like to thank the coordination team (Carlos, Guney, Saba and Charlotte) for making our life easier during the GPAC seminars.

It has been a great privilege to meet people from all over the world and learn from them at the GPAC program. The weeks we spent at Maastricht were a great opportunity to share ideas and have great discussions about our research. I would like to say thank you to the fellows of my cohort (Paul, Clovis, Mutinta, Brenda, Ali, Luiz, Mohammed, Camilo, and Andrea) and the other higher year fellows (Derek, Ana, Corinne, Rafa, Shellie, Richard) for truly enriching this experience. En especial muchas gracias a Camilo y Andrea por su apoyo y amistad durante este proceso...todo en la vida es más fácil cuando se tiene el apoyo de buenos amigos. Gracias también por el apoyo de la colonia Colombiana en Maastricht University...Carlos, Andrea, Omar, Jennifer y Sergio.

Finalmente quisiera dedicarle esta tesis a mi novio y a mi familia. El apoyo y la paciencia que me han tenido en estos años ha sido clave para poder llegar a este final feliz. Amor muchas gracias por todo tu cariño y comprensión durante estos años. Eres mi balance, mi polo a tierra. Hiciste más fácil un trabajo extremadamente difícil. En muchas formas esta tesis también es tuya. A mis papas, muchas gracias por apoyarme en todo. No hay palabras para agradecerles por su ayuda incondicional en cada etapa de la vida que me ha permitido llegar a este momento. Ustedes son mi inspiración. A mi hermano, gracias por alentarme siempre a seguir aún en momentos difíciles. Hermanito muchas gracias por la buena energía que siempre me transmitiste durante este proceso: ¡lo logré!

Summary

Getting a better understanding of the possible effects of civil conflict on human capital is particularly important for the effective implementation of public policy. A civil conflict has significant economic, social and political repercussions that could potentially affect the educational achievement of individuals. Nevertheless, this important issue has been insufficiently addressed in the literature. The limited availability of academic research in this topic is partly due to the difficulties involved in the systematic collection of evidence on why, where and how attacks occur. The availability of a rich database about conflict events in Colombia, maintained by the Resource Centre for Conflict Analysis (CERAC), has opened new possibilities for the analysis of the relationship between conflict and education. In that sense, the Colombian case offers a unique opportunity for research in this topic. The objective of this thesis is to get a better understanding of the relationship between civil conflict and educational achievement, and quantify the magnitude in which civil conflict affects the accumulation of human capital measured by cognitive exams. The research presented in this dissertation concentrates on those students residing in conflict and non-conflict zones in Colombia who take part in the formal education system, as these are the only students that can be observed in the available micro data. Hence, the specific objective of this dissertation is to better understand the relationship between educational achievement and civil conflict amongst those students that participate in the formal educational system.

This dissertation contributes to the existing literature that studies the relationship between civil conflict and education, including the contributions of Swee (2008), Akresh and de Walque (2008), Merrouche (2006), Lai and Thyne (2007), Shemyakina (2011), Debalen and Paul (2012), Bundervoet (2012), Bellows and Miguel (2006), Rodríguez and Sánchez (2010), amongst others. These research efforts have focused on the relationship between school enrollment, gender, years of education and civil conflict, but have not studied the relationship between civil conflict and academic achievement measured by cognitive examinations as we do in this dissertation. The dissertation also contributes to the literature that studies the impact of internal displacement on the welfare of civilians, including the work of Ibañez and Velez (2008), Ibañez and Moya (2010a), Ruiz and Vargas (2013), Calderón and Ibañez (2009), Kirschhoff and Ibañez (2002), Kondylis (2008 and 2010), and Fiala (2009).

The dissertation is divided into eleven chapters. Four of these (Chapters 7, 8, 9 and 10) present econometric exercises to guide the analysis of the relationship between

civil conflict and students' achievement in national examinations. Chapter 7 explores the relationship between civil conflict and educational achievement in Colombia through the use of multilevel analysis. Multilevel analysis is applied in this chapter in order to take into account the hierarchical structure of the data by simultaneously modeling variables at different levels without resorting to aggregation or disaggregation of the dataset. The results obtained in this chapter show that the relationship between the intensity of the conflict and the performance in exit examinations might not be as straightforward as it may seem at first glance. For the year 2001 we found a positive relationship between our contemporaneous measure of the intensity of the conflict and the performance in the mathematics and language examinations, but for the year 2002 we found a negative relationship between the intensity of the conflict and the performance in the language examination.

In Chapter 8 we construct a pseudo-panel of schools from cross-sectional data sets in order to get a better understanding of the relationship between civil conflict and education. To the best of my knowledge this is the first effort to try to understand the relationship between conflict and achievement in standardized exams through the use of a pseudo panel of schools. The results obtained in this chapter provide evidence of the negative repercussions that a civil conflict can have on human capital accumulation. Our results show a negative and significant relation between our two measures of the intensity of the conflict and the levels of performance of schools in the mathematics and language examinations. The results also show a negative and significant relation between our second measure of the presence of conflict (dummy variable) and the levels of performance of schools in the mathematics and language examinations. However, the magnitude of this relationship was smaller than expected.

In Chapter 9 we use two methodologies to estimate the effects of forced displacement on the performance of Colombian students in the national high school exit examination: cluster robust inference and an instrumental variables approach. In cluster robust inference observations are grouped into clusters, with model errors uncorrelated across clusters but correlated within clusters. However, this is our naïve approach in the sense that we were not considering the possibility of endogeneity. Following Calderon and Ibañez (2009), we used an instrumental variables approach as our second method of analysis in order to deal with the possibility of endogeneity as a source of bias. The main finding of this chapter is that being displaced has a negative and significant effect on the levels of achievement of students in mathematical and language exit examinations. The results obtained using the instrumental variables approach show a larger effect of displacement on the performance in the examinations. These results are not surprising if we take into consideration the findings of a number of researchers

who have studied the effects of displacement. The literature has shown that there are a number of needs and vulnerabilities associated to the displacement status that affect the well-being of individuals, and in doing so, might also be affecting the performance of students in standardized tests.

Chapter 10 estimates the effect of civil conflict on student achievement gains in standardized examinations. By considering students who have been exposed to a conflict environment and students who have not been exposed to a conflict environment, we want to find out whether Colombian students affected by conflict have differential gains or losses in performance in comparison to those students who have not been affected, using the results from the Colombian high school exit examination (Saber11) and the Colombian college exit examination (SaberPro). To the best of my knowledge, this is the first attempt to study the relationship between civil conflict and academic achievement measured by cognitive examinations at both high school and university levels. We used a difference in differences estimation strategy in order to quantify the student's learning gains using information at two points in time (high school and university). The main finding of this chapter is that students who have been exposed to conflict have more significant improvements in their academic performance in comparison to those students who have not been affected. This finding is robust to the different measures of conflict (intensity and presence) that were used in the empirical exercise of this chapter.

The evidence obtained in this thesis can be used to guide the formulation of educational policy. Policy makers should work in the creation of new and better incentives to retain students who reside in conflict zones in school with a full time dedication, and if possible, provide financial assistance so that qualified students get the opportunity to attend university. It is possible that those students who are facing difficulties associated to the internal armed conflict are developing high levels of resilience that allows them to continue with their studies successfully, but they need a special follow up. More governmental resources should be spent in order to meet the special needs of those students who have been affected by conflict (i.e. psychological help, school supplies, special remedial tutorials). In particular, some of these resources should be spent in the design of special programs to supervise the academic progress of students at educational institutions. The design of effective policies to guarantee that those students who stay in conflict zones have the adequate incentives and resources to continue with their studies is an essential task that should be prioritized in the agenda of Colombian policy makers.

The government and development agencies should provide funding in order to assure that schools in conflict zones provide a high quality education that is

inclusive. Special efforts should be made at conflict zones to strengthen the capacity of the education authorities, and to promote organizational development at schools. The public authorities should make the necessary arrangements in order to protect schools. This is particularly important in order to guarantee that schools are safe learning spaces where students can feel protected. Furthermore, these spaces should provide specialized guidance for kids so that they receive adequate support for learning and psychosocial development. Authorities should also monitor school attendance. This is particularly important in order to have an early detection system of those students who are more likely to dropout from school.

The limitations of this dissertation are illustrative of the empirical, theoretical and methodological challenges that this area of research faces. For instance, it is important to acknowledge that given the nature of the datasets under analysis, self-selection bias and sample selection bias might have been a concern in the empirical exercises presented in this dissertation. Another limitation that should be acknowledged is associated to constraints in terms of data availability. Throughout this dissertation we pointed out a number of factors that could help us explain the results obtained in the empirical exercises. However, in some cases given data restrictions we were not able to empirically test the importance of these factors in explaining our results. In order to analyse most of the factors pointed out we would need to use mixed methods (i.e. qualitative and quantitative analysis).

Without a doubt there is much more that can be learned about civil conflict and education. In fact, a very important topic that should be further explored is the relationship between academic achievement at a young age (elementary school) and civil conflict. This topic will be explored in my future research agenda.

CONTENTS

Acknowledgements	i
Summary	iii
Chapter 1 - Introduction	1
Chapter 2 – General Historical Overview of Colombia and Its Civil Conflict	16
Chapter 3 – General Historical Overview of Education in Colombia.....	24
Chapter 4 – Literature Review	39
Chapter 5 – Modeling Conflict and Educational Achievement.....	53
Chapter 6 – Measuring Educational Outcomes – External Exit Examinations.....	69
Chapter 7 – Civil Conflict and Educational Achievement: Multilevel Analysis Approach.....	85
Chapter 8 – Educational Achievement at Schools: Assessing the Effect of the Civil Conflict using a Pseudo panel of Schools	175
Chapter 9 – Forced Displacement and Educational Achievement in Bogota’s Public Schools	221
Chapter 10 – Higher Education: Assessing the Effect of the Civil Conflict Amongst Those Individuals Who Attend University	275
Chapter 11 – Conclusions and Policy Implications	389
Valorisation	403
About the author	408
MGSOG Dissertation Series.....	409

CHAPTER 1

INTRODUCTION

International literature has shown that the economic, political and social consequences of civil war¹ are significant (Justino 2010). In fact, as mentioned by Justino (2011), violent conflict is a very considerable deterrent to worldwide development efforts. Armed groups attack the civil population to strengthen territorial strongholds, expand territorial control, weaken the support of the opponent, and accumulate valuable assets. The relationship between conflict, marginalization and social exclusion is clear: generating fear as a war strategy aims at impeding collective action, damaging social networks, and intimidating and controlling civilian population (Justino 2011).

A civil war has important economic, social and political implications that potentially interfere with the development process of the countries affected by it. More specifically, a civil conflict could potentially affect the human capital accumulation process via education. However, this important issue has been insufficiently addressed in the literature. Benefield and Tomlinson (2005) highlight that education in conflict and post-conflict situations, as a recognized field of research, is in its infancy. According to Davies, cited in Benefield and Tomlinson (2005), the connection between conflict and education is an under-analysed area. This is partly due to the difficulties involved in the systematic collection of evidence on why, where and how attacks occur. As mentioned by Justino (2010), high quality monitoring data and systematic reporting of events are very limited. Additionally, in many places affected by conflict censorship of information is also common. Further supporting that observation, Sommers, cited by Benefield and Tomlinson (2005), highlights that the statistical imprecision of data on populations affected by wars represents an important constraint on the ability to accurately estimate war's impact on education. As Bruck et al. (2010) point out, "the current

¹ The terms civil war, non-international armed conflict, internal armed conflict and civil conflict are used interchangeably in this dissertation. According to the International Committee of the Red Cross, there is no real difference between a non-international armed conflict and a civil war. In fact, Article 3 of the Geneva Conventions does not use the term civil war. The Geneva Conventions use the term armed conflict not of an international character to refer to internal conflicts. Furthermore, according to Andersen et al. (2009): "the legal definition of a civil war is a non-international armed conflict".

Other researchers have made a distinction between civil wars and civil conflicts. According to Blattman and Miguel (2010), civil wars are "those internal conflicts that count more than 1000 battle deaths in a single year", whereas civil conflicts are "those that count at least twenty-five battle deaths per annum." However, following Blattman and Miguel's (2010) definitions it would still be appropriate to use the terms civil war and civil conflict interchangeably for the Colombian case.

state of art of empirical research on violent conflict offers therefore considerable opportunity for improving our knowledge of violent conflict itself, its functions and dynamics, as well as the impact of conflict on behaviour, welfare and overall development.”

Although few, there are some publications that have attempted to study and understand the relationship between civil conflict and education, including Swee (2008), Akresh and de Walque (2008) and Bellows and Miguel (2006). However, the research efforts that have been done up to this date have mostly focused on the relationship between school enrolment², gender³, years of education⁴ and conflict, but have not tried to understand the relationship between civil conflict and academic achievement measured by cognitive examinations. To the best of my knowledge, only one effort has been made to try to understand the relationship between conflict and achievement in standardized exams. Such effort was made by Rodríguez and Sánchez (2010). This doctorate thesis contributes significantly to the literature because it introduces substantial improvements to Rodríguez and Sanchez’s (2010) work in terms of methodology and data quality.

As noted previously, the relationship between civil conflict and educational achievement is an area of research that has been understudied, opening an important gap in the literature. Such gap is explained in large part by the limited availability and reliability of micro data about civil conflict events and standardized test results. Taking advantage of a unique database about civil conflict events in Colombia administered by a Colombian think tank (Resource Centre for Conflict Analysis - CERAC), a government kept database containing the results of standardized tests at different levels of schooling, and additional information about schools and teachers from the National Department of Statistics of Colombia (DANE), the objective of this doctorate thesis is to get a better understanding of the relationship between civil conflict and educational achievement, and quantify the magnitude in which civil conflict affects the accumulation of human capital measured by cognitive exams.

As a result of the existing restrictions in terms of availability and reliability of micro data, research agendas have relied on macro level measures of conflict like the number of battle deaths per country per year (Bruck et al. 2010). The results from these research efforts and the policy implications derived can be limited. This is because it is very hard, if not impossible, to decompose the individual effects from a macro level study. For instance, using macro level measures it is very difficult to know who is affected by violence, to what degree, what the welfare

² Swee (2008), Bellows and Miguel (2006), Akresh and de Walque (2008), Lai and Thyne (2007).

³ Annan et al. (2011), Kecmanovic (2012), Valente (2011), Shemyakina (2011), Justino (2011).

⁴ Dabalen and Paul (2012), Bundervoet (2012), Merrouche (2006).

implications are, what are the social and political transformations that occur in societies affected by armed violence, and what are the channels by which violence affects welfare and behaviour (Bruck et al. 2010). The availability of a rich database about conflict events in Colombia, maintained by the Resource Centre for Conflict Analysis (CERAC), opens new possibilities for the analysis of the relationship between conflict and education. In that sense, the Colombian case offers a unique opportunity for research in this topic.

The proposed approach, which takes into consideration educational achievement and not academic desertion, is of relevance given the implications of obtaining high marks in national exams in terms of future academic and labour opportunities. The results of standardized exams are used in many countries as a filter to be admitted to higher levels of education, to qualify for scholarships and to be accepted in job applications. Additionally, recent papers, including Chetty et al. (2011), have documented that test scores predict long-run socioeconomic outcomes (e.g., earnings, employment, arrests, etc.). Given the uncertainty about the magnitude in which civil conflict is affecting the academic performance of individuals, it is not possible to make a definite account about the effects of civil conflict on human capital accumulation via education. At this point, owing to the insufficient amount of research in this topic, it is not possible to conclude whether or not the conflict is creating a gap in terms of future academic and work possibilities.

Investment in Human Capital and Standardized Test Scores

Human capital theory has been established as the standard framework of analysis in the economics of education thanks to the contributions of Becker (1962) and Schultz (1961). The ideas set forth by these two researchers suggest that education should be analyzed as an investment in human capital. Education could be seen as an investment that improves productivity, and in doing so, increases the prospects for higher earnings amongst individuals receiving formal schooling. Becker (1962) defines investment in human capital as those activities that influence future real income through the imbedding of resources in people. In other words, an investment in human capital entails the attainment of competencies that will have an impact on the future income of the individual making the investment. For instance, expenditures on education and health in order to have access to better job opportunities are examples of investment in human capital (Schultz 1961). According to Schultz (1961), human capital theory suggests that individuals are able to improve their production and consumption capabilities by investing in themselves. In doing so, most investments in human capital both raise earnings at older ages because the returns to the investment are added to earnings then, and lower earnings at young age because the costs associated to the investment are

deducted from earnings then (Becker 1962). Formal schooling plays a key role in the expansion of human capital. According to Pritchett and Filmer (1997), the academic literature has acknowledged that expansion in the skills, knowledge, and capacities of individuals, which in turn increases human capital, is a fundamental factor in economic progress and raising living standards.

The academic literature has acknowledged that quantifying the acquisition of cognitive skills (student achievement) is not an easy task. Nevertheless, a group of researchers have chosen to evaluate the acquisition of cognitive skills by analyzing the students' performance on standardized tests (Hanushek 2008). Hanushek (2008) points out that achievement measured by standardized tests has a clear impact on earnings after controlling for differences in the quantity of schooling, the experiences of workers, and other factors influencing earnings. Quoting Hanushek (2008), higher achievement "as measured by tests similar to those currently being used in accountability systems around the world is closely related to individual productivity and earnings". Furthermore, Hanushek (2008) highlights that there are a number of advantages derived from doing research using standardized test scores. For instance, measures of cognitive skills, including standardized test scores, capture variations in the knowledge that schools are aiming to produce and therefore are possibly linked to the relationship between schooling and labor market success. Furthermore, standardized exams allow for differences in performance among students who have had access to the same years of education, and by doing so, opens the possibility to do research designed to understand why these differences occur.

Research Questions

As already mentioned, the goal of this doctorate thesis is to get a better understanding of the relationship between civil conflict and educational achievement, and quantify the magnitude in which civil conflict affects the accumulation of human capital measured by cognitive exams. There are a number of mechanisms that could be at work when examining the relationship between civil conflict and human capital accumulation (educational outcomes). Some examples of those mechanisms include: Are kids scared and thus can't focus on schoolwork? Do kids have to dropout from school because they have to leave town (displacement)? Are families making a special effort to provide a high quality education to their kids in order to keep them away from the influence of the conflict? Are teachers scared and thus can't focus on teaching? Are schools in conflict zones well equipped?

These mechanisms could be grouped into two central categories: those that affect students who stay in the conflict zone and those that affect students who are

displaced because of civil conflict. However, it is only possible to empirically model the case of those students that can be observed, i.e. students that stay in the school system. Therefore, our attention in this study concentrates on those students residing in conflict and non-conflict zones who decide to take part in the formal education system. Hence, the objective of this research is to better understand the relationship between educational achievement and civil conflict amongst those students that participate in the formal educational system.

It is relevant to explore the particular case of the students that stay in the conflict zones because these students might be adapting (or not) to the rough conditions that are prevalent in these areas, and perhaps they have a different (smaller/larger) educational gap (in terms of magnitude) than policy makers think. Although at first glance it might seem straight forward that there is a negative relationship between civil conflict and education, there are reasons to think that this relationship is not as straightforward as it seems. The existing literature studying the relationship between educational enrolment and civil conflict has shown mixed results. For instance, Swee (2008) and Bellows and Miguel (2006) find no significant effects of civil war on enrolment, whereas Lai and Thyne (2007) and Akresh and de Walque (2008) find evidence to show that civil war has a negative effect on enrolment rates. Furthermore, the researchers that have studied the relationship between education, conflict and gender have also found mixed results. That is the case of Shemyakina (2011) and Annan et al. (2011) who have obtained totally different results in their studies of the impact of armed conflict on school enrolment by gender. Shemyakina (2011) shows that exposure to conflict had a negative effect on the school enrolment of girls, and little or no effect on enrolment of boys, whereas Annan et al. (2011) find significant adverse effect for males but not for females. The lack of consistency in the existing literature studying the relationship between conflict and education casts doubts about the results obtained to date, and motivates further research in this topic in order to explore the relationship in more detail.

It is also pertinent to study the case of those students who have been forced to flee from their hometown. The available literature for developing countries has shown that there are serious consequences of forced migration in terms of consumption, employment, and health conditions (Ruiz and Vargas 2013). Forced displacement is triggered by the need to find a safe place to live given the violent environment generated by the armed conflict. Ibañez and Velez (2003) highlight that violence and forced displacement are strongly linked because the strategies adopted by the illegal armed groups (e.g. massacres, life threats, selected homicides) can potentially trigger forced migration. Nevertheless, the lack of research studying the effects of displacement, explained in large part by the unavailability of data to

conduct proper analysis, limits the understanding of this phenomenon (Ruiz and Vargas 2013).

The main questions that occupy the attention of this thesis are summarized in *Table 1*. In the chapters that follow, I will address these questions in an attempt to fill some of the existing gaps in the literature.

Table 1 - Main Research Questions		
	Question	Sub-questions
1	How does the exposure to conflict affect the accumulation of human capital?	Is there a relationship between being exposed to conflict and the performance in standardized exit exams?
		Is that relationship significant?
		If there is a relationship, how strong is it?
2	What are the human capital losses associated with conflict and why do they occur?	Through what mechanisms does exposure to conflict influence performance in standardized exit exams?
3	What types of students are most likely to suffer from conflict?	How strong is the relationship between being exposed to conflict and the educational outcomes of students based on gender?
		How strong is the relationship between being exposed to conflict and the educational outcomes of students based on the geographical region?
		How strong is the relationship between being exposed to conflict and the educational outcomes of students based on the education of the parents and levels of income?

Source: author

These questions are studied using a solid theoretical backbone and an empirical approach supported by econometric methodologies. In order to get a better understanding of the mechanisms through which exposure to conflict influences human capital accumulation via education (*question 2*), this dissertation relies on an extended version of the educational production function which was originally formulated by Hanushek (1979, 1986, 1997) to examine the relationship between schooling inputs and test score outcomes. In this case, Hanushek's educational production function is re-formulated to incorporate conflict events as an environmental factor that determines the educational achievement of students in

standardized exams. This adaptation is discussed in detail in *chapter 5 (Theoretical Model)* of the dissertation. Having this theoretical backbone is a key component in analysing the factors that can explain the existence, or not, of a relationship between conflict and academic achievement in standardized exams. In that sense, the theoretical model maps into the econometric/empirical model. On the other hand, *Questions 1 and 3* will be addressed econometrically using a number of different techniques, which include: multilevel analysis, cluster robust inference, instrumental variables, difference in differences approach and a pseudo panel approach. The application of these econometric techniques will be explained in detail in *chapters 7, 8, 9 and 10* of this dissertation.

Empirical Chapters

Chapter 7

The purpose of chapter 7 is to explore the relationship between civil conflict and educational achievement in Colombia through the use of multilevel analysis. Multilevel analysis is employed in this chapter in order to deal adequately with the hierarchical structure of the data by simultaneously modeling variables at different levels without resorting to aggregation or disaggregation of the dataset. Taking advantage of a unique database administered by a Colombian think tank (Resource Centre for Conflict Analysis - CERAC) that contains records of civil conflict events in Colombia, a government kept database containing the results of standardized tests at different levels of schooling, and additional information about schools and teachers from the National Department of Statistics of Colombia (DANE), we were able to construct a multilevel model to further investigate the relationship between civil conflict and educational achievement.

The results obtained in this chapter show that the relationship between the intensity of the conflict and the performance in exit examinations is not as straightforward as it may seem at first glance. For the year 2001 we find a positive relationship between our contemporaneous measure of the intensity of the conflict and the performance in the mathematics and language examinations, but for the year 2002 we only find a negative relationship between the intensity of the conflict and the performance in the language examination. We contextualize these results using the contributions from other researchers, including Rodriguez and Sanchez (2012), Sánchez and Díaz (2005), Barrera and Ibañez (2004) and Wharton and Oyelere (2012), to help us establish possible transmission channels through which the intensity of the conflict could affect the performance of students in exit examinations, and in doing so, help us get a better understanding of the results. We also incorporate into the analysis a historical perspective of the evolution of the

Colombian conflict in order to get a more complete understanding of the results obtained.

Chapter 8

In Colombia there are no panel datasets available to study the educational achievement of students enrolled in the formal educational system, but there is a very complete series of independent cross-sections managed by the Colombian Institute for the Evaluation of Education (ICFES). Due to the lack of longitudinal data, in this chapter we construct a pseudo-panel from the Saber11 examination cross-sectional data sets. These datasets contain the individual level results from the years 2000, 2002, 2005 and 2007. According to Deaton (1985), a pseudo-panel can be constructed in order to track ‘cohorts’, i.e. a group with fixed membership, just as if panel data were available. In this case, the individual results are grouped in school ‘cohorts’. This chapter differs significantly from previous efforts because we construct a pseudo panel of schools in order to get a better understanding of the relationship between civil conflict and education. In fact, to the best of my knowledge this is the first effort to try to understand the relationship between conflict and achievement in standardized exams through the use of a pseudo panel of schools.

The results obtained in this chapter provide evidence of the negative repercussions that a civil conflict can have on human capital accumulation. Our results show a negative and significant relation between our two measures of the intensity of the conflict and the levels of performance in the mathematics and language examinations at the school level. Several control variables were also included as part of the specifications considered. Most of the control variables, including the variables controlling for the school meeting time, gender, age, number of teachers per student, expenditure of the municipality in education and tuition, were significant in all the specifications considered. According to our findings, there is a negative relationship between being enrolled in a nighttime school and the performance in the exit examination at the school level. The results also indicate that the cost of tuition has a significant relationship with the performance in the mathematics and language examination at the school level. We also found that the variable that captures the gender composition is significant under all the specifications considered. The evidence obtained in this chapter can be used to guide the formulation of educational policy. Special efforts should be made at conflict zones to strengthen the capacity of the education authorities, and to promote organizational development at schools.

Chapter 9

The purpose of this chapter is to identify the effects of forced displacement on the performance of Colombian students in the national high school exit examination. In this chapter we are focusing in those individuals who have been directly affected by conflict given their internal displacement status, instead of looking at the overall effect of conflict like we did in previous chapters. The Secretariat of Education of Bogota, a public entity in charge of managing and executing policies and programs in order to guarantee the rights of individuals to access education in the capital city of Colombia, provided us a detailed database for the years 2011 and 2012, containing information about those students who are attending public schools in Bogota and have been internally displaced because of the Colombian civil conflict. In this chapter we use two different methodologies to estimate the effects of forced displacement on the performance of Colombian students in the national high school exit examination. Our first approach uses cluster robust inference to estimate the effect of displacement. In cluster robust inference observations are grouped into clusters, with model errors uncorrelated across clusters but correlated within clusters. However, this is our naïve approach in the sense that we are not considering the possibility of endogeneity. Following Calderon and Ibañez (2009), we also use an instrumental variable approach as our second method of analysis in order to deal with the possibility of endogeneity as a source of bias.

The results obtained in this chapter provide clear evidence of the negative repercussions that displacement can have on human capital accumulation. For both years, the results of the estimations using OLS cluster robust inference and instrumental variables show that there is a negative and significant effect of displacement on the performance in the mathematics and language exit examinations. However, the results obtained using the instrumental variables approach show a larger effect of displacement on the performance in the examinations. The results show that school age individuals who have been displaced from conflict zones are prone to lower levels of academic achievement, and therefore to lower levels of human capital accumulation. In that sense, our results ratify the conclusions of other authors regarding the need for governmental programs specially designed to accommodate the particular needs of the displaced population in the short and medium term.

Chapter 10

The objective of this chapter is to estimate the effect of civil conflict on student achievement gains in standardized examinations. We are going to study the possibility of differential achievement gains amongst students coming from

conflict zones and non conflict zones by quantitatively analyzing the differences in high school exit examination scores and university exit examination scores. To the best of my knowledge, this is the first attempt to study the relationship between civil conflict and academic achievement measured by cognitive examinations at both high school and university levels. More specifically, by considering students who have been exposed to a conflict environment and students who have not been exposed to a conflict environment, we want to find out whether Colombian students affected by conflict have differential gains or losses in performance in comparison to those students who have not been affected, using the results from the Colombian high school exit examination (Saber11) and the Colombian college exit examination (SaberPro). To tackle this research question we are going to use the theoretical framework employed in educational value added models. Educational value added models are models that use student examination results at different points of time in order to measure their learning gain. Following this theoretical backbone, we are going to apply a difference in differences estimation strategy in order to quantify the student's learning gains using information at two points in time.

The main finding of this chapter is that students who have been exposed to conflict have more significant improvements in their academic performance in comparison to those students who have not been affected. This finding was robust to the different measures of conflict (intensity and presence) that we used in the empirical exercise of this chapter. There are a number of factors that could be taken into consideration to explain the results obtained in the econometric exercise. It is particularly important to establish the factors that could explain why those students who came from conflict zones experienced more significant achievement gains in the standardized examinations, and in doing so, help us get a better understanding of the results obtained in this chapter. There are three factors that could play an important role in the rationalization of the results obtained in the econometric exercise: **(i)** resilience, **(ii)** future expectations and **(iii)** self-selection.

Structure of the Dissertation

The rest of this dissertation is organized as follows: chapter 2 introduces a general historical overview of the Colombian civil conflict; chapter 3 presents a brief historical overview of education in Colombia; chapter 4 reviews the academic literature about conflict and education currently available; chapter 5 presents the theoretical model that supports the empirical exercises of chapters 7, 8, 9 and 10; chapter 6 describes Colombia's exit examination and discusses different issues related to measuring educational outcomes; chapter 7 reviews and analyses the results of an empirical exercise using multilevel analysis; chapter 8 reports and analyses the results of the empirical exercise using a pseudo panel approach;

chapter 9 presents the results of an empirical exercise that analyses the relationship between forced displacement and educational achievement; chapter 10 reports and analysis the results of an empirical exercise that explores the relationship between civil conflict and the academic performance of those students who attend higher education; and chapter 11 presents conclusions and policy implications. The final section of every chapter of this dissertation is devoted to the bibliography and the annexes. It is important to highlight that the empirical chapters of this dissertation are written as publishable papers.

WORKS CITED

- (CERAC), Centro de Recursos para el Análisis del Conflicto (2013). *Conflict data*.
- (DANE), Departamento Administrativo Nacional de Estadística (2008). Metodología de la Investigación de Educación Formal. In Departamento de Metodología y Producción Estadística (Ed.), (Vol. 2, pp. 35). Bogotá: Departamento Administrativo Nacional de Estadística (DANE).
- (ICFES), Instituto Colombiano para la Evaluación de la Educación (2010). Orientaciones para el Examen de Estado de la Educación Media (pp. 30). Bogotá: Instituto Colombiano para la Evaluación de la Educación (ICFES).
- Andersen, D., Barten, U., and Jensen, P. S. (2009). Challenges to Civil War Research-Introduction to the Special Issue on Civil War and Conflicts. *JEMIE*, 8, 1.
- Akresh, R., and De Walque, D. (2008). Armed conflict and schooling: Evidence from the 1994 Rwandan genocide. *World Bank Policy Research Working Paper Series*.
- Annan, J., Blattman, C., Mazurana, D., and Carlson, K. (2011). Civil war, reintegration, and gender in Northern Uganda. *Journal of conflict resolution*, 55(6), 877-908.
- Barrera, F., and Ibáñez, A. M. (2004). Does Violence Reduce Investment in Education?: A Theoretical and Empirical Approach: UNIVERSIDAD DE LOS ANDES-CEDE.
- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *The journal of political economy*, 70(5), 9-49.
- Bellows, J., and Miguel, E. (2006). War and institutions: New evidence from Sierra Leone. *The American economic review*, 96(2), 394-399.
- Benefield, P. and Tomlinson, K. (2005). Education and Conflict: Research and Research Possibilities: National Foundation for Educational Research (NFER).
- Blattman, C., and Miguel, E. (2010). Civil war. *Journal of Economic Literature*, 3-57.

- Brück, T., Justino, P., Verwimp, P. and Avdeenko, A. (2010). Identifying Conflict and Violence in Micro-Level Surveys. *IZA Discussion Paper*, (No. 5067). <http://ssrn.com/abstract=1648352>
- Bundervoet, T. (2012). War, Health, and Educational Attainment: A Panel of Children during Burundi's Civil War: Households in Conflict Network.
- Calderón, V., and Ibáñez, A. M. (2009). *Labor market effects of migration-related supply shocks: Evidence from internally displaced populations in Colombia*: CEDE, Centro de Estudios sobre Desarrollo Económico, Facultad de Economía, Universidad de los Andes.
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, 126(4), 1593-1660.
- Dabalen, A., and Paul, S. (2012). Estimating the causal effects of conflict on education in Cote d'Ivoire. *World Bank Policy Research Working Paper* (6077).
- Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of econometrics*, 30(1), 109-126.
- Dueñas, L. F., and Sanchez, F. (2007). *Does Armed Conflict Affect the Poorest?: The Case of School Dropout in Eastern Colombia*. . Documentos CEDE. Universidad de los Andes. Bogota.
- Glewwe, P., and Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. *Handbook of the Economics of Education*, 2, 945-1017.
- Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources*, 14(3), 351-388.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 24(3), 1141-1177.
- Hanushek, E. A. (1997). Assessing the Effects of School Resources on Student Performance: An Update. *Educational evaluation and policy analysis*, 19(2), 141-164. doi: 10.3102/01623737019002141

- Hanushek, E. A. (2008). Education Production Functions. In L. E. B. Steven N. Durlauf (Ed.), *The New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan.
- Ibañez, A. M., and Velez, C. E. (2008). Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia. *World Development*, 36(4), 659-676. doi: 10.1016/j.worlddev.2007.04.013
- International Committee of the Red Cross (2013). Non-international Armed Conflict. Retrieved May 5, 2015, from <https://www.icrc.org/eng/resources/documents/interview/2012/12-10-niac-non-international-armed-conflict.htm>
- Justino, P. (2010). How Does Violent Conflict Impact on Individual Educational Outcomes? The Evidence So Far. *background paper for the Education For All Global Monitoring Report*.
- Justino, P. (2011). Violent Conflict and Human Capital Accumulation. *IDS Working Papers*, 2011(379), 1-17.
- Kecmanovic, M. (2012). The Short-run Effects of the Croatian War on Education, Employment, and Earnings. *Journal of Conflict Resolution*.
- Lai, B. and Thyne, C.. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, 44(3), 277-292. doi: 10.1177/0022343307076631
- Machel, G., and Nationer, F. (1996). *Impact of armed conflict on children*: UN.
- Merrouche, O. (2006). The human capital cost of landmine contamination in Cambodia: Households in Conflict Network.
- Pritchett, L., and Filmer, D. (1997). What education production functions really show: a positive theory of education expenditures. *Economics of education review*, 18(2), 223-239.
- Rodríguez, C., and Sánchez, F. (2010). Books and guns: the quality of schools in conflict zones. *Documento CEDE* (38).
- Rodriguez, C., and Sánchez, F. (2012). Armed conflict exposure, human capital investments, and child labor: evidence from Colombia. *Defence and peace economics*, 23(2), 161-184.

- Ruiz, I., and Vargas-Silva, C. (2013). The economics of forced migration. *The Journal of Development Studies*, 49(6), 772-784.
- Sánchez, F., and Díaz, A. M. (2005). *Los efectos del conflicto armado en el desarrollo social colombiano, 1990-2002*: CEDE, Centro de Estudios sobre Desarrollo Económico, Facultad de Economía, Universidad de los Andes.
- Schultz, T. W. (1961). Investment in human capital. *The American economic review*, 51(1), 1-17.
- Secretariat of Education of Bogota (2012). Bases de Datos Colegios. Bogotá - Colombia.
- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, 95(2), 186-200. doi: 10.1016/j.jdeveco.2010.05.002
- Swee, E. L. (2009). On war and schooling attainment: The case of Bosnia and Herzegovina. *Households in Conflict Network Working Paper*, 57.
- Valente, C. (2011). What did the Maoists ever do for us? education and marriage of women exposed to civil conflict in Nepal.
- Wharton, K. and Oyelere, R. (2012). *Conflict and Its Impact on Educational Accumulation and Enrollment in Colombia: What We Can Learn from Recent IDPs*. IZA Discussion Paper. IZA. Retrieved from <http://ssrn.com/abstract=1922011>

CHAPTER 2

GENERAL HISTORICAL OVERVIEW OF COLOMBIA AND ITS CIVIL CONFLICT

This chapter presents a general historical overview of Colombia and its civil conflict. The aim of this chapter is to introduce the historical context of the Colombian civil conflict in order to set the scene for the quantitative analysis that will be developed in future chapters. Additionally, a brief historical account of the most prominent guerrilla groups and drug trafficking cartels of Colombia is included to complement the contents of this chapter.

How It All Began

Colombia's pre-Columbian history began over 20,000 years ago, when independent groups occupied small areas scattered throughout the Andean region and along the Pacific and Atlantic coasts. The Chibcha, Calima, Muisca, Nariño, Quimbaya, Arawak, San Agustín, Sinú, Tayrona, Tierradentro, Tolima and Tumaco groups lived in a set of separate but organized, agriculturally based communities (Hudson 2010). In 1499 a Spanish expedition led by Alonso de Ojeda, first came to the now-a-days Colombian territory. The colonists founded the first important mainland settlement, Santa María la Antigua de Darién on the Gulf of Urabá in 1510. Twenty-eight years later, in 1538 Jimenez de Quesada founded Santa Fe de Bogota (Hudson 2010). As the Spanish empire expanded in the New World, a new territorial division was created in 1717, the Virreinato del Nuevo Reino de Granada. Santa Fe de Bogota became the capital of the Virreinato, which comprised the present territories of Colombia, Panama, Ecuador and Venezuela (Hudson 2010).

The conditions imposed by the Spanish Empire in the Virreinato were very harsh. Slavery, the monopoly of commerce, taxes and duties, among other things, gave rise to protests. In 1781, Colombians were able to show their discontent in la Revolución Comunera at El Socorro. This revolt was the starting point for the pro-independence struggle. Simón Bolívar, the hero of the independence, fought several battles to claim Colombia's independence. The last and most decisive battle that led to Colombia's independence took place at Boyacá on August 7, 1819 (Hudson 2010).

After independence, Bolívar became the first President of Gran Colombia (Colombia, Venezuela, Panama and parts of Ecuador), with General Francisco de Paula Santander as his vice president. This union did not last long. The followers

of Bolívar and Santander soon divided, setting the stage for the country's long history of political division and violence. When Ecuador and Venezuela seceded in 1830, Gran Colombia dissolved. What was left emerged as la República de Nueva Granada, with Santander as its first president from 1832 to 1837 (Hudson 2010).

Two political parties emerged after the struggle for independence; the Partido Conservador Colombiano and the Partido Liberal. These two political parties dominated Colombian politics for the next 160 years. In fact, according to Legrand (2001), a unique trait of Colombia is the depth of the affiliation of the citizens to these two political parties. Such depth of affiliation led to a lot of enmity between the members of the two political parties. This rivalry between the two political forces resulted in numerous insurrections and civil wars. Actually, throughout the 19th century Colombia experienced more than eight civil wars. However, the most grave of these wars was the Thousand Days War (*Guerra de los Mil Días*). In 1899 a Liberal revolt led to the Thousand Days War, affecting the entire country and causing over 100,000 casualties. This war went on until 1902, generating great instability in all the Colombian territory.

As a result of the unfortunate events and the very high number of casualties caused by the Thousand Days War, in 1903 Colombia was going through many internal struggles. The difficult times that were prevailing in Colombia eased Panama's, then a Colombian province, secession. On November 3, 1903, Panama declared independence from Colombia (Hudson 2010). This event is definitely a turning point in Colombian history.

After a period of relative stability of about forty years, in April 1948 the assassination of the Liberal leader Jorge Eliecer Gaitán led to another major violent outburst now known as the times of the Violence (*La Violencia*). In fact, the time period going from 1946 to 1965 is now commonly known in Colombia as *La Violencia*. The death of Gaitán led to massive riots in the capital city (Bogotá). An important number of people died and were wounded during this one day riot, now known as the *Bogotazo*. After the death of Jorge Eliecer Gaitán and the *Bogotazo*, the tensions between the Liberal and the Conservative parties intensified. According to Legrand (2001), political tensions between the leaders of the two parties led to clashes in the countryside between peasants supporting the two sides. Armed groups of peasants supporting the Conservative party and supporting the Liberal party were formed during that time. These groups fought each other affecting principally the rural population and the inhabitants of small cities. In fact, during this civil conflict over 300,000 people were killed. According to LeGrand (2001), *La Violencia* is historically one of the civil conflicts in the Western Hemisphere with the greatest number of casualties.

In 1958 the Conservatives and Liberals came together to form the National Front (Frente Nacional) responding to their concern with the social conflict that was getting out of control. The National Front arrangement was a pact between the Conservatives and the Liberals in which the two parties would alternate the presidency of the country. In other words, although elections were still held everyone knew who would win: a Liberal candidate, then a Conservative candidate, and so on (Legrand 2001). Additionally, the presidents had to invite opposition figures to hold cabinet positions as part of the pact (Hudson 2010). This agreement, which lasted until 1974, helped to greatly reduce the violence in the early sixties.

However, as mentioned by Restrepo et al (2004), “during the Frente Nacional and until 1989, the Colombian conflict was essentially a Cold War proxy fight, with some partisan guerrillas transforming themselves into established guerrilla groups, most of them associated with various communist factions.” The exclusion of dissident political forces from the Frente Nacional contributed to the emergence of guerrilla groups in the mid-1960s, who were looking for more participatory power. The Liberal movements created during the time of the *Violencia* gave rise to left wing guerrilla movements in the sixties. The Cuban revolution and their successful guerrilla tactics inspired these new movements. In 1964, the Fuerzas Armadas Revolucionarias de Colombia (FARC) was founded and later on became the largest guerrilla group in Colombia. In 1965, the Ejército de Liberación Nacional (ELN) and the Ejército de Liberación Popular (EPL) began operations. Some years later in 1974 the Movimiento 19 de Abril (M-19) was created (Hudson 2010). It should be highlighted that the two groups that have had the most significant participation in the Colombian civil conflict are the FARC and the ELN.

Guerrilla Groups

Fuerzas Armadas Revolucionarias de Colombia (FARC)

The Colombian Communist Party was founded in the twenties and it became immediately involved with the rights of tenant farmers, public land claims and rights over communal lands (Legrand 2001). As a result, the Communist Party gained support in many areas of the countryside. In the fifties, during the time of the *Violencia*, these rural areas became known as independent peasant republics and served as refuge zones for many peasants that were displaced because of political violence. In 1964, the National Front attacked the independent peasant republics and its inhabitants had to move to new locations. Most of the people moved to the northern part of the southern jungles of Colombia (Legrand 2001). The Fuerzas Armadas Revolucionarias de Colombia (FARC) originated in these

new settlements, and this region of the country became the power base of this guerrilla group. According to Legrand (2001), the FARC was then a real peasant movement that wanted to respond to the violent attacks of the government and military repression. It was the armed branch of the Colombian Communist Party.

However, the motivations of the Fuerzas Armadas Revolucionarias de Colombia changed drastically in the late seventies and early eighties. Whereas in the sixties and in the seventies, the FARC was a self-defence movement living in rural territories, in the eighties it started attacking police stations and army bases, and ambushing army patrols (Legrand 2001). To do so, these groups got money from taxes imposed to coca growers and the ransoms from kidnappings. It should be highlighted that the conditions to raise coca, the principal component of cocaine, were adequate in the territories where the Fuerzas Armadas Revolucionarias de Colombia had presence. Motivated by the expansion of the international drug trade, the inhabitants of these territories started to raise coca commercially in the late seventies.

The Colombian government began peace negotiations with the Fuerzas Armadas Revolucionarias de Colombia in 1982. This was one of the many attempts that have been made in Colombia to make a peace agreement. In 1984, after many meetings and a period of negotiation, many members of the FARC agreed to put down their arms and create a political party (Legrand 2001). The new political party was known as the Patriotic Union (*Unión Patriótica* - UP). However, the negotiations turned out to be unsuccessful as hostilities continued. Some of the members of the Patriotic Union Party who were running for political office or were playing an important role in administrative matters were assassinated by hired killers (Legrand 2001). According to Lee (2012), the FARC replied to the attacks against the members of the Patriotic Union Party and the unsuccessful negotiations by engaging in crime and escalating their military action. These unfortunate events radicalized the position of the FARC about the need for an armed struggle (Lee 2012). According to Johnson and Jonnson (2013), in the period going from 1982 to 1999 the FARC went from approximately 2,000 members to 18,000 full time members and 12,000 militia members. Such growth had direct repercussions in the intensity of the conflict. In the mid-nineties, the FARC perpetrated a number of pungent attacks against the Colombian armed forces and the civilian population. During that period the guerrilla groups seized control of military and police outposts, and took numerous military personnel as hostages (Johnson and Jonnson 2013).

In 1999, the FARC once again began peace negotiations with the government setting a new set of conditions. The guerrilla group asked for control of a part of the Colombian territory that would serve as a demilitarized. However, this zone

did not serve its purpose of facilitating the peace talks. As mentioned by Lee (2012), by the end of the nineties, the power of this guerrilla group had grown considerably in the area and it was using the demilitarized zone for guerrilla training and drug trafficking. Lee (2012) points out that the FARC's misuse of the demilitarized zone justified a major offensive of the Colombian government in order to recover this territory. Johnson and Jonnson (2013) highlight that, as it was evident, in 2002 the negotiations fell apart and this gave way to years of intensive combat. The objective of the democratic security program that started in 2002 was to recover the territory that was under control of the illegal armed groups by military means, reduce or eliminate the production of drugs, and increase the size and competencies of the armed forces (Johnson and Jonnson 2013). This program had the technical and monetary support of the government of the United States. Between 2002 and 2008, the government was successful in reducing significantly the size and capability of the FARC. Additionally, in the northern territory of Colombia the FARC was being attacked as well by paramilitary forces. The Colombian government also established a disarmament and reintegration program to convince the guerrilla members to abandon the conflict. As mentioned by Johnson and Jonnson (2013), by the year 2009 the size of the FARC had fallen to about 8,500 members.

In the last few years, the FARC has recovered some of the power that it lost during the years of the democratic security program. According to Johnson and Jonnson (2013), this could be explained in part by the FARC's development of effective countermeasures like decreasing the use of electronic communications, moving away from populated areas, and relying more on snipers and landmines.

In the year 2012, the Colombian government announced once again the beginning of peace negotiations with the FARC. The agenda that was set covered core elements of the Colombian conflict, like, political participation, drug trafficking, rural development, and the victims (Johnson and Jonnson 2013). The official rounds of negotiations started in November 2012 and have continued in the year 2013.

The National Liberation Army (ELN)

The National Liberation Army (*Ejército Nacional de Liberación - ELN*) is another major guerrilla group in Colombia, which is still active. According to Legrand (2001), the ELN was formed in Santander, a Department of Colombia, in the sixties by university students who had gone to Cuba. In that sense, the ELN was created under the influence of the Cuban revolution. The guerrilla group started its operations in 1962 with a core group formed by 16 men (Sanchez et al 2005). The ELN also has its origins in the Liberation Theology movement led by priests of the

Latin American Catholic Church. In fact, the priest Camilo Torres was one of the first members of the ELN. However, he died in his first combat as an ELN member becoming an inspiration for other members of this guerrilla group (Legrand 2001).

According to Sanchez et al (2005), initially the objectives of the ELN were to defeat the national oligarchy, to object the power of the armed forces, to go against the North American imperialistic economic, political, and military interests, and to furnish the poor population with greater power. This guerrilla group had a strong support amongst groups of society with idealistic ideas, including university students. As mentioned by Sanchez et al (2005), the ideology of this group was a key factor to attract new members.

In the seventies, the ELN was led by the priest Manuel Perez and Nicolas Rodriguez, aka Gabino. Between 1966 and 1973, the ELN grew considerably reaching 270 members by 1973 (Sanchez et al 2005). However that same year, the retaliation of the Colombian armed forces for the attack on the police post of Anorí almost annihilated the guerrilla group. The ELN survived the attack and was able to reconstitute itself after some years of slow recovery. After facing this intense crisis, the guerrilla group was revitalized in the eighties largely due to the money inflow from extortions that were targeted at multinational companies trading in natural resources (Legrand 2001). The strategies of the ELN also started to resemble those of the FARC during those years (Sanchez et al 2005).

Now-a-days, the ELN is the second largest guerrilla group in Colombia. It is still principally financed by extortion, but it also receives income from alliances with drug trafficking cartels, and kidnapping (Sanchez et al 2005). This group continues causing damage to the infrastructure and carrying out violent actions against the civilian population and the State forces.

Paramilitary Groups

The paramilitary groups in Colombia were mostly gathered under the umbrella of the United Self Defense Groups of Colombia (*Autodefensas Unidas de Colombia - AUC*). At the end of the eighties and beginning of the nineties, the paramilitary groups became notorious because of their strong connections with the drug cartels (Restrepo et al 2004). The Death to the Kidnappers paramilitary group not only killed members of the guerrilla groups, but they also killed members of political groups that were against drug trafficking. Starting in the mid-nineties, the paramilitary groups fought the guerrilla and terrorized principally the rural population and human rights workers suspected of supporting or sympathizing with the guerrillas (Hudson 2010). The paramilitary groups played a very important role in the Colombian civil war because of the devastating actions that

they carried out in the nineties and in the beginning of this century against the guerrilla groups and the civilian population. In the year 2003, the Colombian government started demobilisation talks with the paramilitary groups. At the end of that year the demobilisation process began (Restrepo et al 2004). In the year 2005, the Justice and Peace Law (*Law 975*) was approved by the Colombian Congress as a framework to judge the crimes of the paramilitaries. This law has received a lot of criticism by human rights organizations and other entities because of its concessions to paramilitary combatants. It was widely perceived as a law that granted impunity to violators. New paramilitary groups continued functioning after the demobilisation of the United Self Defence Groups of Colombia (*AUC*). These groups are known as criminal gangs (*bandas criminales*). The criminal gangs attack both civilians and members of the public administration. More specifically, they carry out extortions against the civilian population and kidnap civilians and public authorities.

Concluding Remarks

This chapter presented a historical context of the Colombian civil conflict. The purpose of this chapter was to set the historical scene for the quantitative analysis that will be developed in future chapters. An explanation of the origins of the most prominent guerrilla groups of Colombia was included as part of this brief historical account to provide the readers with a better understanding of the Colombian conflict. However, in the coming chapters additional details will be provided about the Colombian conflict to set the scene for each particular analysis.

WORKS CITED

- Hudson, R. (2010). *Colombia: A Country Study*. Washington D.C.: Library of Congress.
- Johnson, K., and Jonsson, M. (2013). Colombia: Ending the Forever War? *Survival*, 55(1), 67-86.
- Lee, C. (2012). The FARC and the Colombian Left Time for a Political Solution? *Latin American Perspectives*, 39(1), 28-42.
- LeGrand, C. C. (2003). The Colombian crisis in historical perspective. *Canadian Journal of Latin American and Caribbean Studies*, 28(55-56), 165-209.
- Offstein, N. (2003). An Historical Review and Analysis of Colombian Guerrilla Movements: FARC, ELN, and EPL. *Desarrollo y Sociedad*, 52(1), 99-142.
- Restrepo, J., Spagat, M., and Vargas, J. F. (2004). The Dynamics of the Colombian Civil Conflict: A New Data Set.
- Sánchez, F., Solimano, A., and Formisano, M. (2005). Conflict, violence and crime in Colombia. *Understanding Civil Wars*, 2, 119-160.
- Sánchez, G., and Bakewell, P. (1985). La violencia in Colombia: new research, new questions. *The Hispanic American Historical Review*, 65(4), 789-807.

CHAPTER 3

GENERAL HISTORICAL OVERVIEW OF EDUCATION IN COLOMBIA

This chapter presents a general historical overview of education and educational policy in Colombia. The aim of this chapter is to describe the evolution of the educational sector and the different policies adopted in order to set the scene for the quantitative analysis that will be developed in future chapters. Historical statistics of the educational sector in Colombia were obtained from a number of different sources and were employed in this chapter in order to enrich the descriptive account.

Education in Colombia: Its Evolution

According to Ramirez and Tellez (2006), the expansion of primary and secondary education in Colombia during the first half of the twentieth century took place at a very slow rhythm. In fact, according to these two researchers, the country had low levels of schooling compared to other countries with similar levels of development (Ramirez and Tellez 2006). A document from the Colombian Development Plan highlights that up to the fifties, Colombia was in relative terms the most backward country in Latin America with respect to education. During that time period, analphabetism levels went up to 49% for the rural population and the coverage of primary education was as low as 50% (Departamento Nacional de Planeación 1990). In line with the previous observation, the expected years of education for the population stayed unchanged in 1.4 years per person during the first half of the twentieth century (Departamento Nacional de Planeación 1990).

In the late fifties, and especially during the times of the Frente Nacional, a great expansion of the educational system took place. This expansion partly helped overcome the backwardness that had ruled the educational system in Colombia for almost one hundred years (Departamento Nacional de Planeación 1990). In fact, according to a document from the Colombian Development Plan, the mean years of schooling received per generation went up continuously, reaching 6 years by the mid-seventies. Ramirez and Tellez (2006) point out that the first stage of this expansion took place even though the country was going through a hard time because of the negative repercussions associated to political violence. The confrontation between the conservative party and the liberal party that was going on during this time period, and the considerable number of casualties that resulted from this confrontation did not stop the expansion of the educational system (Ramirez and Tellez 2006).

Table 1		
Analphabetism Population Older than 15 Years Old (as a percentage)		
	1951	1964
Urban Analphabetism	21.1	15
Rural Analphabetism	49.6	41.3
Analphabetism amongst males	35	25.2
Analphabetism amongst females	40.2	30.5
Total Analphabetism	37.7	27

Source: Departamento Nacional de Planeación (1969)

During the fifties and the sixties, important efforts were made to improve the planning process of the educational system. More specifically, the missions that were led during the fifties by Professor Lauchlin Currie and Louis-Joseph Lebreton, two specialists in development and education, and the mission led by the World Trade Organization at the end of the sixties, provided important contributions to improve the educational planning process during that time period (Ramirez and Tellez 2006). Helg (1987) highlights the importance of these missions. According to this author, the missions provided an impartial analysis of the current situation of the educational system and formulated potential policies to help in the process of its expansion, overcoming potential problems associated to the political tensions of the time period. Important problems that the Colombian educational system was facing at the end of the first half of the twentieth century in terms of coverage, quality, efficiency, inequality between the rural and the urban sectors, infrastructure, and poor educational levels of the teachers, were highlighted (Ramirez and Tellez 2006). Specifically, the reports stressed the need to prioritize actions leading towards improvements in the conditions for primary education and rural education. The feedback from the different missions, led to important changes that ultimately resulted in a new era of educational planning in Colombia. Evidence of this, in the fifties the Office of Educational Planning was created within the Ministry of Education. As mentioned by Ramirez and Tellez (2006), this new office became responsible for the development plans of the educational sector based on the diagnosis of the sector and the needs of the population.

In 1957, a constitutional amendment was approved by popular vote, stating that at least 10% of the National budget should be allocated in public education. The results of this plebiscite showed that the Colombian population perceived education as a vital priority. It was clear that the population wanted an opportunity to receive at least a basic level of schooling. As a consequence, the

national government defined that primary education was fundamental for the development of the country and implemented policies to improve the opportunities to have access to at least a basic level of schooling. In line with that announcement, in 1960, Law 111 was enacted determining that primary education should be financed centrally by the national government, although departmental authorities had to keep the responsibility of naming teachers and managing schools (Ministerio de Educación 2010).

Some of the efforts made during this time period seem to have improved the educational conditions of the population. According to a document from the Colombian Development Plan, the percentage of analphabetism amongst people older than 15 years old went from 37.7% in 1951 to 27% in 1964 (see table 1). Additionally, the mean years of schooling per person amongst people older than 15 years old, during that same period, went up from 1.9 years to 3 years (Departamento Nacional de Planeación 1969). Unfortunately the big differences between the rural and urban regions continued during this time period. Specifically, in 1964 the mean years of schooling were 5.1 and 1.7 years for the urban and rural population, respectively (Departamento Nacional de Planeación 1969). This phenomenon could be associated with the urbanization process that was going on in Colombia during this time period. According to Ramirez and Tellez (2006), the degree of urbanization increased between 1951 and 1964, going from 39% to 52%. These two researchers highlight that the dynamic of the urbanization process in Colombia had its highest point in the fifties and at the beginning of the sixties (Ramirez and Tellez 2006).

Table 2			
National Government Expenditure in Education (1960/1965/1968)			
Billion pesos (nominal)			
Year	National Government Expenditure in Education	Total National Government Expenditure	Expenditure in Education as a Percentage of Total Expenditure
1960	449.1	4431.4	10.14
1965	905.8	7718.4	12.45
1968	1341	10851.1	12.4

Source: Departamento Nacional de Planeación (1969)

The improvement in the educational indicators that was highlighted above could also be explained by the higher allocation of fiscal resources for education during these years. These resources were available during this time period thanks to the favourable conditions of the economy. According to Ramirez and Tellez (2006), public expenditure in education as a percentage of GDP grew considerably since

the mid-fifties, going from around 1% in the sixties to almost 3% in the seventies. It should also be highlighted that starting in the sixties, the participation of the national budget in education, in relation with the total expenditure, has been higher than the 10% that was approved in the 1957 plebiscite (Departamento Nacional de Planeación 1969). More specifically, according to Ramirez and Tellez (2006), the budget for education as a percentage of the total expenditure increased during the fifties, the sixties and the seventies, going from 6.7% in 1950 to 19.8% in 1976.

According to the Ministry of Education, the increment in public resources devoted to education and the expansion of private education that resulted from higher demand pressures led to an important expansion of educational opportunities from the fifties up to the eighties (Ministerio de Educación 2010). To be more specific, in thirty years school attendance grew significantly considering that gross coverage at the primary level went up from 43% in 1951 to 97% in 1980, and at the secondary level went from 4.5% to 35.5% (Ministerio de Educación 2010). In the period from 1964 to 1977, enrolment grew by 92% at the primary level of instruction, and by 314% at the secondary level. As a result of this improvement in the levels of enrolment, between 1964 and 1978, the percentage of the economically active population that did not have any formal education went down from 27.1% to 16.4%. The percentage of people that had a higher level of instruction than primary education went up from 12.3% to 28.7% (Departamento Nacional de Planeación 1978).

An important outcome of the increment in the percentage of expenditure of the national government geared towards education in the seventies is the creation of the National Institutes of Diversified Middle Education (*Institutos Nacionales de Educación Media Diversificada - INEM*). These institutes were created to develop an alternative methodology for secondary education, which made emphasis in the training for work. The instruction imparted in these institutions combined the subjects of a classic secondary education with vocational courses in different areas related to industrial, commercial and agricultural duties (Iregui et al. 2006).

The Seventies and the Eighties

In the second half of the seventies, the process of centralization, nationalization and expansion of the educational system started to show important problems. According to the Ministry of Education, these problems were associated to political, administrative, financial and pedagogical issues (Ministerio de Educación 2004). According to a document from the Colombian Development Plan, the growth rate of the educational system in Colombia started to go down considerably during this period. In primary education, the annual growth rate of enrolment was around 2.3% between 1971 and 1980 (Ramirez and Tellez 2006). In secondary

education, where the problems were more evident, the growth rates of enrolment only reached 7.8% between 1971 and 1980, and 2.8% between 1981 and 1990 (Ramirez and Tellez 2006). This setback in the expansion of the educational system could be partly explained by the institutional problems generated by the centralized system, by the low levels of coordination, by the deficient communication, and by the very limited participation of the community in the educational process (Departamento Nacional de Planeación 1978).

It should be highlighted that the primary education enrolment rate in the cities was 98% in 1982 (Departamento Nacional de Planeación 1982). This confirms the observation that an important evolution in the educational system took place in the fifties, sixties, and seventies, considering that access to education was practically secured for the urban population by the beginning of the eighties. However, the story was very different for the rural population. In many rural regions a high percentage of kids didn't have access to educational facilities or any other options to receive a formal education. Specifically, according to a document from the Colombian Development Plan, the primary education enrolment rate for rural students was only 67% in 1982. It is also important to highlight that dropout rates were very high at the time. According to a document from the planning agency of Colombia, amongst the urban population, only 50% of the population that had enrolled in 1976 had finished primary education by 1979 (Departamento Nacional de Planeación 1982). Amongst the rural population, the percentage goes down to 17%. Additionally, according to that same document, in 1980 38% of the teachers did not have the teaching qualifications that were necessary to adequately teach their classes (Departamento Nacional de Planeación 1982).

It is also important to analyse the statistics of secondary education at that time. According to the Colombian National Planning Agency (*DNP*), the enrolment rate for secondary school (students 13 to 18 years old) was around 44% in 1982. Just like in the case of primary school, there are a high number of dropouts in secondary school. Specifically, only 54% of the students that enrolled in 1975 finished secondary school by 1980 (Departamento Nacional de Planeación 1982). The low quality of instruction at the secondary education level was also a big concern during this time period. According to the Colombian National Planning Agency (*DNP*), in 1980 46% of secondary school teachers did not have the teaching qualifications necessary for adequate schooling.

In terms of analphabetism, the situation improved considerably if the statistics of the fifties are compared to those of the eighties. More specifically, analphabetism rates amongst the population older than 10 years old went from 37.7% in 1951 to 12.2% in 1985 (Departamento Nacional de Planeación 1986).

Table 3						
Primary School Education in Colombia						
	Students		Schools		Teachers	
Period	Average Value	Average Annual Growth	Average Value	Average Annual Growth	Average Value	Average Annual Growth
1903-1910	194,811	12.8	3,126	12.5		
1911-1920	320,520	5.4	5,078	4.4	5,417	
1921-1930	438,856	4.1	6,828	3	8,646	3.3
1931-1940	575,118	1.2	8,499	1.7	11,627	3.4
1941-1950	716,266	3	10,919	2.9	16,511	4.2
1951-1960	1,266,611	7.7	15,922	5	32,309	8.8
1961-1970	2,450,884	6.9	24,073	3.4	66,917	6.8
1971-1980	3,934,603	2.3	31,356	2.2	122,468	4.5
1981-1990	4,079,153	0.4	36,433	1.5	134,592	0.6
1991-2000	4,681,813	2.2	51,760	4.4	185,272	3.8

Source: Ramirez and Tellez (2006)

The infrastructure available for educational purposes in the eighties was very precarious. Available information from the year 1984 about public schools infrastructure shows that out of a total of eighty three thousand classrooms available for primary education, only forty per cent was in good conditions (Departamento Nacional de Planeación 1986). This was indeed a very important restriction. The story is even worst for secondary education. At that level of instruction, out of a total of 17,100 available classrooms, 65% had problems.

However, it is important to highlight the important progress that was made in terms of teacher qualifications at the time. The educational level of public teachers improved significantly in the period that went from 1950 to 1990. In 1950, two thirds of the teachers only had primary education and one third had some training for teaching (*normalista*). In 1989, 68% of the teachers had some training for teaching (*normalista*), 14% had a high school title, 3% had a technical degree, and 15% had a university degree (Departamento Nacional de Planeación 1990). According to the Colombian National Planning Agency (DNP), the qualifications of the teachers improved in part because the requirements necessary to advance in the salary scale became more demanding in the eighties (Departamento Nacional de Planeación 1990). However, even though it is very clear that this indicator had

an important improvement in the period that is being analysed, these percentages are still indicative of the poor educational levels of school teachers.

Table 4						
Secondary School Education in Colombia						
	Students		Schools		Teachers	
Period	Average Value	Average Annual Growth	Average Value	Average Annual Growth	Average Value	Average Annual Growth
1903-1910	20,574	6.6	280	6.05		
1911-1920	25,632	5.68	307	3.03		
1921-1930	27,177	2.99	323	4.22		
1931-1940	49,257	5.61	648	3.02	4,871	5.16
1941-1950	63,293	4.11	733	1.28	6,677	2.78
1951-1960	154,685	12.37	1,319	10.89	13,976	11.65
1961-1970	498,332	12.95	2,759	7.54	33,960	8.49
1971-1980	1,352,712	7.82	4,091	1.38	70,220	5.83
1981-1990	1,901,860	2.8	5,523	5.74	90,255	1.67
1991-2000	3,074,631	4.67	10,126	6.93	158,497	6.81

Source: Ramirez and Tellez (2006)

Public expenditure in education as a percentage of GDP has grown considerably since the mid-fifties, going from around 1% in the sixties, to 3% in the seventies and to around 3.5% in the eighties (Ramirez and Tellez 2006). In the first half of the eighties expenditure in education grew significantly, going from 3.12% in 1980 to 3.63% in 1984. However, this percentage started to go down and by 1988 it reached 3.18% (Departamento Nacional de Planeación 2002). According to the Colombian National Planning Agency (DNP), compared to international standards, public expenditure in education as a percentage of GDP in Colombia (3.2%) was still very low during that time period. It was below the Latin American and Caribbean average (4.3%) and the developing countries average (4.2%) (Departamento Nacional de Planeación 1986).

At the end of the eighties and beginning of the nineties, problems associated with the supply of educational services became evident. According to the Ministry of Education, the education system at the time was not capable of supplying adequately the educational services demanded by the population although public expenditure in education had gone up from 1% of GDP in 1951 to around 3.5% in the eighties (Ministerio de Educación 2010).

The indicators discussed above gave clear signs that new changes in the educational system were necessary. To make those changes, a decentralization strategy was implemented. Decentralization efforts became evident at the end of the eighties. Law 24 of 1988 and Law 29 of 1989 established the bases to decentralize education at the municipal level. According to the Ministry of Education, these laws established that many functions should be passed on from the Ministry of Education to the municipalities, but the system should still be governed by nationwide norms and policies (Ministerio de Educación 2010). For instance, Law 24 of 1988 established a new order to help solve some of the administrative problems of the sector, and give up the over-centralized and disorganized system that was in charge of the Colombian public education at the time (Ramirez and Tellez 2006).

The Nineties: Time for Reform

Starting in the nineties, education in Colombia started gaining even more importance in the public policy debate thanks to the strengthening of the decentralization process that came about as a result of the new Constitution of 1991 and the establishment of important decrees related to the rights to education of the population. The Constitution of 1991 changed the centralized organization of educational administration that was instituted in the 1886 Constitution. It also established that education is a right for every Colombian citizen. More specifically, the new constitution established that education is a public service with a social function, and it is mandatory for kids 5 to 15 years old. That was a very important change considering that the 5 years of mandatory education that were established in 1936 are inadequate to prepare students for life (Ministerio de Educación 2010). According to Iregui, Melo and Ramos (2006), the process of political and administrative decentralization was strengthened, and the resources available for education and health care gradually increased thanks to the changes introduced in the Constitution of 1991 and Law 60 of 1993. A number of important changes were introduced: a new legal framework was established to calculate the monetary transfers from the Government to the regions and the criteria for the distribution of these resources, and from the regional and local point of view, the departments and municipalities began to play a more active role in the administration and execution of these resources (Iregui, Melo and Ramos 2006). According to the Ministry of Education, Law 60 of 1993 was enacted to distribute the legal and administrative competencies of the national government and the departmental and municipal governments, and to establish clear allocation rules within the territorial entities (Ministerio de Educación 2010). More specifically, this law determined that the Municipalities, the Departments and the Districts were in charge of the administrative tasks within the educational system, and the national government

was responsible for the basic costs of the system through transfers (*situado fiscal*) (Ministerio de Educación 2010).

The Constitution of 1991 also established how education should be financed. According to the Constitution, “the Nation and the territorial entities should participate in the direction, financing and administration of public education services, in the terms determined by the Constitution and the law” (Constitución Política 1991). It stated that a percentage of the current revenue of the State has to be delivered to the territorial entities to finance the educational and health needs of the population. This percentage has to be distributed “proportionally according to the number of current users and potential users of the services, taking into account the fiscal effort and administrative efficiency of the respective territorial entity” (Constitución Política 1991). According to Iregui, Melo and Ramos (2006), as a result of this process the resources allocated to education increased gradually and significantly, going from 3.3% of GDP in 1994 to 5.1% of GDP in 2004.

Table 5		
National Government Expenditure in Education (1990-2004)		
Year	National Government Expenditure in Education as a Percentage of GDP	As a Percentage of Total National Government Expenditure
1990	2.5	15
1992	3.4	16
1994	3.3	16
1996	4	18.8
1998	4.5	22.2
2000	4.2	25.5
2004	5.1	

Source: Ministerio de Educación (2001)

According to the Colombian National Planning Agency (DNP), the enrolment rate for primary education grew steadily at an annual rate of 2% since 1985 up to the late nineties. On the other hand, the enrolment rate for secondary education grew at a rate of 4.3% in 1993 and 5.1% in 1997 (Departamento Nacional de Planeación 1998). Additionally, the gross coverage rate for primary school in 1985 was of 105% and in 1997 of 113%, an annual growth rate of 0.7% (Departamento Nacional de Planeación 1998). A gross coverage rate that exceeds a 100% implies that the educational system has the resources (teachers, classrooms, etc.) to satisfy the needs of more children than those that are in school age. According to the Colombian National Planning Agency (DNP), for secondary school the gross coverage rate was 52.36% in 1985 and 80.42% in 1997, an annual growth rate of

2.34%. This means that by 1997, 20% of the resources were still missing to cover all the kids that were at an age to attend secondary school (Departamento Nacional de Planeación 1998).

At the beginning of 1994, another law was approved to regulate the provision of educational services according to the principles that were set in the Constitution of 1991. This new law was called the “General Law of Education”. According to the Ministry of Education, this new law completely re-organized the educational system (Ministerio de Educación 2010). Following the mandate of the “General Law of Education”, between 1995 and 1997, the National Ministry of Education advanced in the process of decentralization and gradually passed on the administration of the schooling system to the territorial entities. To verify that these territorial entities were capable of managing the school system, and for decentralization to be successful, a system to certify the capacity of the local education offices was established in the Decree 2886 of 1994 (Ministerio de Educación 2010). After this law was approved schools had more flexibility to determine their own curriculum. For instance, schools were now allowed to independently establish their own institutional education project. Schools started to have more autonomy and had the ability to innovate as part of the educational process (Ministerio de Educación 2010).

However, not everything worked out as planned. According to the Ministry of Education, the unstable fiscal situation of Colombia led to a great volatility in the monetary transferences for education. There were years when transferences grew rapidly, and therefore new teachers were hired in the different regions of Colombia, increasing the labour obligations of the local governments. In other years, transferences grew much less and these were not enough to cover the automatic increases in the costs of labour. In those years, the national government had to intervene to cover the deficit. In 1998, Colombia went through one of the worst economic crisis in its history. The current revenue of the government went down, and this led to an immediate decrease in the transferences. This was especially troublesome considering that the regional entities had increased significantly their expenditure during recent years (Ministerio de Educación 2010).

Even More Changes Came with the New Millenium

According to the Ministry of Education, in order to overcome the structural problem of governmental expenditures, in which education was a central component, the Constitution was amended in 2001 (Ministerio de Educación 2010). Law 715 of 2001 was enacted to incorporate those changes. The administration of the monetary transferences to finance the health care system and education was re-

organized in the General System of Participations (*Sistema General de Participaciones*). This system defined indicators to allocate the budget following the idea that resources should be distributed according to the population that is served and should be served, urban population and rural population, administrative and fiscal efficiency, and equity (Ministerio de Educación 2010). Additionally, Law 716 elucidated the competencies of the territorial entities (Departments, Municipalities and Districts) and the nation. By identifying the entities in charge of the different duties, the process of monitoring the results of the different policies was simplified (Departamento Nacional de Planeación 2002).

An important initiative was also carried out at the time in order to improve the quality of the teachers and the benefits of the teaching career. The Decree 1278 of 2002 was enacted to establish a new statute to professionalize the teaching career. This decree redefines the criteria to enter, to get a promotion and to receive retirement benefits. Additionally, a new structure was set up to determine the salary scale. This new structure was designed so that promotions are linked to performance and not only to time served. It is also important to highlight that this decree establishes that the renovation of a teaching contract depends exclusively of periodic evaluations of the performance of the teacher (Iregui, Melo and Ramos 2006).

Some educational variables had improvements in the period going from the year 1992 to the year 2005. The mean years of schooling went up 1.5 years during this time period, going from 6.4 years to 7.9 years (Departamento Nacional de Planeación 2007). This is a considerable improvement for a ten years period. However, as it is shown in table 6, unfortunately some important differences persisted between the mean years of schooling in the main cities and the other areas of Colombia. More specifically, the gap between the main cities and the other areas was relatively persistent at around 4 years of schooling (Departamento Nacional de Planeación 2007).

The literacy indicators also show an improvement during these years. Table 7 shows that the rates of analphabetism for the population 15 years and older went down from the year 1992 to the year 2005, going from 9.3% to 6.7%, respectively (Departamento Nacional de Planeación 2007). In the case of analphabetism, the gap between the rural and the urban population is also evident. For instance, the rates of analphabetism in 1992 for the rural population and the urban population were 18.9% and 5.7%, respectively. In the year 2005, the considerable differences between these two zones persisted (Departamento Nacional de Planeación 2007). In that year, the rates of analphabetism for the rural population and the urban population were reported to be 13.8% and 4.3%, respectively (Departamento Nacional de Planeación 2007).

Table 6					
Mean Years of Schooling for Population 15 Years and Older					
Year	Main City	Other Areas	Male	Female	Total
1992	7.4	3.8	6.4	6.4	6.4
1993	7.6	3.8	6.6	6.6	6.6
1994	7.8	3.9	6.7	6.8	6.8
1995	7.9	4.1	6.8	6.9	6.8
1996	7.8	3.8	6.7	6.7	6.7
1997	8	3.9	6.9	7	7
1998	8.1	3.8	7	7	7
1999	8.1	4.2	7.1	7.2	7.1
2000	8.3	4.4	7.3	7.3	7.3
2001	8.4	4.5	7.3	7.5	7.4
2002	8.4	4.7	7.4	7.5	7.5
2003	8.5	4.9	7.5	7.6	7.6
2004	8.6	5.1	7.7	7.8	7.7
2005	8.9	5	7.8	7.9	7.9

Source: Departamento Nacional de Planeación (2007)

Table 7			
Analphabetism Rate as a Percentage			
Year	Urban	Rural	Total
1992	5.7	18.9	9.3
1993	4.7	18.4	8.4
1994	4.3	17.8	7.9
1995	4.5	16.8	7.8
1996	5	20	9
1997	4.7	19.9	8.6
1998	4.9	20	8.7
1999	5	18.6	8.5
2000	4.9	18	8.3
2001	4.5	16.6	7.5
2002	5.2	16.2	7.9
2003	5.1	15.4	7.6
2004	4.8	14.2	7.1
2005	4.3	13.8	6.7

Source: Departamento Nacional de Planeación (2007)

In recent years, there has been important progress in the Colombian educational system. According to the Ministry of Education, 41000 new places for students attending pre-school to secondary school were created in the year 2011. Additionally, in that same year, 170000 adults learned how to read. This led to a reduction in the analphabetism rate, going down from 4.89% in 2010 to 4.53% in 2011 (Ministerio de Educación 2012). Also contributing to the evolution of the system, in the year 2012 the Ministry of Education established that all students attending official schools did not have to pay tuition for pre-school up to secondary school. The idea behind this policy is that all students get access to education, even if they do not have financial means to cover educational expenses (Ministerio de Educación 2012).

Concluding Remarks

As it was highlighted throughout this chapter, the Colombian educational system has had important improvements in the last fifty years. However, there are still many challenges in the years to come for the policy makers in charge of the Colombian educational policy. Some of these challenges are associated to the civil conflict that is still affecting the Colombian population. The objective of this thesis is precisely to get a better understanding of how the conflict has affected educational outcomes in order to take adequate policy decisions.

WORKS CITED

- Constitución Política de Colombia (1991). *Ministerio de Justicia*. República de Colombia.
- Departamento Nacional de Planeación (1961). *Desarrollo Económico y Social*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1969). *Planes y programas de desarrollo: 1969-1972*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1970). *Las Cuatro Estrategias: 1970-1974*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1974). *Para Romper la Brecha: 1974-1978*. Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1975). *Para Cerrar la Brecha: Plan de Desarrollo Social, Económico y Regional, 1975-1978*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1978). *Plan de Integración Social: 1978-1982*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1982). *Cambio con Equidad: 1982-1986*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1986). *Plan de Economía Social: 1986-1990*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1990). *La Revolución Pacífica: 1990-1994*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1994). *El Salto Social: 1994-1998*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (1998). *Cambio para Construir Paz: 1998-2002*. República de Colombia: Departamento Nacional de Planeación.
- Departamento Nacional de Planeación (2002). *Hacia un Estado Comunitario: 2002-2006*. República de Colombia: Departamento Nacional de Planeación.

- Departamento Nacional de Planeación (2007). La Educación en Cifras. In Dirección de Desarrollo Social (Ed.), *Sistema de Indicadores Sociodemográficos de Colombia*. Bogotá, Colombia: Departamento Nacional de Planeación.
- Ministerio de Educación (2001). Informe Nacional Sobre el Desarrollo de la Educación en Colombia. Bogotá, Colombia: Ministerio de Educación Nacional.
- Ministerio de Educación (2004). El Desarrollo de la Educación en el Siglo XXI - Informe Nacional de Colombia. In O. A. d. Planeación (Ed.). Bogotá, Colombia: Ministerio de Educación.
- Ministerio de Educación (2010). Revolución Educativa 2002 - 2010 Acciones y Lecciones. Bogotá, Colombia: Ministerio de Educación.
- Ministerio de Educación (2012). Rendición de Cuentas Ministerio de Educación. Bogotá, Colombia: Ministerio de Educación.
- Helg, A. (1987). La Educación en Colombia 1918–1957. *Una historia social, económica y política – Fondo de Cultura Económica*.
- Iregui, A. M., Melo, L., and Ramos, J. (2006). La educación en Colombia: análisis del marco normativo y de Los indicadores sectoriales. *Revista de economía del Rosario*, 9(2).
- Ramírez, M. T., and Téllez, J. P. (2006). La educación primaria y secundaria en Colombia en el siglo XX: Banco de la República.

CHAPTER 4

LITERATURE REVIEW

This chapter presents a review of the literature about conflict and education that is available internationally. The aim of this chapter is to justify the importance of the research topic explored in this dissertation given the literature that is currently available. Another objective of this chapter is to highlight the gap in the literature that this book attempts to fill.

It should be acknowledged that this literature review focuses on the economic literature that is currently available. However, there are other disciplines that have also studied the relationship between civil conflict and education, including Sociology and Political Science. Nevertheless, the economic literature has a stronger quantitative tradition. For our purposes, and considering the quantitative nature of the methods applied in the analyses throughout the dissertation, we have decided to center our attention on the economic literature that is currently available.

This chapter is divided into two different sections in order to cover the different themes studied in this thesis: conflict and education internationally, and conflict and education in Colombia.

2. Conflict and Education

A number of researchers have contributed to the literature that studies the relationship between civil conflict and education. What is notable is that, to the best of my knowledge, only one effort has been made to try to understand and quantify the impact of civil conflict on education achievement (measured by standardized exams). Such effort, which will be documented in section B of this review of the literature, was made by Rodríguez and Sánchez (2010). Yet, this area of research is much understudied, opening an important gap in the literature. The objective of this doctorate thesis is precisely to fill that gap in the literature. In what follows, I will review the existing research concerning education and civil conflict in order to contextualize the topic under study and justify the relevance of this dissertation. The academic studies mentioned in this review will be grouped according to themes in common.

Different studies have analysed the relationship between conflict, school enrolment and completion. Swee (2008) uses data on war casualties of the Bosnian

War 1992–1995 at the municipality level, and exploits the variation in war intensity and birth cohorts of children, to identify the effects of the civil war on schooling attainment. The empirical strategy exploits the variation in birth cohorts and in war intensity represented by the number of war casualties per capita across Bosnian municipalities to adopt a difference in differences approach that takes into account unobserved pre-war differences across municipalities. The author finds that individuals in the cohorts affected by war are less likely to complete secondary schooling, if they resided in municipalities that experienced higher levels of war intensity, but there are no significant effects of war on the completion of primary schooling. Using these results, Swee (2008) argues that the Bosnian war influenced school attainment through youth soldiering and no other direct mechanisms like exodus of teachers and damaged or destroyed infrastructure. The author discusses different mechanisms through which civil war might affect schooling attainment given the availability of data on individuals' physical and mental health, war damage to housing units, and out-migration during the war. In another study, Akresh and de Walque (2008) examine the impact of the Rwandan genocide on children's human capital investment. The authors use two cross-sectional household surveys collected before and after the Rwandan genocide in order to examine if and how genocide affect children's school enrolment and the probability of completing a particular grade (Akresh and de Walque 2008). To do so, the authors compare children in the same age group who were and were not exposed to genocide. Akresh and de Walque (2008) point out that civil war, and more specifically genocide in this particular case, are among the most destructive of social phenomena especially for children of school-going age. The authors find that children exposed to the genocide experienced a drop in educational achievement of almost one-half year of completed schooling, and are 15 percentage points less likely to complete third or fourth grade (Akresh and de Walque 2008). In that sense the results obtained by Akresh and de Walque (2008) contradict the findings of Swee (2008). The results obtained by Swee (2008) for the Bosnian case show no significant effect on the completion of primary schooling whereas Akresh's and de Walque's (2008) results for the Rwandan genocide show significant negative effects of civil conflict on completing schooling at the primary level.

Supporting the results of Swee (2008), Bellows and Miguel (2006) find no significant effects of civil war on enrollment. These two researchers study the aftermath of the civil conflict in Sierra Leone that started in 1991 and lasted until 2002. Using governmental data from the Institutional Reform and Capacity Building Project (IRCBP), the authors study local socioeconomic and institutional outcomes in 2004 and 2005. As mentioned above, Bellows and Miguel (2006) find that there are no meaningful effects of the war on 2004 and 2005 school enrolment. The authors highlight that areas that suffered from more violence have somewhat

better post-war outcomes, but these effects are not significant. In fact, Bellows and Miguel (2006) point out that following their results there is no evidence of persistent adverse effects of civil war violence on local institutions. However, the impact of the civil conflict on aggregated living standards could be very negative even if the results show positive localized effects in some institutional dimensions. On the other hand, the results obtained by Lai and Thyne (2007) support Akresh's and de Walque's (2008) findings. Lai and Thyne (2007) study the effects of civil wars and the post-civil war environment on education expenditure and enrolment. Through the use of two causal mechanisms and UNESCO education data, these two authors explore those effects. The causal mechanisms explored are the loss of infrastructure and personnel, and the drawing away of funds for increased military expenditures to fight the civil war (Lai and Thyne 2007). The results of this empirical research show evidence of the destructive effects of civil war on a system of education, exposed by the declines in expenditure and enrolment rates during periods of civil war.

It should be emphasized that the studies by Swee (2008), Bellows and Miguel (2006), Lai and Thyne (2007), and Akresh and de Walque (2008) show that there is no consensus in the existing literature about the relationship between educational enrollment and civil conflict. These results could be indicative of the possibility that the characteristics and effects of a civil conflict vary from case to case. Such possibility justifies the need to carry out more research in order to better understand the relationship between education and civil conflict for the case of the countries affected by these circumstances.

Previous efforts, including those of Dabalen and Paul (2012), Bundervoet (2012) and Merrouche (2006), have attempted to explore the relationship between conflict and years of education exploring different identification strategies. For instance, Dabalen and Paul (2012) estimate the causal effect of civil war on years of education in a cohort of students that were exposed to armed conflict in Cote d'Ivoire. The authors calculate the average causal effect of civil war on education using a household survey collected in 2008 and data on local incidences of conflict from the Armed Conflict Location and Event Database (ACLED). Dabalen and Paul (2012) use two different identification strategies. In the first strategy, the year and department of birth is used to determine an individual's exposure to war and a difference in differences approach to quantify differences in the years of education for kids that live in war affected regions. The second strategy uses victimization indicators to implement the propensity score matching technique, in an effort to try to minimize selection bias and confounding in the causal effect. The results robustly show that war has a detrimental impact on education.

In line with Dabalen and Paul (2012), Bundervoet (2012) also studies the relationship between conflict and years of education. However in this case, the researcher examines the impact of, war induced, poor childhood health on educational attainment in early adolescence through the use of a panel of children from Burundi. The question that this paper seeks to answer is whether the health effect of the war has had consequences for schooling later in life. The author finds that children that were malnourished at the baseline had on average less years of education than children that are their same age but were healthier at the baseline. Bundervoet (2012) highlights that this relationship is stronger amongst older kids i.e. those that were exposed for a longer period of time to conflict. The observed results could be explained by an enrolment effect (those children who were malnourished at the baseline were less likely to ever enroll in school) and a grade repeating or early drop out effect. The findings of Bundervoet (2012) show a micro-level pathway between war and poor longer-run economic performance, i.e. if exposure to war in early childhood leads to ill health, ill health in early childhood leads to poor school performance, then –insofar as low educational attainment is correlated with worse outcomes later in life- exposure to war in early childhood will have lifelong adverse economic consequences. Another researcher has also studied the relationship between conflict and years of education exploring a different identification strategy. Merrouche (2006) uses instrumental variable regressions and a difference in differences approach to identify the effects of landmine contamination on years of schooling using data from Cambodia. This researcher found that exposure to landmines in Cambodia resulted in an average loss of 0.4 years of education. The author claims that there are three possible factors that could explain this finding: the downgrading of educated people during the regime, direct effects of landmines on the returns to education, and the destruction of physical capital and technological delay. Although Merrouche (2006), Dabalen and Paul (2012) and Bundervoet (2012) use different empirical strategies, the results obtained lead to similar conclusions. Following the results of these researchers, it seems to be the case that conflict indeed has a negative effect on years of schooling.

Some researchers, including Annan et al. (2011), Kecmanovic (2012), Valente (2011), Shemyakina (2011), and Justino (2011), have focused their work on the relationship between education, conflict and gender. Annan et al. (2011) study the gender-based impacts of war using a quasi-experiment carried out in Northern Uganda. These researchers find that abduction generates significant adverse effects in human capital for males, but not for females. In the specific case of males, they receive 0.51 less years of education for every year of abduction. According to Annan et al. (2011), this difference could be explained by the fact that for most females the alternative to abduction is still very unfortunate: lack of employment opportunities, low educational investment and in some cases no educational

opportunities at all. However, the authors observe that females who are forced to bore children during their abduction have more than a year less education than other females, abducted or not abducted. Even though educational opportunities are limited for all females, forced motherhood is also associated with a lower probability of returning to school (Annan et al. 2011). Also for Uganda, Annan and Blattman (2010) find similar results for a group of children. According to these researchers, abducted male youth receive 0.75 fewer years of education. This result is particularly important given the fact that lower levels of education lead to a lack of transferable skills, a deficiency in human capital, and ultimately a decrease in lifetime earnings ability (Annan and Blattman 2010).

Valente (2011), Shemyakina (2011), and Justino (2011) have also explored the impact of armed conflict on school enrolling and schooling accumulation, analysing the heterogeneous impact by gender. Shemyakina (2011) studies the specific case of Tajikistan because this country, located in the Former Soviet Union region, was involved in a devastating civil war from 1992 to 1998. Shemyakina (2011) uses the difference in differences technique to determine whether the exposure to conflict affected the probability of completion of mandatory schooling. The results show that exposure to the conflict had a significant negative effect on the school enrolment of girls, and little, or no, effect on enrolment of boys (Shemyakina 2011). It is worth noting that the results obtained by Shemyakina are totally different to those obtained by Annan et al. (2011) and Annan and Blattman (2010). However, there is also research available to support the findings of Shemyakina (2011). Valente (2011) considers the case of Nepal, a country that experienced a civil conflict between 1996 and 2006 due to a Maoist insurgency (Valente 2011). In order to estimate the impact of conflict intensity on education and marriage in Nepal, the author exploits within and between district variations in the intensity of violence. Valente (2011) shows that conflict intensity had a small positive effect on female educational attainment, whereas male schooling was not significantly affected by the conflict measures. Subsequent work by Shemyakina, and her co-author Singh (2013), also ratify the findings of Shemyakina (2011). Singh and Shemyakina (2013) explore the gender specific long-run effects of the Punjab insurgency on educational attainment of adults who were between ages 6-16 years at the time of insurgency. They find a substantial and statistically significant effect of terrorism and terrorists' killings of civilians on educational attainment by girls; for one standard deviation increase in the number of killings due to terrorism and in the number of terrorist incidents between 1981 and 1993, women from the affected cohort attained about 0.69 and 0.60 fewer years of education than comparable women that were not exposed to conflict. These numbers are substantial considering that women in Punjab have only 4.83 years of education on average. Singh and Shemyakina (2013) highlight that these results could be explained by a decrease in the educational expenses by households that

had a greater share of girls to boys of school-going age due to insurgency. This implies that given the constrained amount of resources available during conflict, parents may prefer investing in the education of sons, while investments in daughters are reduced.

Nevertheless, the debate continues as additional research by Justino (2011) find evidence to support the findings by Annan et al. (2011) and Annan and Blattman (2010) in the long run. Justino (2011) study the short term impact of conflict in Timor Leste on the level and access to education of boys and girls. Exploiting the variation in the number of killings over time and across districts to identify conflict affected individuals, the authors calculate the impact of the 1999 violence wave on school attendance and grade deficit rates in 2001 and the longer term impact of the conflict on primary school completion of cohorts of children observed in 2007 (Justino 2011). Justino (2011) find that the short-term effects of the conflict are mixed, but in the longer term there is a strong negative impact of the conflict on primary school completion among boys of school age exposed to a violence wave. It is also possible to find additional support for the findings by Annan et al. (2011) and Annan and Blattman (2010) in Kecmanovic (2012). This researcher studies the effect that the war in Croatia from 1991 to 1995, had on the educational, employment and income possibilities of the 1971 male birth cohort. The 1971 birth cohort was chosen because it is very likely that an important number of men from this cohort were drafted into the armed forces. The author uses the occurrence of war and the subsequent effects as a natural experiment, and applies the difference in differences technique to find the impact of conflict on education. Kecmanovic (2012) finds lower levels of education amongst the cohort of males affected by war in Croatia compared to women and older cohorts. The author explains this finding by pointing out the importance of males in the conflict as soldiers.

In sum, all the empirical studies analysing the relationship between education, conflict and gender that were reviewed suggest that there is a significant gender effect in education associated with conflict. However, as it was pointed out, the results obtained by the different researchers studying this topic are not consistent. The lack of consistency casts doubts about the results obtained to date, and motivates further research in this topic in order to explore the relationship in more detail.

The empirical studies reviewed above focus on evidence from developing countries. Nevertheless, it is also possible to find empirical research focusing on developed countries. However, it is important to point out that available research for developed countries seeks to specifically understand the effects of World Wars on education and labour market outcomes. Ichino and Winter-Ebmer (2004) and Akbulut-Yuksel (2014) provide evidence from a developed country, Germany, to

explore the link between conflict and education. Akbulut-Yuksel (2014), using a methodology that is similar to the one employed by Akresh and de Walque (2008), provides causal evidence on the long-term consequences of large-scale physical destruction on the educational attainment, health status and labour market outcomes of German children after World War II. On the other hand, using WWII as an instrumental variable, Ichino and Winter-Ebmer (2004) provide evidence on the causal effect of education on earnings by comparing Austria and Germany, where the civilian population was severely affected, with Sweden and Switzerland, where civilians were not directly affected by conflict. Both papers find that Germans who were in the schooling cohorts during World War II received less education than their counterparts. Akbulut-Yuksel (2014) shows that German children that were school-aged during WWII had 0.4 fewer years of education in adulthood. The author highlights that an important mechanism to explain this impact was the destruction of schools and the absence of teachers. Ichino and Winter-Ebmer (2004) find that individuals who were 10 years old during or immediately after WWII acquire less education and earned a lower salary in adulthood compared to other cohorts within Germany and Austria, and compared to children of the same cohort born in non-war countries. The authors argue that exposure to war affects individuals only through deteriorating their human capital. It is worth noting that the results from Akbulut-Yuksel (2014) and Ichino and Winter-Ebmer (2004) are not consistent with the results obtained by Swee (2008) and Bellows and Miguel (2006), which were discussed previously, regarding the relationship between conflict and educational attainment in developing countries. In contrast with Akbulut-Yuksel (2014) and Ichino and Winter-Ebmer (2004), Swee (2008) and Bellows and Miguel (2006) find no significant effect of violent conflict on schooling. These results suggest that the characteristics and effects of a conflict vary from case to case. This justifies the need to carry out more research in order to better understand the relationship between education and conflict for different cases around the world.

The articles mentioned above attempt to capture the state of the art in the academic research of the relationship between conflict and education. Nevertheless, the enthusiastic reader could complement this review of the literature by referring to other remarkable efforts that have made by other researchers. Justino (2010) makes an important contribution to the literature by reviewing and analysing available empirical research on the impact of violent conflict on the level and access to education. In her paper, Justino (2010) highlights that the lack of empirical studies limits the understanding of the current educational situation in conflict zones. In that sense, Justino's work suggests that there is room for new research in this area. Justino's statement further justifies the relevance of this dissertation.

As reviewed in this section, a number of researchers have contributed to the literature that studies the relationship between civil conflict and education. However, there is still a lot of ground to cover in this area of research. As it was highlighted throughout this review of the literature, there are many contradictory results in the existing literature that justify the need to continue doing research in those topics that are already under research. Additionally, there are many topics that are understudied and could potentially play a very important role in policy analysis. The objective of this doctorate thesis is precisely to study one of those understudied topics: the relationship between conflict and educational achievement (measured by standardized exams).

3. Conflict and Education in Colombia

There are some articles that have explored the relationship between education and conflict using data from Colombia. Just like the studies reviewed previously, these articles focus on the relationship between school enrolment, desertion and conflict, but do not try to quantify the effects of civil conflict on academic achievement (measured by standardized tests). The only exception is Rodríguez and Sánchez's (2010) working paper. In fact, these two researchers highlight that no study in the literature has analysed how armed conflict might influence schooling achievement in standardized exams of students attending schools in conflict regions. Rodríguez and Sánchez (2010) investigate the impact of armed conflict on the quality of high school education, measured through standardized exams. The authors use an instrumental variable approach to control for possible measurement errors and problems of endogeneity. Rodríguez and Sánchez (2010) instrument the number of attacks in a municipality with the homicide capture rate at the state level. The authors find that an increase in one standard deviation on the average attacks by armed groups to which students are exposed during secondary education decreases the ICFES test score in 0.46 standard deviations. Rodríguez and Sánchez (2010) argue that achievement in conflict zones is lower because these regions attract less qualified teachers.

The working paper of Rodríguez and Sánchez (2010) is pioneer in analysing the effects of conflict on schooling achievement in standardized exams, but there is still a lot of ground to cover in this area of research. There are different methodologies and approaches that can be explored to understand the relationship between conflict and the results in standardized exams. This dissertation will attempt to fill the existing gap in the literature by exploring some methodologies including multilevel/hierarchical analysis.

Several studies have studied the relationship between educational accumulation

and conflict for the Colombian case. Wharton and Oyelere (2012) analyse the existing gap between educational accumulation and enrolment created by the Colombian conflict. The authors show that children who live in a municipality with high levels of conflict have a gap in enrolment and accumulation, but that gap is much smaller than the gap of those children who are directly affected by the conflict because of displacement. The results of Wharton and Oyelere (2012) suggest that there is a significant education accumulation and enrolment gap for children of internally displaced families that can reach levels of over half a year in secondary school. In line with Wharton and Oyelere's (2012) work, Dueñas and Sanchez (2007) study the impact of violence on dropout rates focusing on households in the eastern part of Colombia. The authors develop a theoretical model to show how violence affects wages and education costs. Dueñas and Sanchez (2007) also use a duration model to show that the presence of illegal armed groups increases dropout rates. Dueñas and Sanchez (2007) highlight that these effects are higher for the poorest households. Following the work of Dueñas and Sanchez (2007), Rodriguez and Sanchez (2012) estimate the effect that exposure to armed conflict has on school drop-out decisions of Colombian children between the ages of six and seventeen. The principal objective of this article is to determine how armed conflict existent in the municipality where the child resides affects her/his schooling attainment. The authors measure civil conflict as the total number of offensive actions (attacks against infrastructure and the civil population and clashes with governmental forces) undertaken by the guerrillas, paramilitary groups or common criminality in each municipality and year. They use household and violence data sets to implement a duration analysis methodology. They find that armed conflict reduces the average years of schooling in 8.78% for a pooled sample of Colombian children, and it reduces the average years of schooling in 17.03% for children between sixteen and seventeen years old. Another piece of research that comes up with similar conclusions is the one written by Sánchez and Diaz (2005). In that article the two researchers study the social effects of illegal armed activities from 1990 to 2002. Using a matching mechanism, Sánchez and Diaz (2005) show that municipalities affected by the activities of illegally armed groups, had slower growth rates of primary and secondary school enrolment compared to those in municipalities that were not affected by illegal activities. Additionally, these two researchers show that the dropout rate in conflict-affected municipalities is around two per cent higher than in peaceful municipalities (control group).

It should be highlighted that the conclusions set forth by Dueñas and Sanchez (2007), Rodriguez and Sanchez (2012), Sánchez and Diaz (2005), and Wharton and Oyelere (2012) in the four articles surveyed above coincide in that there is a negative relationship between conflict and educational accumulation. Even though the researchers employ different empirical methods, they all come up with similar

results for the Colombian case.

The research that has been done in Colombia is not limited to empirical exercises. There is also research exploring the theoretical underpinnings of the relationship between conflict and education. Barrera and Ibañez (2004) develop a dynamic theoretical model that identifies three different mechanisms through which violence influences educational investment. The authors highlight that their paper is the first effort in the literature to develop a formal model of this relationship. The formulation of theoretical models to explain the relationship between conflict and education is a particularly understudied topic in the literature. The three mechanisms that Barrera and Ibañez (2004) identify are: a) violence can directly influence the utility of households, changing the consumption of education, b) violence many times destroys physical capital and generates uncertainty, reducing the levels of investment and production, and c) violent events can influence the returns on education and therefore affect the investment in education. The authors also present empirical evidence to test their theoretical model. In line with their theoretical model, Barrera and Ibañez (2004) find that the enrolment rates are lower in municipalities where homicide rates are above the national average and the likelihood of school enrolment decreases when homicide rates go up. These results predict a negative relationship between conflict and educational accumulation as the other studies that were reviewed for the Colombian case.

Other researchers have also studied the relationship between conflict and education but from a different perspective. These researchers have studied how the intensity of the conflict might be affected by the level and quality of education. For instance, Cortés and Montolio (2013) study the effect of the degree of publicness of goods on violent conflict in Colombia. The authors use three goods with different degrees of publicness to find the sign of the effect: coca crops, road density and average education quality. Cortés and Montolio (2013) use an instrumental variable approach to control for endogeneity because not only provision of goods may affect conflict but also conflict may affect goods provision. The authors find that the likelihood of suffering a guerrilla attack is reduced with higher education quality. Cortés and Montolio (2013) highlight that the likelihood of conflict decreases with the degree of publicness. The provision of public goods reduces the likelihood of conflict. It should be highlighted that the approach of the work of Cortés and Montolio (2013) differs greatly from the approach proposed for this thesis. This dissertation studies how conflict affects educational achievement, whereas Cortés and Montolio (2013) study how educational achievement/quality affects the intensity of conflict.

As it was exposed in this section of the literature review, there are some papers that have explored the relationship between education and conflict using data

from Colombia. By reviewing the literature it was possible to show that the existing articles focus on the relationship between school enrolment, desertion and conflict. The relationship between conflict and academic achievement in cognitive exams has been understudied, opening an important gap in the literature. Such gap is explained in large part by the limited availability and reliability of micro data about civil conflict events and standardized test results. In fact, only one effort by Rodríguez and Sánchez (2010) has been made to try to quantify the effects of civil conflict on academic achievement measured by standardized tests. This doctorate thesis contributes significantly to the literature because it introduces substantial improvements to Rodríguez and Sanchez's (2010) work in terms of methodology and data quality. Taking advantage of a unique database about civil conflict events in Colombia administered by a Colombian think tank (Resource Centre for Conflict Analysis - CERAC) and a government kept database (the Colombian Institute for the Evaluation of Education (ICFES) database) containing the results of the Colombian High School exit exams, the objective of this doctorate thesis is to get a better understanding of the relationship between civil conflict and educational achievement through the use of several econometric techniques.

4. Concluding Remarks

The review of the existing literature presented in this chapter confirms that there is still a lot of ground to cover in this area of research. In that sense, this review of the literature confirms the remarks of Benefield and Tomlinson (2005) regarding the lack of academic research studying the connection between conflict and education: "education in conflict and post-conflict situations, as a recognized field of research, is in its infancy." This dissertation will attempt to fill some of the existing gaps in this literature by exploring methodologies that have not been used to explore the understudied relationship between civil conflict and academic achievement measured by standardized tests.

WORKS CITED

- (CERAC), Centro de Recursos para el Análisis del Conflicto (2013). *Conflict data*.
- (ICFES), Instituto Colombiano para la Evaluación de la Educación (2010). Orientaciones para el Examen de Estado de la Educación Media (pp. 30). Bogotá: Instituto Colombiano para la Evaluación de la Educación (ICFES).
- Akbulut-Yuksel, M. (2014). Children of war: The long-run effects of large-scale physical destruction and warfare on children. *The Journal of Human Resources*. 49: 634-662.
- Akresh, R., and de Walque, D. (2008). Armed conflict and schooling: evidence from the 1994 Rwandan genocide.
- Angrist, J. D., and Kugler, A. D. (2008). Rural windfall or a new resource curse? Coca, income, and civil conflict in Colombia. *The review of economics and statistics*, 90(2), 191-215.
- Blattman, C., and Annan, J. (2010). The consequences of child soldiering. *The review of economics and statistics*, 92(4), 882-898.
- Annan, J., Blattman, C., Mazurana, D., and Carlson, K. (2011). Civil war, reintegration, and gender in Northern Uganda. *Journal of conflict resolution*, 55(6), 877-908.
- Barrera, F., and Ibáñez, A. M. (2004). Does Violence Reduce Investment in Education?: A Theoretical and Empirical Approach: UNIVERSIDAD DE LOS ANDES-CEDE.
- Bellows, J., and Miguel, E. (2006). War and institutions: New evidence from Sierra Leone. *The American economic review*, 96(2), 394-399.
- Benefield, P. a. Tomlinson, K. (2005). Education and Conflict: Research and Research Possibilities: National Foundation for Educational Research (NFER).
- Bundervoet, T. (2012). War, Health, and Educational Attainment: A Panel of Children during Burundi's Civil War: Households in Conflict Network.
- Cortés, D., and Montolio, D. (2013). Publicness of goods and violent conflict: Evidence from Colombia.

- Dabalen, A., and Paul, S. (2012). Estimating the causal effects of conflict on education in Cote d'Ivoire. *World Bank Policy Research Working Paper*(6077).
- Dueñas, L. F., and Fabio Sanchez. . (2007). *Does Armed Conflict Affect the Poorest?: The Case of School Dropout in Eastern Colombia*. . Documentos CEDE. Universidad de los Andes. Bogotá, Colombia.
- Ichino, A., and Winter-Ebmer, R. (2004). The long-run educational cost of World War II. *Journal of Labor Economics*, 22(1), 57-87.
- Justino, P. (2010). How Does Violent Conflict Impact on Individual Educational Outcomes? The Evidence So Far. *background paper for the Education For All Global Monitoring Report*.
- Justino, P. (2011). Violent Conflict and Human Capital Accumulation. *IDS Working Papers*, 2011(379), 1-17.
- Kecmanovic, M. (2012). The Short-run Effects of the Croatian War on Education, Employment, and Earnings. *Journal of Conflict Resolution*.
- Lai, B. and Thyne C. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, 44(3), 277-292. doi: 10.1177/0022343307076631
- Merrouche, O. (2006). The human capital cost of landmine contamination in Cambodia: Households in Conflict Network.
- Rodríguez, C., and Sánchez, F. (2010). Books and guns: the quality of schools in conflict zones. *Documento CEDE*(38).
- Rodriguez, C., and Sánchez, F. (2012). Armed conflict exposure, human capital investments, and child labor: evidence from Colombia. *Defence and peace economics*, 23(2), 161-184.
- Sánchez, F., and Díaz, A. M. (2005). *Los efectos del conflicto armado en el desarrollo social colombiano, 1990-2002*: CEDE, Centro de Estudios sobre Desarrollo Económico, Facultad de Economía, Universidad de los Andes.
- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, 95(2), 186-200. doi: 10.1016/j.jdevco.2010.05.002

- Singh, P., and Shemyakina, O. (2013). Gender-Differential Effects of Conflict on Education: The Case of the 1981-1993 Punjab Insurgency. *Available at SSRN* 2253382.
- Swee, E. L. (2009). On war and schooling attainment: The case of Bosnia and Herzegovina. *Households in Conflict Network Working Paper*, 57.
- Valente, C. (2011). What did the Maoists ever do for us ? education and marriage of women exposed to civil conflict in Nepal.
- Wald, N., and Bozzoli, C. (2011). Bullet Proof? Program Evaluation in Conflict Areas: Evidence from Rural Colombia.
- Wharton, K. a. U. O., Ruth. (2011). *Conflict and Its Impact on Educational Accumulation and Enrollment in Colombia: What We Can Learn from Recent IDPs*. IZA Discussion Paper. IZA. Retrieved from <http://ssrn.com/abstract=1922011>

CHAPTER 5

MODELING CONFLICT AND EDUCATIONAL ACHIEVEMENT

The academic literature related to education and cognitive skills that is currently available has documented that student's results on cognitive tests are correlated with educational attainment and income (Todd and Wolpin 2007). Those findings reveal the importance of understanding the determinants of students' achievement on cognitive tests. A useful assumption that is often made in order to understand the relationship between educational achievement and the determinants under study is that each household maximizes a utility function subject to constraints (Glewwe and Kremer 2006). The main arguments of that utility function are consumption of goods and services, and schooling. The constraints that are imposed to the utility function are the educational production function, the impact of schooling on future labor income, credit constraints, and a life-cycle budget constraint (Glewwe and Kremer 2006). This utility maximizing setup discloses the importance of the educational production function in the analysis of educational achievement.

Most of the studies done in the economics of education are modeled through the use of educational production functions. In using these functions, researchers are establishing an analogy between the learning process of an individual and the production process of a firm (Todd and Wolpin 2003). The existing literature exploring educational production functions examines the relationship between schooling inputs and test score outcomes in order to get a better understanding of the importance of given inputs in the educational process (Todd and Wolpin 2007). As Hanushek (2008) highlights, the underlying model that evolved as a result of the different research efforts that have been done throughout the years in this field is straightforward: the achievement of students, which is the output of the educational process, is directly related to some inputs that can be controlled directly by policy makers and other inputs which cannot be easily controlled by them. For example, policy makers can control inputs such as the characteristics of teachers and school infrastructure, but they cannot control the innate endowments and family related characteristics of the students (Hanushek 2008).

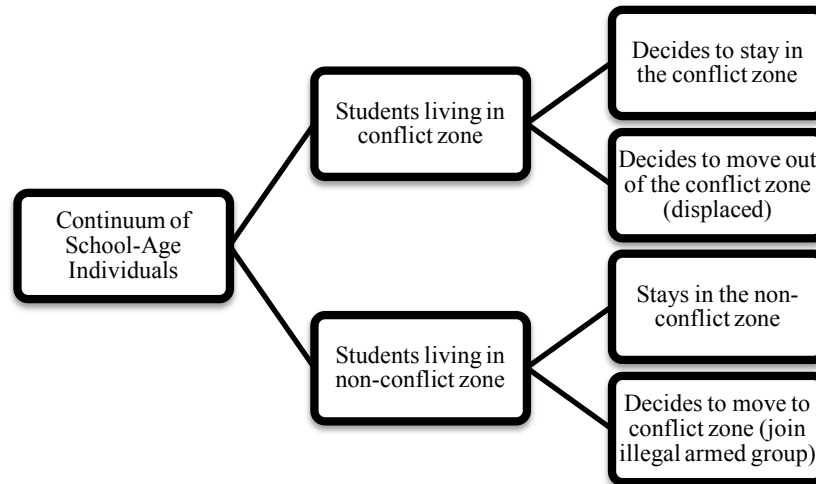
In this research, I will consider an extension of a simple model of achievement proposed by Todd and Wolpin (2003), in which the student's achievement level is modeled through the use of an educational production function. A key element that is taken into consideration in this model is that although achievement can be measured at discrete points in time, the educational process is cumulative

(Hanushek 2008). In other words, those inputs that have been used in the past affect the current levels of achievement.

2. The Educational Decision Making Process of School-Age Individuals

A sub-group of the population consists of a continuum of school-age individuals. These people either reside in conflict zones or in non-conflict zones. The families of school-age individuals living in conflict zones face an important decision: to stay in the conflict zone or to migrate to a non-conflict zone. Students living in a non-conflict might also move voluntarily or involuntarily to conflict zones in order to join illegal armed groups, therefore, dropping out of school. The basic decision making structure is portrayed in *diagram 1*.

Diagram 1

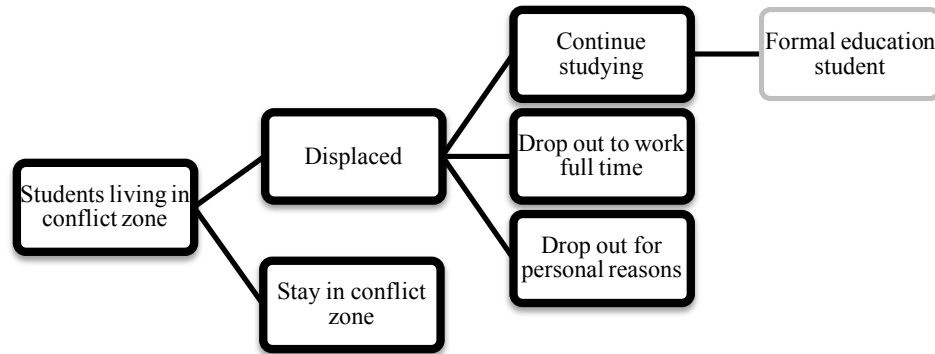


Source: author's elaboration

Those students that are displaced from their hometown often times face harsh conditions in their new place of residency, forcing them to decide whether to continue with their studies or go on to a different activity. More specifically, they have to decide whether to continue studying, to dropout for personal reasons or to find a job to contribute to the income of the household (*diagram 2*). In taking that decision, the individuals weigh the costs associated with a given level of education, the expected benefits from a higher level of education, the expected benefits from dropping out for other personal reasons (e.g. teenage pregnancy), and the expected income derived from finding a full time job at the new place of

residency. The decision that individuals take seeks to maximize their expected utility.

Diagram 2



Source: author's elaboration

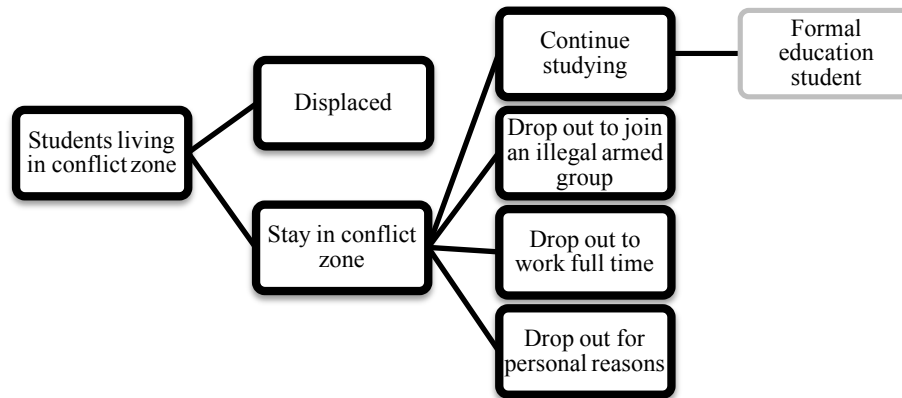
The students that stay in a conflict zone are also confronted with a decision regarding the desired level of education. Those individuals have to decide whether to continue studying, to join an illegal armed group with presence in their town, to drop out for other personal reasons or to find a job to contribute to the income of the household (*diagram 3*). As in the case of the displaced population, in order to take that decision the individuals weigh the costs associated with a given level of education, the expected benefits from a higher level of education, the expected benefits from dropping out for other personal reasons, the expected benefits from joining an illegal armed group and the expected income derived from finding a full time job.

The school age individuals that live in non-conflict zones also weigh the costs associated with a given level of education, the expected benefits from a higher level of education, the expected benefits from dropping out for other personal reasons (e.g. teenage pregnancy), and the expected income derived from finding a full time job in order to decide whether or not to continue with their formal education (*diagram 4*).

Those students residing in conflict and non-conflict zones who decide to continue with their formal education are the ones that attract our attention in this study. As a matter of fact, given data restrictions, it is only possible to study empirically the case of those students that can be observed, i.e. formal education students that stay in the conflict zone and formal education students living in non-conflict zones. The

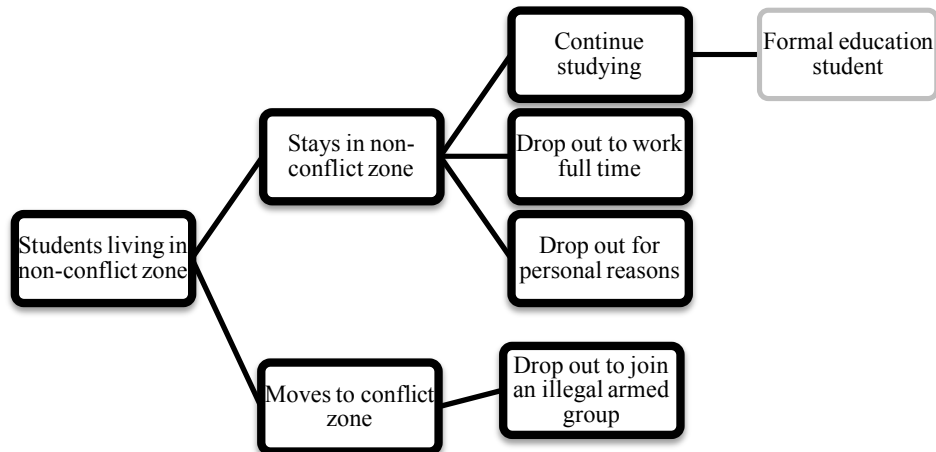
goal of this research is therefore to better understand the relationship between educational achievement and civil conflict amongst those students that participate in the formal educational system.

Diagram 3



Source: author's elaboration

Diagram 4



Source: author's elaboration

3. A Simple Model of Achievement

This section introduces a simple model of achievement in order to rationalize the analytical relationship between educational achievement and civil conflict. The model, proposed by Todd and Wolpin (2003), sets forth the idea that as students progress through the formal education curricula, their level of achievement in cognitive exams is a function of the history of family inputs, the school inputs and innate ability. However, for our purposes, in this case the model has been extended to incorporate the role of conflict as a determinant of educational achievement. The intensity of conflict and/or terrorist attacks can be seen as an environmental factor that causes stress amongst school aged individuals, and therefore should be included as a determinant of educational achievement.

Let A_1 denote the level of achievement at a point in time just before entering first grade, F_0 the family inputs that are put into the achievement production during the pre-school years $t = 0$, C_1 is an ambient factor (in this case the intensity of civil conflict events) which affects the student at a point in time just before entering first grade, and μ a measure of endowed ability determined at the time of conception. Achievement at the entrance of first grade is given by the following educational production function (g):

$$A_1 = g_0(F_0, C_1, \mu) \quad (1)$$

In this model we assume that family inputs in the period before entering school are determined by the family's permanent resources (W), and the endowed ability of the student. The intensity of conflict, modeled as an environmental factor, is included as a determinant of the educational production function because cognitive achievement might be affected in numerous ways due to exposure to conflict. Conflict directly affects educational achievement because it reduces family resources available for education. This is the case because in conflict zones labor opportunities tend to be limited and households have a higher level of exposure to violent attacks that can result in the destruction of family assets. Additionally, conflict zones are very prone to attacks that cause damage or destruction of the available public and private infrastructure. This also reduces the prospects of private investment in the conflict zone. But the effects are not restricted to the direct effects mentioned previously. Camacho (2008) points out that conflict generates disproportionate amounts of fear and stress amongst the population. This researcher highlights that the fear and stress generated, leads to indirect effects that might be more significant than the direct effects. Indeed, Camacho (2008) finds that the intensity of random landmine explosions during the first trimester of pregnancy has a significant negative effect on child birth weight. Other

indirect effects of conflict, highlighted by Camacho (2008), include: reduction in quality of life, anxiety, depression, sadness, disabilities, psychological stress, and other illnesses that deteriorate the human capital accumulation process.

Achievement at the entrance of second grade depends on the historical family inputs (F_0 and F_1), on school inputs (S_1), on environmental factors (in this case the intensity of civil conflict events) affecting the student at a point in time just before entering second grade (C_2), and on the μ a measure of endowed ability determined at the time of conception:

$$A_2 = g_1(S_1, F_1, F_0, C_2, \mu) \quad (2)$$

As mentioned previously, in this model we assume that although achievement can be measured at discrete points in time, the educational process is cumulative. This implies that in order to estimate this achievement production function it is necessary to take into consideration both school and family inputs, current and past (Todd and Wolpin 2003). That implication is perhaps the most important idea that this extension of Todd and Wolpin's model sets forth in terms of analyzing educational achievement amongst formal education students. In fact, it is important to keep in mind that the basic idea that guides most of the theoretical modeling and empirical exercises in this subject is that as students progress through the formal education curricula, their level of achievement in cognitive exams is a function of the history of family inputs, the school inputs and innate ability.

However, missing data for past inputs poses an important problem at the time of estimation. It should be acknowledged that one of the most problematic issues in analyzing student achievement is the lack of information on past family inputs and on children's heritable endowments. As Todd and Wolpin (2007) point out, it would be ideal to have access to data on all past and present family and school inputs but unfortunately historical information is most of the times missing in the datasets that are available. Given this limitation, Todd and Wolpin (2003) have developed a conceptual framework that could be implemented under different kinds of data limitations in order to estimate the relationship between achievement outcomes and relevant inputs. In the next section we will explore Todd and Wolpin's (2003) conceptual framework, and how it can be applied for our purposes.

4. Conceptual Framework

Following Todd and Wolpin's (2003) notation, let T_{ija} be a measure of achievement for student i who lives in municipality j at age a . As it was emphasized previously, we assume that knowledge acquisition is a production process in which current and past inputs are combined with the student's genetic endowment to produce a certain level of cognitive achievement (Todd and Wolpin 2003).

In this conceptual framework we assume that inputs are given by the choices that are being made by parents, schools and the actors of the conflict (causing environmental stress). Following Todd and Wolpin's (2003) notation, the vector of inputs supplied by the parents at a given age is represented by F_{ija} , the vector of school-supplied inputs is represented by S_{ija} and the vectors of their respective environmental factors (in this case the intensity of civil conflict events) affecting student i who lives in municipality j at age a . The student's endowed ability is the production function is represented as:

order to denote that the impact of inputs and of the genetic endowment depends on the age of the student.

4.1 The Contemporaneous Specification

As mentioned previously, missing data for past inputs poses an important problem at the time of estimation. In order to overcome that problem, Todd and Wolpin (2003) came up with a specification that only depends on contemporaneous measures of school and family inputs. This specification is often times adopted because of data limitations associated with limited availability of historical input measures or historical test score results. Considering that this specification only depends on contemporaneous observations, it is known as "Contemporaneous Specification". A number of assumptions are necessary to justify the application of this specification. According to Todd and Wolpin (2003), the assumptions on the production technology and on the input decision rules that have to be met are:

- (i) Only contemporaneous inputs matter to the production of current achievement.

or

- (ii) Inputs are unchanging over time, so that current input measures capture the entire history of inputs.

Either condition (i) or condition (ii) has to be met, and additionally condition (iii) has to be satisfied:

- (iii) Contemporaneous inputs are unrelated to (unobserved) endowed capacity.

Given these assumptions, the contemporaneous specification can be written as:

ε'_{ija} is an additive error that includes any omitted factors, including past inputs and endowed capacity, and the possibility of measurement error. Therefore, in equation (4) the residual term includes all the omitted factors. That implies that although this specification has weak data requirements, it requires strong assumptions in order to justify its application (Todd and Wolpin 2003).

As Hanushek (1979) highlights, the most consistent and obvious divergence between the empirical specification and the theoretical educational production function is the non-existence of a measure of innate or endowed abilities. However, omitting such an important variable can be problematic because it can bias the estimated regression coefficients. The size of the bias depends on how strong is the effect of the variable on achievement and also on the correlation of the omitted variable with the other variables that are included in the model. Hanushek (1979) points out that if endowed capacity is not correlated with the variables that are included in the model, the estimated coefficients of the model would not be biased. The only effect in such case would be an increase in the residual variance. Nevertheless, it is very unlikely that endowed capacity is not correlated with the other determinants of achievement. This implies that omitting the endowed capacity of the student from the empirical specification will likely cause an upward bias in the estimated impact of the other variables on achievement (Hanushek 1979).

4.2 Value-Added Specification

Todd and Wolpin (2003) propose another specification to deal with the problem of missing data for past inputs. This specification is known as the “Value-Added Approach” for estimating achievement production functions. The only difference between the value added approach and the contemporaneous specification, is that in the value added specification a lagged achievement measure is also included. In other words, the value added specification associates an achievement outcome measure to contemporaneous school and family input measures and a lagged achievement measure (Todd and Wolpin 2003). This specification assumes that the lagged achievement measure is a sufficient statistic for unobserved input histories as well as the unobserved endowment of mental capacity (Todd and Wolpin 2003).

Just like the contemporaneous specification, the value-added approach requires strong assumptions on the underlying production technology. Additionally, it should be acknowledged that by including a lagged test score as a conditioning variable, the model becomes more prone to endogeneity bias when data on some of the relevant inputs are missing. However, this specification is useful when longitudinal is not available.

Following Todd and Wolpin’s (2003) notation, let X denote the vector of family and school inputs and $X(a)$ denote the input histories up to age a . C_{ija} denotes the intensity of civil conflict events for student i who lives in municipality j at age a . Under this specification we are assuming that equation (3) can be written as a function only of a baseline test score and contemporaneous inputs (Todd and Wolpin 2003). Without loss of generality, let’s assume that the lagged test is taken by the student at time $a - 1$. Given this notation, the value-added model would be specified by:

If we assume that the arguments in equation (3) are additively separable and the parameters are non-age varying, we have the following estimating equation:

5. Measuring the Inputs of the Educational Production Function

To date, the literature has identified several factors that are considered to be determinants of students' educational outcomes. Hanushek (1986) points out that the inputs included in an educational production function should be relevant to the students under study. Following the structure of the educational production function described previously, in this research educational achievement of a given student at a particular point of time is presented as a function of the characteristics of the student's family, the characteristics of the schools and teachers, the innate abilities of the student, and conflict related factors (Hanushek 1986).

According to the literature, family background variables play a very important role on educational performance (Häkkinen et al. 2003, Haveman and Wolfe 1995, Woßmann 2003). Family inputs are usually measured by the socio-demographic characteristics of the families. In this dissertation, three different variables are going to be included to account for family background: parental education, family income, and family size⁵. The inclusion of the parental education variable is particularly important considering that studies, including those of Woßmann (2003) and Häkkinen et al. (2003), have found that students whose parents have a higher level of education have a better performance than those students whose parents have a low level of education. This could be the case because of genetic transmission of skills from parents to kids, and because highly educated parents tend to provide a good cultural environment that is adequate for learning. The literature has also shown that family income is related to educational achievement. The research done by Belley and Lochner (2007) and Davis-Kean (2005) has shown that relationship by examining cross-sectional and longitudinal datasets. Likewise, according to Downey (1995) and Blake (1981, 1989) the family size also plays an important role on educational performance. These researchers have found that resource dilution as the family size increases causes an inverse relationship between the number of siblings and educational performance. This finding is explained by the fact that parents have limited resources (time, energy, money, etc.) that have to be divided among children as the size of the family increases (Downey 1995).

The academic literature has found evidence to show that the characteristics of the schools and teachers have a significant role in determining educational achievement. For instance, Gamoran and Long (2006) have found that the characteristics of the school have an important impact on academic performance in developing countries, like Colombia. However, these research efforts are not

⁵ Additional variables might be included in some of the econometric exercises of the dissertation. The inclusion of those additional variables will be explained as needed.

completely conclusive. The existing literature has not reached consensus as to which variables have a significant impact and which do not. For instance, Hanushek and Luque (2003) did not find a significant impact of the school size on educational achievement, while Barnett et al (2002) found a significant impact between the two variables. This shows that the results are not pointing to the same conclusion in all cases. Unfortunately, the same is the case for those studies that have tried to establish a relationship between private/public schools and academic achievement (Hanushek 1986, Vandenberghe and Robin 2004, Somers et al. 2004). Considering that the existing literature has not reached consensus about which variables are significant, it is very important to control for the characteristics of the schools and the teachers that are considered to be relevant for the sample under study, as these might have an effect on the achievement of the students. To do so, in this dissertation several variables, for which there is data availability, are going to be included to account for school and teacher characteristics. Some of the variables included are school facilities available, number of teachers available in the school, and education level of the teachers⁶.

Schneider (2002) highlights the importance of school facilities as determinants of academic outcomes. The logic behind the analysis is particularly simple: how can students be high achievers if schooling infrastructure is not adequate? In fact, the research done by Earthman and Lemasters (1996, 1998) has found that there is a positive relationship between the available infrastructure and higher test scores. Earthman (2002) points out that school building components have a measurable influence on student learning. Indeed, according to his study, "in cases where students attend school in substandard buildings they are definitely handicapped in their academic achievement...correlation studies show a strong positive relationship between overall building conditions and student achievement" (Earthman 2002). The number of teachers available in the school is also included as an explanatory variable because it is possible that a low number of teachers per student negatively affect the learning environment. If a large group of students receives class from only one teacher, most likely the attention received will be less and this might have a negative effect on the achievement of students in evaluations. Additionally, the low number of teachers per student may be a factor discouraging school attendance. Nevertheless, the literature has not shown conclusive results when this variable is included as a determinant. Harbison and Hanushek (1992) reviewed the literature of the impact of teacher-pupil ratios and found that out of sixteen studies with statistically significant effects, eight were positive and eight were negative. These results cast doubt about the reliability of previous studies, but serves as an additional motivation to include this variable as

⁶ Additional variables might be included in some of the econometric exercises of the dissertation. The inclusion of those additional variables will be explained as needed.

part of this research. The education level of the teachers is another factor that possibly affects the achievement of students in standardized evaluations. Darling-Hammond (2000) points out that there is substantial evidence indicating that teachers who have more preparation for teaching are more confident and successful with students. In fact, Harbison and Hanushek (1992) found that teacher education had a significantly positive impact for math on primary school students in Brazil. Glewwe and Kremer (2006), which provides a very complete survey of the literature about schools, teachers and education outcomes in developing countries, highlights that developing countries reply to the shortage of trained teachers by hiring more untrained teachers. Such behavior might have a negative impact on the achievement of certain students. These studies provide justification for the inclusion of the education level of the teachers as a possible determinant of educational achievement.

As mentioned previously, the presence and intensity of conflict are also included as an environmental factor affecting the educational production function. Those measures of conflict are considered in this research because cognitive achievement might be affected in numerous ways due to ambient stress caused by exposure to conflict. Conflict directly affects educational achievement because it reduces the resources that are available for education. Nevertheless, the effects are not restricted to the direct effects associated to resource allocation. There are also indirect effects that might cause significant repercussions on the achievement of students. Camacho (2008) highlights that the indirect effects of conflict include: fear and stress amongst the population, reduction in quality of life, anxiety, depression, sadness, disabilities, psychological stress, and other illnesses that deteriorate the human capital accumulation process.

It is also expected that the innate or endowed abilities of the student play an important role in determining the level of achievement in standardized tests. However, the non-existence of a measure of innate or endowed abilities limits our capacity to capture this presumably important determinant. Such omission might lead to biased results. In the statistical analysis of the educational production function it is important to have in mind that omitting the endowed capacity of the student from the empirical specification will likely cause an upward bias in the estimated impact of the other variables on achievement (Hanushek 1979). The interpretation of the results should take that possibility into consideration, because the dismissal of a possible bias might lead to erroneous conclusions.

6. Concluding Remarks

Throughout this chapter, an extension of a simple model of achievement, originally proposed by Todd and Wolpin (2003), was considered as a theoretical framework to support the empirical exercises of the other chapters of the dissertation. This extended version of Todd and Wolpin's (2003) model sets forth the idea that as students progress through the formal education curricula, their level of achievement in cognitive exams is a function of the history of family inputs, the school inputs and innate ability. However, for our purposes, the model also incorporates the role of conflict as a determinant of educational achievement. The intensity of conflict and/or terrorist attacks is seen as an environmental factor that causes stress amongst school aged individuals, and therefore is included as a determinant of educational achievement. To complement the explanation of the theoretical model, we also employed the available literature to identify several factors that are considered to be determinants of students' educational outcomes. In the coming chapters, some of the theoretical ideas which were set forth in this chapter will be explored in more detail.

WORKS CITED

- Barnett, R. R., Glass, J. C., Snowdon, R. I., and Stringer, K. S. (2002). Size, performance and effectiveness: cost-constrained measures of best-practice performance and secondary-school size. *Education Economics*, 10(3), 291-311.
- Belley, P., and Lochner, L. (2007). The changing role of family income and ability in determining educational achievement: National Bureau of Economic Research.
- Blake, J. (1981). Family size and the quality of children. *Demography*, 18(4), 421-442.
- Blake, J. (1989). *Family size and achievement* (Vol. 3): University of California Pr.
- Camacho, A. (2008). Stress and birth weight: evidence from terrorist attacks. *The American Economic Review*, 98(2), 511-515.
- Darling-Hammond, L. (2000). How teacher education matters. *Journal of teacher education*, 51(3), 166-173.
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294.
- Downey, D. B. (1995). When bigger is not better: Family size, parental resources, and children's educational performance. *American Sociological Review*, 746-761.
- Earthman, G. I. (2002). School facility conditions and student academic achievement.
- Earthman, G., and Lemasters, L. (1996). Review of Research on the Relationship between School Buildings, Student Achievement, and Student Behavior.
- Earthman, G. I., and Lemasters, L. (1998). Where Children Learn: A Discussion of How a Facility Affects Learning.
- Gamoran, A., and Long, D. A. (2007). Equality of Educational Opportunity A 40 Year Retrospective *International studies in educational inequality, theory and policy* (pp. 23-47): Springer.

- Glewwe, P., and Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. *Handbook of the Economics of Education*, 2, 945-1017.
- Häkkinen, I., Kirjavainen, T., and Uusitalo, R. (2003). School resources and student achievement revisited: new evidence from panel data. *Economics of education review*, 22(3), 329-335.
- Haladyna, T. M., Nolen, S. B., and Haas, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher*, 20(5), 2-7.
- Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources*, 14(3), 351-388.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 24(3), 1141-1177.
- Hanushek, E. A. (2008). Education Production Functions. In L. E. B. Steven N. Durlauf (Ed.), *The New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan.
- Hanushek, E. A., and Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics of education review*, 22(5), 481-502.
- Harbison, R. W., and Hanushek, E. A. (1992). *Educational performance of the poor: lessons from rural Northeast Brazil*: Oxford University Press.
- Haveman, R., and Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of economic literature*, 33(4), 1829-1878.
- Horn, S. P. (1995). Educational Assessment Reassessed: The Usefulness of Standardized and Alternative Measures of Student Achievement as Indicators for the Assessment of Educational Outcomes William L. Sanders. *education policy analysis archives*, 3, 3-1995.
- Linn, R. L. (2001). A century of standardized testing: Controversies and pendulum swings. *Educational Assessment*, 7(1), 29-38.
- Schneider, M. (2002). Do School Facilities Affect Academic Outcomes?

- Somers, M. A., McEwan, P. J., and Willms, J. D. (2004). How effective are private schools in Latin America? *Comparative Education Review*, 48(1), 48-69.
- Todd, P. E., and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement*. *The Economic Journal*, 113(485), F3-F33.
- Todd, P. E., and Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital*, 1(1), 91-136.
- Vandenberghe, V., and Robin, S. (2004). Evaluating the effectiveness of private education across countries: a comparison of methods. *Labour economics*, 11(4), 487-506.
- Wößmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2), 117-170.

CHAPTER 6

MEASURING EDUCATIONAL OUTCOMES – EXTERNAL EXIT EXAMINATIONS

As Hanushek (1986) highlights, it is essential to use adequate measures of educational outcomes in order to analyze school production. However, measuring these outcomes is not an easy task. It is possible to find many different measures of educational outcomes in the literature, including: school attendance rates, school enrollment rates, exit exam scores, and dropout rates. Nevertheless, for the purposes of this dissertation exit exam scores are going to be employed as a measure of educational outcomes. Although some academic works have criticized the use of test scores as a measure of outcome, there are significant reasons to employ this measure. Hanushek (1986) defines that the goal of measuring educational output is to find a quantitative measure that is readily available and is also associated to long-run objectives of schooling. In that sense, exit examination scores are valued in and out themselves.

According to Hanushek (1986), most educators and policy makers believe that central examinations are important measures of education. As a matter of fact, the results of national exams are used to evaluate educational programs and in some countries they are also used as criteria to allocate funds to schools. Additionally, the scores on standardized exams are now being used in many places as a requisite in order to graduate from high school. Moreover, certain scores are required in order to continue with further schooling. It should also be noted that in terms of educational research a very important characteristic of standardized exams is their common availability. In sum, the characteristics mentioned above comply with the objectives of measuring educational output set forth by Hanushek (1986).

In what follows, a definition of external exit examinations and an overview of the historical development of these examinations will be provided. The specific case of the Colombian High School Exit Examination (ICFES Exam) will be explored in detail. The historical context of the Colombian central exam will be discussed, and an analytical justification of the suitability of this exam as an external exit examination will be provided. A brief review of the academic literature that has analyzed the results of this exam will also be included.

2. Curriculum-based External Exit Examination

As mentioned previously, exit exam scores are going to be employed as a measure of educational outcomes in this dissertation. The literature has highlighted certain aspects that characterize exit examinations. The definition proposed by Bishop (1997) will be followed in this case to identify the characteristics that exit examinations should have. According to Bishop (1997) a curriculum-based external exit examination should possess the following characteristics:

- i. Produces signals of student accomplishment that have real consequences for the student.
- ii. Defines achievement relative to an external standard, not relative to other students in the classroom or the school.
- iii. Is organized by discipline and keyed to the content of specific course sequences.
- iv. Signals multiple levels of achievement in the subject.
- v. Covers almost all secondary school students.

It is important to explore these characteristics in more detail. The first characteristic points out that the implementation of an exit examination should help improve the signaling of academic achievement, and in doing so, the recompenses derived from learning should be greater and more visible. The results obtained in a curriculum-based external exit exam should play a significant role, either in terms of further schooling or professional development, in such a way that real incentives are generated. The second characteristic highlights that curriculum-based external examinations should shift away the attention from relative measures of achievement, including class rank and school specific grades to a more objective measure of achievement that is relative to an external standard (Bishop 1997). This second characteristic is particularly important considering that it eliminates perverse incentives that some students might have to persuade other students to study less in order to maintain lower average grades, or even to get the teachers off track (Bishop 1998). The third characteristic points out that exit examinations should evaluate separately the specific contents of the different disciplines that students learn in school. This characteristic implies that the responsibility for preparing the students for a particular exam is concentrated on one or very few teachers (Bishop 1998).

As Bishop (1998) highlights, characteristic four is particularly important because of its impact on the incentive effects of exams. By signaling multiple levels of achievement in the subject, curriculum-based external exit examinations create stronger incentives for effort. This is so because most students tend to work harder

when the full range of achievement is reported rather than just a pass or fail report. According to Bishop (1998), “few students will find the reward for exceeding a single absolute cutoff an incentive for greater effort.” The fifth characteristic points out that the coverage of the exit examinations should be as universal as possible. This characteristic plays an essential role because if exams were only administered in a limited set of elite schools or amongst those students specializing in a particular field, they would influence standards amongst those schools and students that are being evaluated, and only have limited or null effects on the bulk of students (Bishop 1997).

3. The Origins of Examinations

Different countries have established national exit exams at different time frames and for differing reasons. In fact, according to Klein and Van Ackeren (2011), the different contexts within which the examinations were established and the differing institutional traditions and cultural beliefs have produced examinations with specific functions in each country.

The origins of central examinations go back to 210 B.C. in China when a system of examinations was used to select government officials (Madaus 1993). However, according to Keeves (1994), it was only in the 16th century that the Jesuits introduced a system of competitive examinations for selection purposes in their schools and universities in Europe. Later in the 18th century, Europe followed the Chinese tradition of using examinations to appoint individuals to the civil service, in order to substitute the system that was being used up to that date which was based on patronage and nepotism. Through the use of these examinations, it was hoped that the best candidates would be chosen for the job. According to Madaus (1991), Germany (Prussia) was the first European country to implement civil service examinations. In fact, by the year 1748 recruitment to government positions in the judiciary offices and administrative services was guided by the performance in the civil service examinations. After the revolution, France also implemented the system of civil service examinations in order to select the individuals with the highest qualifications in order to establish a powerful absolutist state. It should be highlighted that although France implemented civil service examinations at a later date than Germany, this country had more frequent examinations covering more subjects. According to Madaus (1991), the examinations that were held in France during that time were very demanding and the failure rate was high.

Even though the Jesuits introduced a system of competitive examinations for educational selection purposes in the 16th century, it was only in the late 18th century and early 19th century that central examinations to determine admission

to university became widely used in countries like Germany and France. The Abitur exam, which was introduced in Germany in 1788, helped determine which students would be admitted to university. As mentioned by Madaus (1991), this examination was implemented to help improve the quality of the universities by filtering out students with low levels of achievement. France started some years later with the introduction of the Baccalaureat in 1808 by Napoleon. According to Madaus (1991), the Baccalaureat was used as a filter in the admission process of the grandes ecoles and as a determinant to participate in the government service and other professions.

Public examinations were also established in the 19th century in Great Britain. Just like in Germany and France, the examinations in Great Britain were established to recruit individuals with the highest levels of achievement into the professions, the public service and university. The establishment of examinations dates back to 1815, when the Society of Apothecaries introduced a written qualifying examination to be admitted into the profession (Madaus 1991). The use of central examinations to be admitted to the civil service was also established in Great Britain in the 19th century. In fact, as Madaus (1991) highlights, by 1870 most of the civil service appointments were based on examinations. The use of examinations grew considerably during the second half of the 19th century as the utilitarian ideas of Adam Smith and Jeremy Bentham expanded (Madaus 1991). According to utilitarian ideas, self-interest is the principal reason to study and because studying requires hard work it is necessary to provide incentives for those students that are successful by giving out certificates and prizes (Madaus 1991). Additionally, it was believed that the outcomes from learning could be measured with some accuracy. This gave rise to the essence of written examinations: the reproduction of factual knowledge (Madaus 1991).

Central examinations have been used to measure the levels of achievement of individuals for a very long time. Nevertheless, it is worth noting that the prevalence and the accountability purpose of these tests have grown considerably during the past two decades (Volante 2007). In fact, the use of central examinations for accountability purposes has been a very powerful trend in educational policy in the last 20 years (Volante 2007). According to Volante (2007), the implementation of a test-based accountability system in Great Britain during the eighties by Prime Minister Margaret Thatcher has been a point of reference for other proponents of the test-based/ standards-based reform around the world, including the United States, Canada, and Australia. However, this list of countries should not be restricted to developed countries. Developing countries, like Colombia, also implemented changes to their High-School exit exam starting in the eighties, in line with this test-based reform to improve the accountability of the educational system.

4. Colombian High School Exit Exam – ICFES Saber 11

The Colombian nation-wide High School exit exam (ICFES Exam) was first created more than forty years ago in order to provide support to universities and other higher education institutions in their admissions processes. Throughout its forty years, the exam has gone through a series of important changes to capture the needs of the educational system in Colombia. In what follows, a brief explanation of the changes and the evolution of the ICFES exit exam will be provided.

In 1966 the Colombian Association of Universities (*Asociación Colombiana de Universidades*) and the Universities Fund (*Fondo Universitario*) signed an agreement to jointly create the College Admissions and Job Guidance Service (*Servicio de Admisión Universitaria y Orientación Profesional*). This Service was created to design, administer and evaluate different instruments that could be used by Colombian universities in their admission processes. This was the first step towards the creation of the nation-wide ICFES exam for college admission. The College Admissions and Job Guidance Service administered admissions exams in different universities from 1966 until 1968. However, those exams were only offered and administered when specific universities requested them, and therefore did not have national coverage (Ministerio de Educación 2013).

In 1968, the Colombian Institute for the Evaluation of Education (*ICFES*) was created. The creation of this institute led to the formal establishment and implementation of the nation-wide high school exit exam, ICFES exam, in Colombia. One of the offices at the newly created ICFES, the National Testing Service (*SNP*), was made responsible for administering the first nation-wide college admissions and high school exit exam. It should be highlighted that the National Testing Service (*SNP*) was an improved version, after a number of re-organization efforts, of the College Admissions and Job Guidance Service. The first nation-wide exams implemented at that time tested skills in different areas, including: mathematics, verbal skills, abstract reasoning skills, spatial relationships, social sciences, philosophy, chemistry, physics, biology and English (Ministerio de Educación 2013).

However, the nation-wide college admission and exit exam was not mandatory for many years. It was only in 1980 when the national exams for admission to higher education were officially mandated as a pre-requisite for graduation for all Colombian high school seniors. As a result, from that year on the exit exam evaluated almost all the universe of students graduating from high school. Starting that year, the results from the exit exam also became a requirement to be admitted to any undergraduate program in Colombia (Ministerio de Educación 2013).

In the eighties, new changes were introduced to the ICFES exit exam. From 1980 to 1999, the exam was organized such that nine different subjects that were grouped in five areas were tested. This nation-wide exam was conceived as an exam to test specific contents in the nine different subjects. During those years, the results were reported by subject and by area. Additionally, a total score, calculated as the sum of the scores in the five areas that were tested, was also reported. The five areas (and the subjects) that were tested in that exam were:

- Natural sciences: biology, chemistry, and physics.
- Language: verbal aptitude, Spanish, literature.
- Mathematics: mathematical aptitude and mathematical skills.
- Social sciences: history, geography and philosophy.
- Elective: the students could choose one test from a list that included different subjects (for example, English, French, accounting, democracy, abstract reasoning).

In the year 2000, a number of very important changes were introduced to the nation-wide college admission and high school exit exam. More specifically, the emphasis of the exam changed. As mentioned previously, the nation-wide exam was originally conceived as an exam to test specific contents, but in 2000 it became an exam to test competencies. The changes that were introduced were the result of the implementation of new educational goals in the Colombian educational system. The exam had to be consistent with the new educational requirements based on competencies established in the Colombian legislation in the General Education Law of 1994⁷. According to ICFES, the exam should be geared towards evaluating competencies, i.e. knowing what to do in different contexts (ICFES 2010). This implies that successful students need to be able to use and bring together all their knowledge and abilities to solve a specific problem.

The specific objectives of the exam are established by law (Decree 869 of 2010). These objectives include:

- To be an admissions criteria for higher education.
- To inform students about their competencies in the different areas that are being tested, in order to guide their professional career options.
- To support the auto-evaluation and continuous improvement efforts of schools.
- To become the basis for the development of new social, educational and cultural investigations.
- To be a criteria to allocate educational grants.

⁷ For more information about this Law, refer to the *history of Colombian education* section of this thesis.

The newest version of the exam is divided in two main components: the core subjects component (*eight different subjects*) and the flexible component. The core subjects' component is compulsory for all test takers and it evaluates the following subjects: language (*Spanish*), mathematics, biology, chemistry, physics, social sciences (*starting in the year 2006*), philosophy, and English (*mandatory since 2007*). It should be highlighted that before the year 2006, social sciences was divided in two independent subjects: geography and history. It is also important to mention that before the year 2007, the English exam was not a mandatory part of the core subjects' component of the exam. The flexible component has two different types of tests: advanced skills and interdisciplinary. The advanced skills exam tests with a higher level of complexity one of the core subjects. Students can choose from language (*Spanish*), mathematics, biology and history. The flexible component exam tests multidisciplinary problem solving skills. The results for each of the subjects of the core subjects' component are reported in a 0 to 100 scale (ICFES 2010).

5. The Colombian Exit Examination: A Curriculum-based External Exit Examination?

The definition of a curriculum-based external exit examination proposed by Bishop (1997), which was discussed earlier in this document, can be used to identify the characteristics that exit examinations should have. The Colombian exit exam administered by ICFES has all these characteristics, and in doing so, complies with the definition of a curriculum-based external exit examination set forth by Bishop (1997). In what follows, these characteristics will be discussed for the case of the Colombian exit examination.

- i. Produces signals of student accomplishment that have real consequences for the student.

The results obtained in the ICFES exam play a very significant role in terms of further schooling and professional development in Colombia. In order to get admitted to university or technical school, Colombian students should obtain a minimum score in the exit exam. The required score depends on the policies of each educational institution, but in general the most prestigious institutions require high levels of achievement. The Colombian exit examination indeed improves the signaling of academic achievement, and in doing so generates real incentives to make greater efforts to have a better performance.

- ii. Defines achievement relative to an external standard, not relative to other students in the classroom or the school.

The Colombian exit examination defines achievement relative to a national standard, not relative to students in the classroom or the school. In that sense, the ICFES exam is shifting away the attention from relative measures of achievement, including class rank and school specific grades, to a more objective measure of achievement that is relative to an external standard (Bishop 1997).

- iii. Is organized by discipline and keyed to the content of specific course sequences.

This third characteristic establishes that exit examinations should evaluate separately the specific contents of the different disciplines that students learn in school. The Colombian exit examination has this characteristic because it is organized by disciplines. Indeed, the newest version of the ICFES exit exam has a set of compulsory tests that evaluate eight different core subjects (language, mathematics, biology, chemistry, physics, social sciences, philosophy, and English).

- iv. Signals multiple levels of achievement in the subject.

The results for each of the subjects evaluated in the Colombian exit examination are reported in a 0 to 100 scale. Additionally, the results are categorized by levels of achievement depending on the score (significantly high, high, medium, low, and significantly low). These reports are structured so that they can signal multiple levels of achievement in the subject, therefore complying with characteristic four. By creating those signals, the Colombian exit examination can generate stronger incentives because most students tend to work harder when the full range of achievement is reported rather than just a pass or fail report.

- v. Covers almost all secondary school students.

According to this fifth characteristic, the coverage of the exit examinations should be as universal as possible. In the eighties, the Colombian exit examination became officially mandated as a pre-requisite for graduation for all Colombian high school seniors. As a result, from that year on the exit exam has evaluated almost all the

universe of students graduating from high school. This implies that the Colombian exit exam complies with this fifth characteristic.

As shown above, the Colombian exit examination has the five characteristics proposed in the definition of a curriculum-based external exit examination by Bishop (1997). Therefore, throughout this dissertation the Colombian exit examination will be considered and defined as a curriculum-based external exit examination.

6. Exit Examinations – Weaknesses and Strengths

The literature has pointed out a number of potential weaknesses and strengths related to the use of exit examinations. In this section, some of these weaknesses and strengths are discussed by referring to the existing literature.

Studies including those of Haladyna et al. (1991), Horn (1995) and Linn (2001), have discussed difficulties in the use of exit exam scores. Most of the criticisms to the use of standardized test scores, which have been mentioned in the literature, are related to test score pollution. Test score pollution is a problem that is often times associated to contaminated results because of teachers that are teaching to the exam and schools that are administering exams under non-standard conditions (Haladyna et al. 1991). The use of national examinations might limit the approaches to learning. It is also possible that teachers limit their attention to the material that is being tested in the central examination. This is particularly troublesome considering that the topics that can be covered in a national examination are limited, and therefore it is possible that important educational objectives might not be part of the curriculum that is being taught in the classrooms. As Madaus (1991) highlights, the curriculum that is being taught in the schools can be affected in such a way that examinations may end up determining the curriculum instead of the curriculum determining what should be evaluated in the examinations. This problem is magnified when the perception of the parents and the general public regarding the objective and the quality of education is centered in the results of these evaluations. It is also important to highlight that exams might be constraining curricular variety, and in that sense certain contents that are specific to the needs of certain students and locations might not be included (Madaus 1991). This is very problematic considering that it might lead to incongruences between the specific educational needs and what is being taught in order to perform in the evaluations.

Issues related to problematic testing design are also common in the literature (Horn 1995). There are a number of limitations related to the use of central

examinations highlighted by Madaus (1991) related to problematic testing design. Madaus (1991) highlights that central examinations are often times criticized because the conditions under which students are examined tend to be artificial and the amount of time available to answer the exam is limited. Such conditions are not adequate for all students, and can possibly generate anxiety and stress. Another criticism that has been mentioned in the literature is the possibility that national examinations generate incentives to memorize specific topics instead of developing a deeper understanding of the subject.

It should be noted that it is possible that the negative effects that are usually attributed to exit examinations could be weaker than expected because they are possibly interacted with the schooling contexts in which the examination occurs (Sloane and Kelly 2003). For instance, it is not always clear if the anxiety that students many times show in exit examinations is due to the examination, or to a feeling of insecurity because of inadequate preparation or inadequate instruction. According to Sloane and Kelly (2003), "it is important that teachers and policy makers not blame the thermometer for the fever."

There are also some strong points in favor of nationwide exit examinations. According to Madaus (1991), one of the reasons behind the initial implementation of examinations is that they help reduce the effects of patronage, and in that sense they open higher education to a wider population of students. Madaus (1991) also highlights that national examinations guarantee a degree of homogeneity in educational standards nationwide. That is the case because the majority of countries have one central authority responsible for the organization of the central examinations, to assure that students throughout the country take the same exam for the different subjects that are being evaluated (Klein and Van Ackeren 2011).

According to Klein and Van Ackeren (2011), statewide examinations can potentially affect and regulate work at the school and classroom level positively. Such influence is explained by the imposition of minimum standards and the implementation of a new curriculum according to which students are going to be evaluated. These two factors generate a need to increase the levels of commitment amongst students and teachers, and in doing so, might result in an improvement in the level of instruction. Klein and Van Ackeren (2011) highlight that national examinations are an explicit and transparent source of information about the competencies that students have acquired throughout their studies, and in that sense they provide feedback on where schools have to improve. It can also be argued that national examinations place additional incentives for students to study and have higher levels of performance because the results can serve as an instrument to make them more accountable for their work. National examinations can be a way to provide tangible incentives to students and teachers, and remove

tensions between them because assessment cannot be affected by personal relationships (Madaus 1991).

Exit examinations place the attention of everyone in the school on the common goal of ensuring that all students possess the skills to pass the exam (Jacob 2001). Such focus can eventually lead to higher levels of achievement. That is the case because the special consideration that is placed on the exam results provides incentives to students to do their schoolwork more effectively, thus improving their achievement and school performance.

Administering exit examinations can also be a way for schools and national authorities to identify potential problems in the schooling system. According to Jacob (2001), it might be the case that competency testing allows teachers to identify student and school weaknesses, and focus remediation efforts.

Even though some academic works, including those of Haladyna et al. (1991), Horn (1995) and Linn (2001), have criticized the use of standardized test scores as a measure of outcome, exit examinations scores are a very useful tool in educational research. Nevertheless, the criticisms mentioned above should not be completely dismissed and the information obtained from the results of the exit examinations should be interpreted cautiously taking into consideration potential weaknesses.

7. Use of the results of the Colombian Exit Exams in the Literature

There are a number of academic papers that have used the results from the Colombian exit exam to analyse the determinants of the achievement levels of students, and assess the quality of academic institutions in Colombia. In what follows, some of these papers are surveyed to establish the state of the art in the empirical literature using the results of the exam.

Tobón, Posada and Ríos (2009) analyse the determinants of school performance in the last year of secondary school students in Medellín (Colombia) using ICFES standardized tests and data collected in a survey. The authors of this paper use a hierarchical model to quantify the contribution of a number of variables, which are usually included in the empirical literature. Tobón, Posada and Ríos (2009) evaluate the performance at three different levels of analysis (student, family, school). The empirical results show that school level variables, including management, explain a high percentage of the variability in the student's performance, while those constructed variables associated to student's characteristics have a low impact. Tobón, Posada and Ríos (2009) also find that parent human capital and the initial student conditions are much more important

than the teacher human capital.

Following the same line of analysis of Tobón, Posada and Ríos (2009), Correa (2011) analyses the determinants of school performance in the last year of secondary school students in Cali (Colombia) using ICFES standardized tests and data collected in a survey. Using a multilevel (two level) linear model, the author finds that there are significant differences in achievement, which can be explained by the distinct characteristics of schools. Additionally, according to Correa (2011) the strength of the relationship between achievement and individual and family characteristics varies positively and significantly between schools. This result implies that there are notable differences in the capacity of schools to help students overcome human capital inequalities.

Chica et al. (2010) have also studied the determinants of academic achievement in the Colombian exit examination. These authors used a generalized ordered logit to study the determinants of achievement. Chica et al. (2010) analyse the results of students in the mathematics and language (Spanish) test of the Saber 11 exam that was administered in the second semester of 2009. The authors highlight that socioeconomic variables, especially income level and the level of education of the parents, have a positive and significant relation with achievement in the standardized exam.

On a slightly different note, Gaviria and Barrientos (2001) study the determinants of educational quality in secondary education in Colombia using the results of the ICFES exit exam. Three different aspects related to academic achievement are considered separately: family characteristics, school characteristics, and public financing. The authors highlight that parental education has a significant effect on the achievement of students. Additionally, Gaviria and Barrientos (2001) point out that the ratio of teachers to students plays an important role. They note that despite a substantial increase in public investment in education, the quality differences amongst private and public institutions remains unaltered. In that sense, Gaviria and Barrientos (2001) suggest that public education is primarily affected by inadequate incentives and a problematic organizational structure, and only secondarily by the lack of financial resources.

The results of the Colombian exit exam have also been used to study the difference between public and private schooling. Núñez, Steiner, Cadena, and Pardo (2002) study the differences in achievement in the exit exam between students from private and public institutions in Colombia. They find that students from private schools have significantly higher levels of achievement in these exams. They attribute the difference to non-observable characteristics, like the structure of incentives for teachers and the teaching career. The authors highlight that students

from public schools administered by the private sector tend to perform better than students from traditional public schools. This suggests that the articulation between the public and the private sectors in education might be a possible alternative to improve quality. It is important to highlight that Núñez, Steiner, Cadena, and Pardo (2002) control for possible endogeneity issues in the choice of school.

Bonilla (2011) explores another characteristic of schools to determine if it has any effects on the quality of schooling. More specifically, this author uses instrumental variable regressions to evaluate if double-shift schooling has a negative effect on the quality of education in Colombia. Bonilla (2011) estimates the local average treatment effects of double-shift schooling using information from the Colombian exit examination. The results suggest that attending single-shift schools has a positive impact on academic performance, particularly if compared to students attending the afternoon shift. The findings from this study suggest that it is very important to control for different schooling shifts in econometric exercises.

Other authors, like Sánchez (2011), have used the results of the Colombian exit examination to understand the relationship between student characteristics and academic achievement. More specifically, Sánchez (2011) uses information from the Colombian exit exam to study the relationship between ethnicity and academic achievement. This research documents the existence of an academic gap between ethnic and non-ethnic students and decomposes this gap in factors related to observable characteristics, such as family and school, and non-observable factors. Sánchez (2011) uses quantile regressions to determine the existence of test score gaps throughout the distribution of academic performance. Results indicate that for Departamentos where there is a statistically significant gap, an important part of it can be attributed to nonobservable factors.

8. Concluding Remarks

Although some academic works have criticized the use of test scores as a measure of outcome, there are significant reasons to employ it for that purpose. Throughout this chapter, we provided a definition of external exit examinations and an overview of the historical development of these examinations. The case of the Colombian High School Exit Examination (ICFES Exam) was explored in detail. A discussion of the historical context of the Colombian central exam allowed us to justify the suitability and validity of this exam as an external exit examination. This analysis was carried out following the definition of external examinations provided by Bishop (1997). Checking the suitability of the information that is used in the empirical chapters of this thesis to understand educational outcomes is

particularly important. After all, as Hanushek (1986) highlights it is essential to use adequate measures of educational outcomes in order to analyze and understand the determinants of performance.

WORKS CITED

- (ICFES), Instituto Colombiano para la Evaluación de la Educación (2010). Orientaciones para el Examen de Estado de la Educación Media (pp. 30). Bogotá: Instituto Colombiano para la Evaluación de la Educación (ICFES).
- Bishop, J. H. (1997). The Effect of National Standards and Curriculum-Based Exams on Achievement. *The American economic review*, 87(2), 260-264.
- Bishop, J. H. (1998). The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement. *The Journal of Economic Education*, 29(2), 171-182. doi: 10.1080/00220489809597951
- Bonilla, L. (2011). Doble jornada escolar y calidad de la educación en Colombia. *Documentos de Trabajo sobre economía Regional*, 143.
- Chica Gómez, S. M., Gutiérrez, D. M. G., and Hassan, A. R. (2012). Determinantes del rendimiento académico en Colombia. Pruebas ICFES-Saber 11o, 2009*. *Revista Universidad EAFIT*, 46(160), 48-72.
- Correa, J. J. (2011). Determinantes del rendimiento educativo de los estudiantes de secundaria en Cali: Un Análisis multinivel. *Sociedad y Economía*(6), 81-105.
- Gaviria, A., and Barrientos, J. H. (2001). Determinantes de la calidad de la educación en Colombia.
- Haladyna, T. M., Nolen, S. B., and Haas, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher*, 20(5), 2-7.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 24(3), 1141-1177.
- Horn, S. P. (1995). Educational Assessment Reassessed: The Usefulness of Standardized and Alternative Measures of Student Achievement as Indicators for the Assessment of Educational Outcomes William L. Sanders. *education policy analysis archives*, 3, 3-1995.
- Jacob, B. A. (2001). Getting tough? The impact of high school graduation exams. *Educational evaluation and policy analysis*, 23(2), 99-121.

- Keeves, J. P., and Planning, I. I. f. E. (1994). *National examinations: design, procedures and reporting*: UNESCO, International institute for educational planning.
- Klein, E. D., and van Ackeren, I. (2011). Challenges and problems for research in the field of statewide exams. A stock taking of differing procedures and standardization levels. *Studies in Educational Evaluation*, 37(4), 180-188.
- Linn, R. L. (2001). A century of standardized testing: Controversies and pendulum swings. *Educational Assessment*, 7(1), 29-38.
- Madaus, G. F., and Kellaghan, T. (1991). Student Examination Systems in the European Community: Lessons for the United States.
- Madaus, G. F. (1993). A national testing system: Manna from above? An historical/technological perspective. *Educational Assessment*, 1(1), 9-26.
- Ministerio de Educación Nacional (2013). Examen de Estado para el Ingreso a la Educación Superior. Retrieved August 21, 2013, from <http://www.colombiaaprende.edu.co/html/home/1592/article-156080.html>
- Núñez, J., Steiner, R., Cadena, X., and Pardo, R. (2002). ¿ Cuáles colegios ofrecen mejor educación en Colombia? *Archivos de Economía*, 193, 1-56.
- Sanchez-Jabba, A. (2011). Etnia y rendimiento académico en Colombia. *Revista de Economía del Rosario*, 14(2).
- Sloane, F. C., and Kelly, A. E. (2003). Issues in high-stakes testing programs. *Theory into Practice*, 42(1), 12-17.
- Tobón Orozco, D., Posada Duque, H. M., and Ríos Gallego, P. (2009). Determinants of the performance of the schools in Medellin in the High-School Graduation-Year Test (ICFES). *Cuadernos de Administración*, 22(38), 311-333.
- Volante, L. (2007). *Evaluating test-based accountability systems: An international perspective*. Paper presented at the Proceedings of the Annual Meeting of the Association for Educational Assessment.

CIVIL CONFLICT AND EDUCATIONAL ACHIEVEMENT: A MULTILEVEL ANALYSIS APPROACH

A civil conflict has significant economic, social and political repercussions that could potentially affect the educational achievement of individuals. Nevertheless, this important issue has been insufficiently addressed in the literature. The purpose of this chapter is to explore the relationship between civil conflict and educational achievement in Colombia through the use of multilevel analysis. Multilevel analysis is employed in this chapter in order to deal adequately with the hierarchical structure of the data by simultaneously modeling variables at different levels without resorting to aggregation or disaggregation of the dataset. Taking advantage of a unique database administered by a Colombian think tank (CERAC^{*}) that contains records of civil conflict events in Colombia, a government kept database containing the results of standardized tests at different levels of schooling, and additional information about schools and teachers from the National Department of Statistics of Colombia (DANE), we were able to construct a multilevel model to further investigate the relationship between civil conflict and educational achievement.

There are a number of mechanisms at work when examining the relationship between civil conflict and educational outcomes. For instance, are kids scared and thus can't focus on schoolwork? Are kids working harder in order to succeed under harder conditions? Are families making a special effort to provide a high quality education to their kids in order to keep them away from the influence of the conflict? Are schools in conflict zones well equipped? Do kids have to dropout from school because of forced displacement? Are teachers scared and thus can't focus on teaching? Are teachers working harder in order to succeed under harder conditions? These mechanisms could be grouped into two central categories: those that affect students who stay in the conflict zone and those that affect students who drop out because of civil conflict. However, it is only possible to empirically model the case of those students that can be observed, i.e. students that stay in the school system. Therefore, our attention in this chapter concentrates on those students residing in conflict and non-conflict zones who are able to take part in the formal education system. Hence, the objective of this research is to better understand the relationship between educational achievement and civil conflict amongst those students that participate in the formal educational system.

*The contents of this chapter correspond to the first publishable paper of this PhD dissertation.

^{*} Resource Centre for Conflict Analysis.

It is relevant to explore the particular case of the students that stay in the conflict zones because these students might be adapting (or not) to the rough conditions that are prevalent in these areas, and perhaps they have a different (smaller/larger) educational gap (in terms of magnitude) than policy makers think. Although at first glance it might seem straight forward to think that there is a negative relationship between civil conflict and education, there are reasons to think that this relationship is not as straightforward as it seems. The existing literature studying the relationship between educational enrolment and civil conflict has shown mixed results. For instance, Swee (2008) and Bellows and Miguel (2006) find no significant effects of civil war on enrolment, whereas Lai and Thyne (2007) and Akresh and de Walque (2008) find evidence to show that civil war has a negative effect on enrolment rates. Furthermore, the researchers that have studied the relationship between education, conflict and gender have also found mixed results. That is the case of Shemyakina (2011) and Annan et al. (2011) who have obtained totally different results in their studies of the impact of armed conflict on school enrolment by gender. Shemyakina (2011) shows that exposure to conflict had a negative effect on the school enrolment of girls, and little or no effect on enrolment of boys, whereas Annan et al. (2011) find significant adverse effect for males but not for females. The lack of consistency in the existing literature studying the relationship between conflict and education casts doubts about the results obtained to date, and motivates further research in this topic in order to explore the relationship in more detail.

This chapter contributes to the existing literature that studies the relationship between civil conflict and education, including the contributions of Swee (2008), Akresh and de Walque (2008), Merrouche (2006), Lai and Thyne (2007), Shemyakina (2011), Debalen and Paul (2012), Bundervoet (2012), and Bellows and Miguel (2006). These research efforts have focused on the relationship between school enrolment, gender, years of education and civil conflict, but have not studied the relationship between civil conflict and academic achievement measured by cognitive examinations. To this date, the only attempt to understand the relationship between civil conflict and educational achievement in standardized examinations was made by Rodriguez and Sánchez (2010), demonstrating that this area of research is very understudied. This chapter differs significantly from Rodriguez and Sánchez (2010) because we incorporate the hierarchical structure of the data into the analysis. In fact, to the best of my knowledge this is the first effort to try to understand the relationship between conflict and achievement in standardized exams through the use of multilevel analysis.

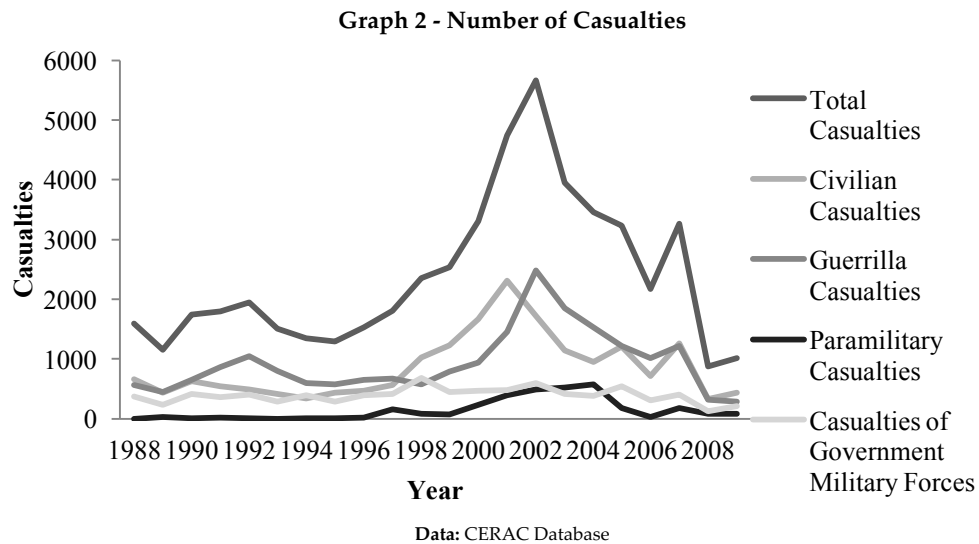
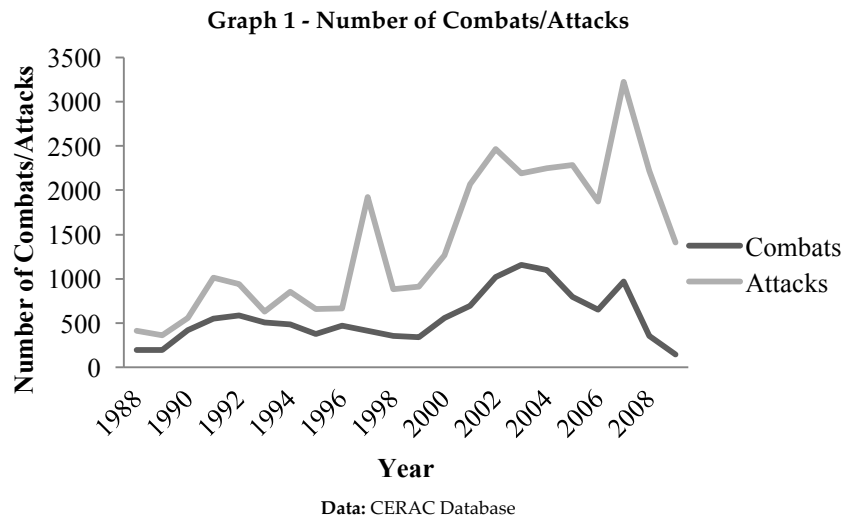
The rest of this chapter is organized as follows: section 2 presents a brief description of the Colombian conflict during the “escalation period”; section 3 presents the theoretical framework; section 4 describes the data and the variables that are being used in the analysis; section 5 describes the methodology employed for the analysis and discusses the specification of the model; section 6 reports the main results; section 7 presents a discussion of the results; section 8 discusses possible policy implications of this analysis; and section 9 presents the conclusions.

2. Colombian Conflict During the “Escalation Period”

The Colombian conflict is considered to be a long duration conflict with variable intensity. As mentioned in Chapter 2, two violent actors have been the major players behind the conflict: the guerrilla groups and the paramilitary groups. These illegal groups went through a period of re-organization between 1991 and 1994 because of a major cut back in available resources due to the end of the Cold War. This re-organization led to changes in their operative and financial structure. More specifically, during the first half of the nineties these illegal groups expanded their territorial presence throughout the country. However, according to Restrepo et al. (2006) during these years violence did not increase. It was only in the year 1996 that the dynamic of conflict started to change significantly, meaning that the re-organization that took place during the first half of the nineties had a lagged effect on this dynamic. According to Restrepo et al. (2006), during this period an important increment in the intensity of conflict, and also in the frequency of combats and attacks took place (Graph 1). The re-organization period gave a new air to these illegal groups, allowing them to grow and gain power.

Restrepo et al. (2006) have called this period the escalation of violence period given the increase in the intensity of the conflict and the frequency of the combats and the attacks. According to these authors, the escalation period took place between the years 1996 and 2002, year in which the armed conflict reached the highest intensity. The recrudescence of violence in Colombia during those years could be explained by a number of factors. According to Restrepo et al. (2006) some of the factors that should be highlighted are: the unification of the majority of paramilitary groups into one organization (Autodefensas Unidas de Colombia), the implementation of a terror strategy against civilians, and the modernization of the Colombian military forces through the implementation of new technologies in order to have a better response against the attacks of illegal groups. It should also be noted that during the escalation period, the illegal groups implemented a strategy that focused in attacking civilians as an instrument to terrorize the society. By the end of the escalation period, the number of civilian fatalities started

to decrease, although the number of combats and attacks remained high (Graphs 1 and 2).



In this chapter, the relationship between civil conflict and educational achievement will be explored concentrating on the time period identified above when the affectation to civilians was at its peak (2001 – 2002). The idea behind the analysis done in this chapter is to verify whether there is a significant relationship between the intensity of conflict during this period and the achievement of students in standardized exams.

3. Theoretical Framework

The economics of education literature has traditionally modeled schooling decisions (quantity and quality of education) using an educational production function. Let's consider the education production function specification proposed by Hanushek (1979) and Todd and Wolpin (2003). Following Todd and Wolpin's are given by the choices that are being made by parents and schools. Following Todd and Wolpin's (2003) notation, the vector of inputs supplied by the parents at is represented as:

Missing data for past inputs poses an important problem at the time of estimation. In order to overcome that problem, Todd and Wolpin (2003) came up with a specification that only depends on contemporaneous measures of school and family inputs. This "contemporaneous" specification is often times adopted because of data limitations associated with limited availability of historical input measures or historical test score results. Behind the implementation of a includes any omitted factors, including past inputs and endowed capacity, and the possibility of measurement error. Therefore, in equation (1) the residual term includes all the omitted factors. That implies that although this specification has less data requirements, it requires strong assumptions in order to justify its application (Todd and Wolpin 2003).

It should also be acknowledged that an important problem that comes up in the specification of an education production function to be estimated empirically is the non-existence of a valid measure of innate abilities. As Hanushek (1979) highlights, the most consistent and obvious divergence between the empirical specification and the theoretical educational production function is the non-existence of a measure of innate or endowed abilities. The absence of a valid measure of innate abilities can be problematic because it can bias the estimated regression coefficients. The size of the bias depends on how strong is the effect of the variable on achievement and also on the correlation of the omitted variable with the other variables that are included in the model. Omitting the endowed capacity of the student from the empirical specification will likely cause an upward bias in the

estimated impact of the other variables on achievement (Hanushek 1979). The possibility of this upward bias should not be ignored, and therefore, it should be considered carefully in the interpretation of the results of the empirical exercise.

4. Data and Variables

Two cross-sectional datasets, containing the nation-wide results of the Colombian High School exit examination from the second semester of the year 2001 and the second semester of the year 2002, are employed in the analysis. These two cross-sections were put together using information from three different sources. Information about Colombian exit examinations scores and the socioeconomic characteristics of students taking the examinations was obtained from the Colombian Institute for the Evaluation of Education (*Instituto Colombiano para la Evaluación de la Educación – ICFES*), educational data about schools and teachers was obtained from the Colombian National Department of Statistics (*Departamento Administrativo Nacional de Estadística - DANE*), and conflict data was obtained from the Resource Centre for Conflict Analysis (*Centro de Recursos para el Análisis del Conflicto – CERAC*).

The information contained in the datasets has a multilevel structure, as depicted in *Table 1*. For the year 2001, the dataset comprises the results from 268842 students, coming from 5262 schools in 1010 municipalities. On the other hand, for the year 2002 the dataset encompasses the results from 284492 students, enrolled in 6167 schools in 1011 municipalities.

Table 1 – Number of Observations		
Level	2001-2	2002-2
Student	268842	284492
School	5262	6167
Municipality	930	1011

Source: author's calculations

4.1 Main Sources of Information

The C-600 and C-100 Censuses – National Department of Statistics

With the support of the Ministry of Education, the Colombian National Department of Statistics (*DANE*) has been collecting information about schools,

students and teachers since 1978. This information has been collected through the C-100 and C-600 statistical censuses.

The C-100 is a statistical census with an annual periodicity, which was carried out by the Colombian National Department of Statistics (*DANE*) until the year 2002, in order to collect information about school infrastructure and the supplies available at the schools. On the other hand, the C-600 is a statistical census with an annual periodicity that collects information about students and teachers from all formal education schools in Colombia. Information for the C-600 census is still being collected (*DANE* 2010).

Databases containing the nation-wide census information for the C600 and the C100 censuses for the years 2001 and 2002 were made available by the Colombian National Department of Statistics (*DANE*) for this thesis. The analysis will therefore be restricted to these two years. As it was discussed in *section 2*, the 2001 – 2002 period is a particularly relevant period of study to analyze the relationship between civil conflict in Colombia and other social variables, including educational performance, given the peak levels of intensity of the armed conflict during that time frame.

Colombian High School Exit Examination – ICFES Saber 11

The Colombian nation-wide exit examination (ICFES Exam) was first created more than forty years ago in order to provide support to universities and other higher education institutions in their admissions processes⁹. Databases containing the nation-wide results of the Colombian high school exit exams (Saber11 exams), as well as information about socioeconomic variables of the students that took the exam, have been made public by ICFES for research purposes. The database for the Saber11 exam covers the years 2000 to 2012. However, as mentioned previously, in this empirical exercise we are only going to analyze the years 2001 and 2002.

CERAC's Colombian Conflict Dataset

Information about conflict events was obtained from the Resource Centre for Conflict Analysis (CERAC). CERAC is a think tank based in Bogotá (Colombia) with three main lines of research: violent conflict, political violence, and violence associated to drug trafficking. Following these lines of research, an important area of study at CERAC is the quantification of violence directly associated to the internal armed conflict in Colombia. To quantify violent armed conflict, the research team at CERAC developed a methodology to collect, document, code and

⁹ More information about this examination and the ICFES dataset was provided in Chapter 6 of this dissertation.

register information about this type of violence. The construction of CERAC's database started in the year 2002. This database contains information about the violent actions, the event in which these actions occur, the victims of the event, the attackers and the group to which they belong. CERAC's methodology is explained in detail in *Appendix B* of this chapter.

In this chapter we are going to use a measure of the intensity of the Colombian conflict to understand the relationship between civil conflict and educational achievement. Specifically, the number of casualties reported in CERAC's database is going to be used as a municipal-level measure of the intensity of conflict.

4.2 Variables

Table 2 summarizes all the variables considered in the analysis. The variables are presented in the table following the hierarchical structure of the data. A description of these variables is provided in what follows.

Table 2 Variable Selection		
Dependent Variable		Score in mathematics and language high school exit exam (high school exit exam)
Explanatory variables at the different levels	<i>Level 1 (student level)</i>	Gender, age, socio-economic variables (household ownership, number of people contributing to the household income, monthly income of the household, level of education of the father, level of education of the mother, number of brothers/sisters, number of family members, number of family members supported by the income of the household, house mortgage), student works.
	<i>Level 2 (School level)</i>	Number of teachers available per student, school basic infrastructure (water, sewage, and energy service), cost of tuition, number of students enrolled, school meeting time (day/night), lab space per student, classroom space per student.
	<i>Level 3 (Municipality level)</i>	Intensity of conflict, area of the municipality in KM ² (log), resources assigned by the central government to education (log), unsatisfied basic needs index, degree of rurality, population of the municipality (log), altitude (log) and distance to the capital city (log).

Source: author

The Dependent Variables

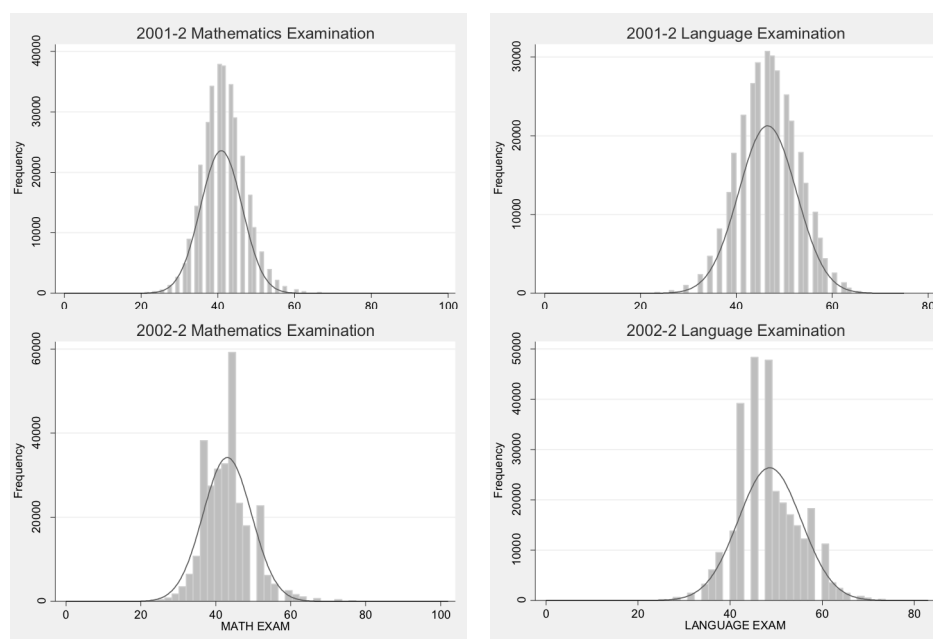
Our analysis focuses on two outcome variables: the score in the mathematics exam and the score in the language exam. Even though the results of the examinations in other subjects are available, we have decided to restrict the analysis to mathematics and language because these two subjects capture two very important competencies: quantitative ability and verbal ability. However, this quantitative analysis could potentially be extended to study other subjects that are also tested in the Colombian High School exit examination. In fact, the results obtained by the students in the other examinations will also be considered to test the robustness of the results of the final models for the mathematics and language exams.

Table 3 presents the statistics of the scores obtained by students in the mathematics and language High School exit examinations. The scoring scale in this exam goes from 0 to 100. The average score in the mathematics examination for the year 2001 is 41.12 and for the year 2002 is 43.21, with a standard deviation of 5.34 and 6.34, respectively. In the case of the language examination, the average score for the year 2001 is 46.82 and for the year 2002 is 48.76, with a standard deviation of 5.97 and 6.81, respectively. Graph 3 illustrates the distribution of the scores in the mathematics and language examinations for the two years that are under analysis.

Table 3 – Language and Math Scores in the High School Exit Examination (mean and standard deviation)				
	2001-2		2002-2	
	Mean	Std. Dev.	Mean	Std. Dev.
Math Score	41.12	5.34	43.21	6.34
Language Score	46.82	5.97	48.76	6.81
Observations	268842		284492	

Source: author's calculations

Graph 3 – Distribution of the Language and Math Scores in the Colombian High School Exit Examination



Source: author's calculations using data from ICFES

Student Level Variables

According to the literature, family and individual background variables play a crucial role on educational performance (Häkkinen et al. 2003, Haveman and Wolfe 1995, Woßmann 2003). Family inputs are often times measured by the socio-demographic characteristics of the families. In this chapter, nine different variables are included to account for family background: household ownership, number of people contributing to the household income, monthly income of the household, level of education of the father, level of education of the mother, number of brothers/sisters, number of family members, number of family members supported by the income of the household, and whether the family has a house mortgage. A number of these variables (household ownership, number of people contributing to the household income, monthly income of the household, and whether the family has a house mortgage) are related to the sources of income of the family. The literature has shown that family income is related to educational achievement. The research done by Belley and Lochner (2007) and Davis-Kean (2005) has shown that relationship by examining cross-sectional and longitudinal datasets. The

inclusion of the parental education variables (level of education of the father and level of education of the mother) is also particularly important considering that studies, including those of Woßmann (2003) and Häkkinen et al. (2003), have found that students whose parents have a higher level of education have a better performance than those students whose parents have a low level of education. This could be the case because of genetic transmission of skills from parents to kids, and because highly educated parents tend to provide a good cultural environment that is adequate for learning. Family size also plays an important role on educational performance. To account for that two variables were considered: number of brothers/sisters and number of family members. Downey (1995) and Blake (1981, 1989) have found that resource dilution as the family size increases causes an inverse relationship between the number of siblings and educational performance. This finding is explained by the fact that parents have limited resources (time, energy, money, etc.) that have to be divided among children as the size of the family increases (Downey 1995).

Table 4 presents information about the educational background of the parents of the students in the sample through the variables *level of education of the student's father* and *level of education of the student's mother*, which can go from no formal studies to completion of graduate school. On average, around 29% of the mothers had completed at least High School, while only around 25% of the fathers had completed that same level of schooling. However, around 7% of the mothers had completed a university degree, whereas around 9% of the fathers had completed a university degree. This is particularly interesting because comparatively more women complete High School, but on other hand, more men tend to complete a university degree. For our sample, a higher percentage of fathers have completed the highest level of education, which is graduate school (around 4%). For the econometric exercise presented in this chapter, we have decided to divide the parental education variable into four categories comprising the different levels of education discussed above. The four categories are the following: no education (parent didn't complete any studies), low level of education (elementary school), mid-level of education (High school and/or technical studies) and high level of education (university and/or graduate school).

Three additional variables were also considered to account for individual characteristics: gender, age and a dummy to capture whether the student works or not. *Table 5* contains information about the gender distribution of the sample. There is a higher percentage of female students in the sample in both years. In the year 2001, 45.4% of the students taking the examination were males and 54.6% were females. In the year 2002, we have a similar trend with 45.7% of the sample being males and 54.3% being females.

Table 6 contains information about students who work. There is a higher percentage of male students in the sample who work in both years. In the year 2001, 7.37% of the male students had a job, while 4.70% of the female students did. For the year 2002, we can observe a similar trend with 10.70% of the male students having a job, while only 6.84% in the case of female students. In both years we can see that more male students participate in the labor market than female students.

Table 7 summarizes all the other relevant family background variables, including the minimum and maximum values and the mean and standard deviation.

Table 4 - Parental Education								
	Mother Education				Father Education			
	2001-2		2002-2		2001-2		2002-2	
	# Students	%	# Students	%	# Students	%	# Students	%
Didn't complete any studies	9995	3.97	10412	3.92	12227	5.09	12591	4.97
Kindergarten completed	2719	1.08	3036	1.14	2602	1.08	2910	1.15
Elementary completed	107978	42.9	112965	42.53	102180	42.53	107301	42.3
High school completed	72933	28.9	77092	29.02	61188	25.47	65140	25.7
Vocational School completed	19984	7.94	21245	8.0	15656	6.52	16898	6.47
Tech. program completed	11734	4.66	12813	4.82	13765	5.73	14527	5.33
University completed	17887	7.11	19137	7.2	22864	9.52	23916	9.03
Graduate school	8464	3.36	8933	3.37	9747	4.06	10242	4.04

Source: author's calculations

Table 5 – Gender				
	2001-2		2002-2	
	Number of Students	%	Number of Students	%
Male	122234	45.43	130310	45.67
Female	146800	54.57	155020	54.33

Source: author's calculations

Table 6 – Students Who Work				
	2001-2		2002-2	
	Number of Students	%	Number of Students	%
Male	8544	7.37	13343	10.70
Female	6633	4.70	10234	6.84

Source: author's calculations

Table 7 – Family Background Variables								
	2001-2				2002-2			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Age	17.56	2.72	12	73	17.8	3.36	12	67
Household ownership (dummy: 1→owns house)	0.74	0.44	0	1	0.72	0.45	0	1
# of people contributing to the household income	1.59	0.81	0	10	1.59	0.78	0	9
# of family members	5.28	1.80	1	10	5.05	1.56	1	9
# of brothers/sisters	2.93	2.12	0	10	2.77	1.88	0	9
# of family members supported by household income	0.71	1.03	0	10	0.73	1.02	0	9
House mortgage (dummy: 1→ no debt)	0.23	0.42	0	1	0.25	0.43	0	1

Source: author's calculations

School level Variables

The academic literature has also found evidence showing that the characteristics of the schools and teachers have a significant role in determining educational achievement. In fact, Gamoran and Long (2006) have found that the characteristics of the school have an important impact on academic performance in developing countries. However, the existing literature has not reached consensus as to which variables have a significant impact and which do not. Considering that the existing literature has not reached consensus about which variables are significant, it is very important to control for the characteristics of the schools and the teachers that are considered to be relevant for the sample under study, as these might have an effect on the achievement of the students. In this chapter eight different variables are included to account for school characteristics: number of students per teacher, school basic infrastructure (water, sewage, and energy service), cost of tuition, number of students enrolled, school meeting time (day/night), lab space per student, and classroom space per student. *Table 8* summarizes the school level variables considered in the analysis, including the minimum and maximum values and the mean and standard deviation.

Hanushek and Luque (2003) did not find a significant impact of the school size on educational achievement, while Barnett et al (2002) found a significant impact between the two variables. This shows that the results are not pointing to the same conclusion in all cases. In the case of those studies that have tried to establish a relationship between private/public schools and academic achievement, including Hanushek (1986), Vandenberghe and Robin (2004), and Somers et al. (2004), the results are not conclusive either. The literature has also highlighted the importance of school facilities as determinants of academic outcomes (Schneider 2002). The justification for this is particularly simple: how can students be high achievers if schooling infrastructure is not adequate? The research done by Earthman and Lemasters (1996, 1998) has found that there is a positive relationship between the available infrastructure and higher test scores. Earthman (2002) points out that school building components have a measurable influence on student learning. Indeed, according to his study, “in cases where students attend school in substandard buildings they are definitely handicapped in their academic achievement...correlation studies show a strong positive relationship between overall building conditions and student achievement” (Earthman 2002).

The number of teachers per student available in the school is also included as an explanatory variable. If a large group of students receives class from only one teacher, most likely the attention received will be less and this might have a negative effect on the achievement of students in evaluations. Additionally, the

low number of teachers per student may be a factor discouraging school attendance. Nevertheless, the literature has not shown conclusive results when this variable is included as a determinant. Harbison and Hanushek (1992) reviewed the literature of the impact of teacher-pupil ratios and found that out of sixteen studies with statistically significant effects, eight were positive and eight were negative. These results cast doubt about the reliability of previous studies, but serves as an additional motivation to include this variable as part of this research.

Table 8 – School Level Variables								
	2001-2				2002-2			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
# of students per teacher	180.92	114.95	0	500	188.68	116.89	6.29	500
Sewage service (dummy: 1→ school has the service)	0.90	0.30	0	1	0.88	0.33	0	1
# of students enrolled	2016.6	2038.6	50	14592	1869.45	1924.33	62	13118
School meeting time (dummy: 1→ school meets during the day; 0→ meets at night)	0.93	0.248	0	1	NA	NA	NA	NA
Lab space per student (m ²)	0.21	0.80	0	27.3	0.23	1.21	0	38.9
Classroom space per student (m ²)	1.53	2.20	0.5	37.14	1.41	2.01	0.5	34.6

Source: author's calculations

Table 9 presents information about the tuition that has to be paid in the different schools that are included in the dataset. For the econometric exercise presented in this chapter, we have decided to divide the cost of tuition variable into four categories comprising the different levels of tuition paid by the students in the sample. The four categories are the following: no tuition (students don't pay), low tuition (students pay less than \$25 dollars per month), mid tuition (students pay between \$25 dollars and \$100 dollars per month) and high tuition (students pay more than \$100 dollars per month). Most of the schools are either free (43% and

35% of the total in 2001 and 2002, respectively) or have a tuition that costs more than US\$100 (32% and 36% of the total in 2001 and 2002, respectively). This observation can be explained by the dichotomy between public and private education. Public education in Colombia in most of the cases is free, while most of the private schools tend to have a high tuition.

Table 9 – Cost of tuition				
	2001-2		2002-2	
	Number of Students	%	Number of Students	%
Tuition Level – No Tuition	81228	42.60	96261	35.32
Low Tuition	24050	12.61	34462	12.65
Mid Tuition	25089	13.15	44055	16.16
High Tuition	60325	31.63	97747	35.87

Source: author's calculations

Municipality Level Variables

Municipality level data was obtained from a number of different sources. As mentioned previously, conflict data was obtained from CERAC. The additional information that was used in order to construct municipal controls was obtained from the National Department of Statistics (DANE), the Colombian Federation of Municipalities (Federación Colombiana de Municipios), and the Instituto Geográfico Agustín Codazzi. Table 10 summarizes the municipality level variables, including the mean and standard deviation, considered for the analysis.

Central government resources are an important source of income for schools. Every year each municipality receives resources that are assigned by the General System of Participation (Sistema General de Participaciones)¹⁰. In order to capture that fact, and also to account for financial differences between municipalities, a variable that captures the resources that are transferred by the central government to the territorial entities to finance education expenditure is included in the model. This variable was transformed to logarithms in order to be included in the specification that was used for the estimation.

¹⁰ Refer back to Chapter 3 for more information about the General System of Participation.

Table 10 – Municipality Level Variables								
	2001-2				2002-2			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Financial Resources Assigned by Gov. to Educ. (thousand of COP¹¹)	1.84e+08	2.61e+08	140	6.79e+08	1.62e+08	2.73e+08	393.5	6.79e+08
Degree of “rurality”	0.22	0.272	0	1	0.20	0.26	0	1
Unsatis. Basic Needs	24.91	18.9	5.43	100	24.51	18.65	5.43	100
Size (KM²) of Mun.	993.6	1599.87	140	65674	973.08	1382.96	139.7	65674
Popul. of Mun.	1752389	2585524	1153	6573291	1741020	2598881	1095	6712247
Total Expend. (Millions of COP)	870795	1459591	61.9	3616924	8.92e+08	1.56e+09	2738	3.86e+09

Source: author's calculations

Municipalities can be characterized according to the percentage of rural population that they have (rurality). To calculate the degree of “rurality” of each municipality, a “rurality” index was calculated by dividing the rural population over the total population of each municipality. This “rurality” index goes from 0 to 1. It is important to include this variable in order to account for differences between the rural and urban composition of each municipality. The unsatisfied basic needs index at the municipality level is also included in order to capture differences in the poverty levels of the different municipalities considered in the analysis. The unsatisfied basic needs index is a poverty index that is calculated by estimating the percentage of the population with basic unsatisfied needs.

¹¹ Colombian Pesos

As part of the analysis it is also important to account for geographic characteristics. Two variables were included in the model in order to control for the size and the population density of the municipalities. The first variable captures the area of the municipality in KM² and the second variable captures the population of the municipality. The altitude and the distance to the capital of the departamento (kilometers) were also included to capture geographical characteristics that could also affect the performance of students in the examination.

International literature has shown that the economic, political and social consequences of civil war are significant (Justino 2010). However, the connection between conflict and education is an under-analysed area of study. This is partly due to the difficulties involved in the systematic collection of evidence on why, where and how attacks occur. As mentioned by Justino (2010), high quality monitoring data and systematic reporting of events are very limited. Additionally, in many places affected by conflict censorship of information is also common. The availability of a rich database about conflict events in Colombia, maintained by the Resource Centre for Conflict Analysis (CERAC), opens new possibilities for the analysis of the relationship between conflict and education. Using information from CERAC, we have included a variable that captures the intensity of the conflict as the central component of this empirical analysis.

Table 11 – Correlation between Scores and the Intensity (deaths) of the Conflict

2001-2			2002-2		
Correlation	Math Exam	Language Exam	Correlation	Math Exam	Language Exam
Intensity 2000	0.0289	0.0408	Intensity 2001	0.0497	0.0706
Intensity 2001	0.0467	0.075	Intensity 2002	-0.1093	-0.1413
Average intensity last 5 years	0.0474	0.0693	Average intensity last 5 years	0.0777	0.1015

Source: author's calculations

The correlations between the scores in the two examinations and the intensity of the conflict are presented in *Table 11*. The descriptive statistics for the years 2001 and 2002 consistently show a positive correlation between the intensity of conflict and the scores in the examinations. The only exception is the negative correlation that was found between the intensity of the conflict in the year 2002 and the results of the examination in that same year. Even though a positive correlation does not

necessarily imply a positive causal relationship between the variables under study, it indeed motivates a further exploration of this relationship. In fact, my hypothesis is that those students that receive a formal education in a conflict zone are not as negatively affected as intuition might suggest. My hypothesis is explained by two channels. First, the parents of those students that stay in conflict zones are most likely to invest as many resources as needed to give their kids a good education that can keep them away from the conflict. The families living in conflict areas are often faced with a very difficult dilemma: stay in their hometown or move to a non-conflict area. Those families that decide to stay are interested in the safety of their kids, and to do so, they will take as many actions to keep the kids away from the conflict and from conflict groups. One of those actions will likely be to follow up as close as possible the educational process of their sons and/or daughters. Such effort could translate in levels of performance that are similar (or possibly higher) to those of kids studying in non-conflict zones. Second, there might be an auto-selection channel in action. It could be the case that those school-aged individuals that are mostly affected by conflict are the ones that do not participate in the formal education system. Those individuals that do not have the resources (monetary, psychological, family support) to overcome or adapt to the rough conditions that are prevalent in conflict areas are more likely to drop out from the formal educational system. There is auto-selection because those individuals with resources to overcome or adapt to the harsh conditions are most likely to continue with their studies. As a consequence, I suspect that the levels of performance of those students in conflict zones are not significantly negatively affected.

5. Methods

5.1 Cluster Robust Ordinary Least Squares

In order to get a first estimate of the relationship between conflict intensity and the achievement in the Colombian exit examination we are going to use an ordinary least squares regression model. However, in order to take into consideration the hierarchical structure of the dataset we will employ a cluster robust version of the model that corrects the standard errors for within-cluster error correlation. Specifically, in our case we have student-level cross-sectional data with school clustering. According to Cameron and Miller (2013), failure to control for within-cluster error correlation can lead to misleadingly small standard errors, and therefore misleadingly narrow confidence intervals, large t-statistics and low p-values. Clustered errors have two principal repercussions: there is a reduction in the precision of $\hat{\beta}$ and the estimator for the variance of $\hat{\beta}$ is usually biased downward from the true variance (Cameron and Miller 2013).

5.2 Multilevel Analysis

Multilevel analysis is employed in this chapter as our preferred method to deal adequately with the hierarchical structure of the data by simultaneously modeling variables at different levels without resorting to aggregation or disaggregation of the dataset. Only analyzing student level factors would be insufficient to identify what is affecting the performance of students in standardized exams. Therefore, given the structure of the dataset, a three level multilevel model is used for our analysis (*diagram 1*).

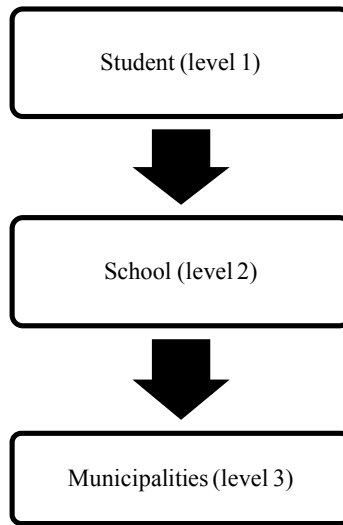
By using multilevel modeling we are assuming that mediating mechanisms exist, and these mechanisms cause variables at one level to influence variables at another level (Garson 2012). The use of individual level techniques for the analysis of a dataset with a hierarchical structure will likely underestimate the standard errors of these macro-level effects (Raudenbush and Bryk 2002). The use of biased standard errors to carry out statistical tests or construct confidence intervals is very problematic because it might lead to misleading results.

An important assumption of a single level model, e.g. multiple regression models, is independence of the measured units¹². However, according to Steele (2008), if we have a grouped data structure and the group effects are not taken into account in the regression model, the independence assumption will be violated.

However, the problem of using aggregated or disaggregated data is not limited to the issue of making misleading inferences. As Garson (2012) highlights, by aggregating and disaggregating data used in regression models there is a risk of running into an ecological fallacy, i.e. what is true at one level is not necessarily true at another level. As Goldstein (2011) points out, if both students and schools are not modeled simultaneously it is impossible to study the extent to which school and student characteristics interact to influence the response measurement or measurements. An approximation of that quantification can only be done through the use of a multilevel model (Goldstein 2011).

¹² The single level models assume that the residuals are uncorrelated

Diagram 1



Source: author's elaboration

5.3 Instrumental Variables Approach

It is credible to think that a student takes into consideration the intensity of conflict when deciding whether to study or not. Some students might have special characteristics (i.e. support from parents) that make them more able to face conflict-related violence, and continue with their studies. As a consequence, it is necessary to consider the possibility of endogeneity of the intensity of conflict measure. To tackle this issue, and in order to test the robustness of the results obtained using multilevel analysis, an instrumental variables approach is considered in this section. Following Camacho and Rodriguez (2013), we are going to use instrumental variables in order to deal with the possibility of endogeneity as a source of bias. The inconvenience associated with the use of instrumental variables in this case is that we would be omitting the hierarchical structure of the dataset, which has been shown to be relevant in the literature that studies students' achievement in standardized examinations. Nevertheless, using an instrumental variables approach is fundamental in testing the robustness of the results obtained using a multilevel analysis approach in our case.

Two instruments, which were proposed by Camacho and Rodriguez (2013) to explain the presence of armed conflict in municipality m at time t , are used to test

the robustness of the results obtained in the previous section. The two instruments proposed are: lagged laboratories dismantle and antinarcotics operations. Camacho and Rodriguez (2013) suggested these instruments for two reasons. First, the dismantling of laboratories and the antinarcotics operations rate in municipality m at time $t-1$ should be highly correlated with armed conflict, because these two measures indicate the presence of the government to offset the activities of illegal groups in the region (Camacho and Rodriguez 2013). The intuition behind this is that there should be a negative relation between these dissuasion techniques employed by the government and armed conflict intensity. Second, it seems very unlikely that the schooling decisions of students today will be based on the two governmental deterrence decisions that are being used as instruments. As Camacho and Rodriguez (2013) highlight, antinarcotics operations and the dismantling of laboratories are normally secret operations that take place under conditions that are not discovered by the civil population.

The technical requirements for a good set of instruments are also verified empirically. In this case, I was able to verify that the two deterrence measures are strongly related with the intensity of the conflict, and additionally these two measures are exogenous to the schooling decisions taken by the students.

We tested two different specifications in this case. In the first specification we include all the control variables and we instrument the intensity of the conflict in the year under analysis¹³ using the lagged measure (t-1) of laboratories dismantle and antinarcotics operations. In the second specification we include all the control variables and we instrument the intensity of the conflict in the year before the students took the National Examination using the lagged measure (t-1) of laboratories dismantle and antinarcotics operations.

6. Results

6.1 Cluster Robust Ordinary Least Squares

Tables 12 and 13 present the ordinary least squares regression (cluster robust) results of the relationship between conflict intensity and achievement in the Colombian exit examination. We considered four different specifications. The first specification includes all the control variables mentioned in *Table 3* but it only uses a contemporaneous measure¹⁴ of the intensity of the conflict. The second specification includes all the control variables, a contemporaneous measure of the

¹³ The year in which the students took the National Examination.

¹⁴ Intensity of conflict for the year in which the central examination was administered.

intensity of the conflict and the intensity of conflict at the municipality for the previous year. The third specification includes all the control variables, a contemporaneous measure of the intensity of the conflict and the intensity of conflict at the municipality for the previous 5 years (separate variable per year). The fourth specification includes all the control variables, a contemporaneous measure of the intensity and the average measure of intensity of conflict for the last 5 years.

Table 12 – Ordinary Least Squares Model (cluster robust) – Mathematics and Language 2001								
2001-2	Model 1 Math	Model 2 Math	Model 3 Math	Model 4 Math	Model 1 Lang.	Model 2 Lang.	Model 3 Lang.	Model 4 Lang.
Intensity 2001	0.037 (0.035)	0.038 (0.035)	0.033 (0.035)	0.038 (0.036)	0.099*** (0.042)	0.100*** (0.042)	0.090** (0.041)	0.099** (0.042)
Intensity 2000		-0.001 (0.001)	0.0004 (0.002)			-0.001 (0.003)	-0.001 (0.003)	
Intensity 1999			-0.005 (0.014)				0.006 (0.017)	
Intensity 1998			0.005 (0.009)				0.006 (0.011)	
Intensity 1997			-0.030** (0.013)				-0.023 (0.015)	
Intensity 1996			0.018*** (0.007)				0.009 (0.008)	
Avg. Intensity Prev. 5 Yrs.				-0.001 (0.007)				0.0003 (0.011)
R ²	0.061	0.061	0.062	0.061	0.1487	0.1487	0.1489	0.1487

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

As it can be observed in *Table 12*, the results for the year 2001 show no significant relationship between our contemporary measure of intensity of the conflict and the performance in the mathematics examination. We only found a significant relationship between two of the lagged measures of intensity of the conflict and the performance in the mathematics examination (*model 3*). However, the relationship found is not conclusive. The results show a significant negative relationship between the intensity of the conflict in 1997 and the outcome in the mathematics examination, and a significant positive relationship between the

intensity of the conflict in 1996 and the outcome in the mathematics examination. In the case of the language examination, we found a significant positive relationship between the contemporary measure of intensity of the conflict and the performance in the examination. None of the lagged measures of intensity that were considered in the different specifications turned out to be significant.

Table 13 – Ordinary Least Squares Model (cluster robust) – Mathematics and Language								
2002								
2002-2	Model 1 Math	Model 2 Math	Model 3 Math	Model 4 Math	Model 1 Lang.	Model 2 Lang.	Model 3 Lang.	Model 4 Lang.
Intensity	-0.015	-0.018	-0.039	-0.026	0.003	0.004	0.014	0.005
2001	(0.041)	(0.040)	(0.041)	(0.040)	(0.047)	(0.047)	(0.047)	(0.047)
Intensity		0.005	0.001			-0.001	0.0001	
2000		(0.006)	(0.005)			(0.007)	(0.007)	
Intensity			0.0002				0.004	
1999			(0.002)				(0.003)	
Intensity			0.002				-0.007	
1998			(0.017)				(0.024)	
Intensity			0.020*				-0.003	
1997			(0.012)				(0.014)	
Intensity			-0.014				-0.024*	
1996			(0.012)				(0.013)	
Avg. Intensity Prev. 5 Yrs.				0.017 (0.013)				-0.003 (0.014)
R ²	0.1333	0.1334	0.1339	0.1335	0.178	0.178	0.1782	0.178

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 13 presents the results for the year 2002. The results show no significant relationship between our contemporary measure of intensity of the conflict and the performance in the mathematics and the language examination in any of the specifications considered. In the case of the mathematics examination, we only found a significant relationship between one of the lagged measures of intensity of the conflict (year 1997) and the performance in the mathematics examination (*model 3*). In the case of the language examination, we found a significant negative

relationship between the measure of intensity of the conflict in the year 1996 and the performance in the examination.

It should be highlighted that the results obtained in this section could be subject to bias because of problems associated to misspecification and endogeneity. Those two problems are very common in OLS regression analysis. In order to deal with these possible sources of bias, we used two additional methods to get an estimate of the effect. The results obtained using multilevel analysis and an instrumental variables approach are presented in the next two sub-sections.

6.2 Multilevel Analysis Results

As mentioned previously, we use multilevel analysis as our preferred method to deal adequately with the hierarchical structure of the data by simultaneously modeling variables at different levels without resorting to aggregation or disaggregation of the dataset. The results of the three level variance components (fully unconditional) model will be discussed first in order to establish the relevance of using a multilevel model. After justifying the use of a three level multilevel model, the results of the complete model will be discussed.

A. *Variance Components Model*

For the second semester of 2001 (*Table 14*), the overall mean score in mathematics (β_0) is estimated to be 40.401, and the variance 28.129. The variance can be decomposed in variance between municipalities (0.837), variance between schools (2.237) and variance between students (25.055). In terms of the VPC statistic, the results show that 3% of the variation in the mathematics test lies between municipalities, 8% lies between schools and 89% lies between students. An LR test was applied in order to check whether the three level model is preferred to a single level model. The results¹⁵ obtained from the LR test show that the three level model is preferred to a single level model ($\chi^2_2 = 19556.23$, $p < 0.000$), to the two level students-within-municipality model ($\chi^2_1 = 13654.056$, $p < 0.001$) and the two level students-within-schools model ($\chi^2_1 = 649.21$, $p < 0.001$). These results imply that students from the same municipality are significantly more alike than students from different municipalities. Likewise, students from the same school are significantly more similar than those students from two different schools. As a consequence, a three level multilevel model should be used to analyze this dataset instead of a two level or single level model.

¹⁵ The complete results for the LR test are reported in the appendix.

Table 14		
Three Level Variance Components Model		
Period 2001 – 2	Model 1 – Mathematics	Model 2 - Language
Parameter	Estimate	Estimate
β_0 intercept	40.401	45.042
Variance	0.837	1.496
	2.237	4.926
	25.055	28.299

Source: author's calculations using the databases from ICFES, DANE and CERAC.

The results obtained for the language exam dataset in the second semester of 2001 also justify the use of a three level multilevel model. The overall mean score in the language exam (β_0) is estimated to be 45.042, and the total variance 34.721. The variance can be decomposed in variance between municipalities (1.496), variance between schools (4.926) and variance between students (28.299). The VPC statistic shows that 4.3% of the variation in the language test lies between municipalities, 14.2% within schools, and 81.5% within students. The LR test shows that for the language exam the three level model is preferred to a single level model ($\chi^2_2 = 51093.61$, $p < 0.000$), to a two level students-within-municipality model ($\chi^2_1 = 31392.3$, $p < 0.001$) and to a two level students-within-schools model ($\chi^2_1 = 1371.3$, $p < 0.001$). Just like in the case of the mathematics examination, these results suggest that students from the same municipality are significantly more alike than students from different municipalities and students from the same school are significantly more similar than those students from two different schools. Given these results, the use of a three level multilevel model is justified.

Table 15		
Three Level Variance Components Model		
Period 2002 – 2	Model 1 – Mathematics	Model 2 - Language
Parameter	Estimate	Estimate
β_0 intercept	42.008	47.029
Variance	1.180	2.091
	4.590	7.812
	33.539	36.090

Source: author's calculations using the databases from ICFES, DANE and CERAC.

For the second semester of 2002 (*Table 15*), the overall mean score in mathematics (β_0) is estimated to be 42.008, and the variance 39.309. The variance can be decomposed in variance between municipalities (1.180), variance between schools (4.590) and variance between students (33.539). In terms of the VPC statistic, the results show that 3% of the variation in the mathematics test lies between municipalities, 11.7% lies between schools and 85.3% lies between students. In this case, the results obtained from the LR test also show that the three level model is preferred to a single level model ($\chi^2_2 = 42598.64, p < 0.000$), to the two level students-within-municipality model ($\chi^2_1 = 28446.248, p < 0.001$) and the two level students-within-schools model ($\chi^2_1 = 1034.263, p < 0.001$). Given these results, a three level multilevel model should be used to analyze this dataset instead of a two level or single level model. For that same time period, in the case of the language exam dataset we can also justify the use of a three level multilevel model. The overall mean score in the language exam (β_0) is estimated to be 47.029, and the total variance 45.993. The variance can be decomposed in variance between municipalities (2.091), variance between schools (7.812) and variance between students (36.090). The VPC statistic shows that 4.5% of the variation in the language test lies between municipalities, 16.9% within schools, and 78.6% within students. The LR test shows that for the language exam the three level model is preferred to a single level model ($\chi^2_2 = 63489.74, p < 0.000$), to a two level students-within-municipality model ($\chi^2_1 = 42128.7, p < 0.001$) and to a two level students-within-schools model ($\chi^2_1 = 1301.5, p < 0.001$). Given these results, the use of a three level multilevel model is justified.

B. Random Intercept Model

The results discussed in the previous section did not make any adjustments for predictor variables because the objective of the variance components model is to decompose the total variance into level specific variance components. However, in this section we will analyze the results of a random intercept model that includes predictor variables at the three different levels. By doing so, this more complete model will help us get a better understanding of the variation at each level and the relationship between the variables included in each specification and the performance in the mathematics and language exams (outcome variables). Four different specifications were considered for each year and outcome variable that was being analyzed.

The first specification considered includes all the control variables mentioned in *Table 3* but it only uses a contemporaneous measure¹⁶ of the intensity of the conflict. The second specification includes all the control variables, a

¹⁶ Intensity of conflict for the year in which the central examination was administered.

contemporaneous measure of the intensity of the conflict and the intensity of conflict at the municipality for the previous year. The third specification includes all the control variables, a contemporaneous measure of the intensity of the conflict and the intensity of conflict at the municipality for the previous 5 years (separate variable per year). The fourth specification includes all the control variables, a contemporaneous measure of the intensity and the average measure of intensity of conflict for the last 5 years.

Mathematics Examination

As it can be observed in *Table 16*, the results for the year 2001 show no significant relationship between our contemporary measure of intensity of the conflict and the performance in the mathematics examination in all but one of the specifications. We only found a significantly positive relationship between the contemporary measure of intensity of the conflict and the performance in the mathematics examination in *model 4*. The results show that on average an increase of one standard deviation in the variable that captures the intensity of the conflict causes an increase of approximately 0.18 standard deviations in the score in the mathematics examination. Intuitively, this could have been the case because the families of those students living in municipalities with a higher intensity of conflict in the months previous to the examination might be making a special effort to keep track of their kids' education in order to keep them away from the influence of the conflict. By having that special follow up, it might be the case that the relationship between the contemporary measure of intensity of conflict and performance in the exit exam is positive. Additionally, it is also possible that those students who have the highest levels of affectation because of conflict, and would have performed very poorly, have left the formal education system (dropped out) before taking this examination. In fact, Rodríguez and Sanchez (2012) have found that violent attacks in Colombian municipalities where students reside increase the probability of school drop-out and child labor. According to these authors, on average the Colombian conflict has reduced the average education of students residing in conflict areas by almost half a year of education (Rodríguez and Sanchez 2012). However, it is impossible to establish whether or not the levels of achievement of those students who have dropped out would have been lower or not in comparison to students in non-conflict areas because the levels of achievement of those students cannot be observed (we don't have a counterfactual).

It is also worth noting that the results obtained in *model 3* show a significant negative relationship between the intensity of the conflict in 1997 and the outcome in the mathematics examination, and a significant positive relationship between the intensity of the conflict in 1996 and the outcome in the mathematics examination. According to these results, on average an increase of one standard

deviation in the variable that captures the intensity of the conflict causes an increase of approximately 0.18 standard deviations in the score in the mathematics examination. Nevertheless, in *model 3* the contemporary measure of intensity of the conflict is not significant.

The results for the year 2002 (*Table 17*) show that the variable that captures the intensity of conflict is not significant in any of the specifications considered. This is true for both the contemporary measure of intensity and for the lagged measures of intensity of conflict. At first glance, these results would imply that there is not a relationship between the performance in the mathematics examination administered in 2002 and the intensity of the conflict. However, these results could be subject to endogeneity bias. For that reason, the relationship between these two variables will be further explored in the next section using an instrumental variables approach.

Table 16				
2001-2	Model 1 Mathematics	Model 2 Mathematics	Model 3 Mathematics	Model 4 Mathematics
Intensity 2001	0.043 (0.027)	0.043 (0.027)	0.041 (0.028)	0.044* (0.022)
Intensity 2000		-0.0008 (0.003)	0.001 (0.003)	
Intensity 1999			-0.004 (0.016)	
Intensity 1998			0.004 (0.009)	
Intensity 1997			-0.033** (0.016)	
Intensity 1996			0.016** (0.008)	
Average Intensity Previous 5 Years				-0.003 (0.010)

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 17				
2002-2	Model 1 Mathematics	Model 2 Mathematics	Model 3 Mathematics	Model 4 Mathematics
Intensity 2002	-0.025 (0.045)	-0.023 (0.046)	-0.033 (0.046)	-0.029 (0.046)
Intensity 2001		-0.003 (0.006)	-0.004 (0.007)	
Intensity 2000			0.003 (0.005)	
Intensity 1999			-0.008 (0.019)	
Intensity 1998			0.019 (0.013)	
Intensity 1997			-0.015 (0.014)	
Average Intensity Previous 5 Years				0.006 (0.013)

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Although our focus in this chapter is to understand the relationship between the intensity of the conflict and the performance in standardized testing, it is worth discussing briefly the results obtained for the other control variables included in the analysis. The tables that contain the complete results for the four specifications considered can be found in the appendix of this chapter. Several school-level control variables were included as part of the specification (*Appendix Tables 26 and 27*). As it can be seen in *Tables 26 and 27*, some of the control variables included in the models for both years were not significant under any of the specifications considered. For instance, the availability of an aqueduct (clean water service), the availability of a sewage service, the amount of lab space per student and the classroom space per student were not significant. On the other side, other variables, including the variable controlling for the school meeting time, were significant in all the specifications. According to the findings of the model, there is a positive relationship between being enrolled in a daytime school and the performance in the exit examination. The cost of tuition has a significant relationship with the performance in the mathematics examination. It should be highlighted that the variable that captures the cost of tuition is a categorical variable that was divided into separate dummy variables to make it suitable for the econometric exercise. The tuition category to which the other categories should be compared to is no tuition (i.e. reference category). The schools with a mid-level tuition have higher levels of performance in the examination compared to those schools with no tuition. Such finding would indicate that those students paying

the lowest amount of money for tuition, or those receiving a free education, are the ones that are comparatively doing worst in the examination.

This model also includes a range of student-level controls that have been associated by the literature with educational achievement. The results for the year 2001 and 2002 are reported in *Tables 26 and 27*. Two variables that are typically included are the age of the students and their gender. These two variables are significant for both years, and the sign of the effect turned out to be as expected. The sign of age is negative, reflecting the fact that younger students are the ones that have higher levels of achievement. The positive sign on the gender variable shows that male students (Gender = 1) are significantly doing better than female students in the mathematics examination. The model also predicts that students whose families do not have a house mortgage tend to perform better in the examination. This variable is relevant because it can help us capture the economic situation of the family of the student. The variable *number of brothers and sisters* contains information about the number of children that are being supported by the family. As expected, the sign of the effect is negative. If the family is supporting more kids, it is expected that the student has a lower performance in the examination. The education of the parents also turns out to be a significant determinant of the performance of the students in the central examination. The model predicts that those students whose parents have a higher level of education are performing better in the exam compared to those students whose parents have a lower level of education. The education level to which the other categories should be compared to is *no education*.

Language Examination

The variable that captures the contemporary measure of intensity of the conflict is significant in all of the specifications considered for the year 2001 dataset (*Table 18*). More specifically, we found a significantly positive relationship between the contemporary measure of intensity of the conflict and the performance in the language examination. According to the results, on average an increase of one standard deviation in the variable that captures the intensity of the conflict causes an increase of approximately 0.34 standard deviations in the score in the language examination. Intuitively, this could have been the case because the families of those students living in municipalities with a higher intensity of conflict in the months previous to the examination might be making a special effort to keep track of their kids' education in order to keep them away from the influence of the conflict. By having that special follow up, it might be the case that the relationship between the contemporary measure of intensity of conflict and performance in the exit exam is positive. As explained previously, it is also possible that those

students who have the highest levels of affectation because of conflict, and would have performed very poorly in the examination, left the formal education system (dropped out) before taking this examination. On the other hand, all of the variables capturing lagged measures of intensity were not significant in the different specifications that were considered in the analysis.

Just like in the case of the mathematics examination, the results for the year 2002 in the language examination (*Table 19*) show that the variable that captures the intensity of conflict is not significant in any of the specifications considered. This is true for both the contemporary measure of intensity and for the lagged measures of intensity of conflict. At first glance, these results would imply that there is not a relationship between the performance in the language examination administered in 2002 and the intensity of the conflict. However, these results could be subject to endogeneity bias. For that reason, the relationship between these two variables will be further explored in the next section using an instrumental variables approach.

Table 18				
2001-2	Model 1 Language	Model 2 Language	Model 3 Language	Model 4 Language
Intensity 2001	0.095*** (0.031)	0.094*** (0.031)	0.088*** (0.031)	0.091*** (0.031)
Intensity 2000		0.002 (0.004)	0.004 (0.004)	
Intensity 1999			-0.004 (0.021)	
Intensity 1998			0.013 (0.013)	
Intensity 1997			-0.028 (0.021)	
Intensity 1996			0.009 (0.011)	
Average Intensity Previous 5 Yrs.				0.009 (0.014)

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Several school-level control variables were included as part of the language specification (*Appendix Tables 28 and 29*). As it can be seen in *Tables 28 and 29*, a number of the control variables included in the models for both years were not

significant under any of the specifications considered. For instance, lab space per student and the number of teachers available per student were not significant under any of the specifications considered. On the other hand, other variables, including the variable controlling for the availability of a sewage system (cod_alc), the cost of tuition and the meeting time, were significant in all the specifications. According to the findings of the model, there is a positive relationship between the availability of a sewage system and the performance in the exit examination. The time of the day in which classes take place also seems to have a positive relationship with the performance in the language examination. The cost of tuition has a significant relationship with the performance in the language examination. Given the fact that the cost of tuition is a categorical variable, just like in the mathematics examination, the tuition category to which the other categories should be compared to is no tuition (i.e. reference category). The schools with a lower tuition have an inferior performance in the examination compared to those schools with a higher tuition. This indicates that those students paying the lowest amount of money for tuition, or those receiving a free education, are the ones that are comparatively doing worst in the examination.

Table 19				
2002-2	Model 1 Language	Model 2 Language	Model 3 Language	Model 4 Language
Intensity 2002	0.007 (0.048)	0.012 (0.048)	0.022 (0.049)	0.012 (0.048)
Intensity 2001		-0.008 (0.007)	-0.005 (0.007)	
Intensity 2000			0.006 (0.005)	
Intensity 1999			-0.012 (0.021)	
Intensity 1998			-0.006 (0.014)	
Intensity 1997			-0.024 (0.016)	
Average Intensity Previous 5 Yrs.				-0.013 (0.013)

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

The results obtained for the student-level controls in the years 2001 and 2002 are also reported in *Tables 28 and 29* of the appendix. The sign of the variable *age* is negative, reflecting the fact that younger students are having statistically significant higher levels of achievement in the exit examination. The variable *gender* also has a negative sign showing that female students (*Gender* = 0) are significantly doing better than male students in the language examination in both years. This shows that in comparative terms, female students do better than male students in the language examination but male students do better than female students in the mathematics examination. In line with the model for the mathematics examination, this model also predicts that students whose families do not have a house mortgage tend to perform better in the examination. The variable *Number of brothers and sisters*, which contains information about the number of children that are being supported by the family, has a negative sign as expected. According to this model, if the family is supporting more kids the student has a lower performance in the examination. As we saw in the model for the mathematics examination, the education of the parents also turns out to be a significant determinant of the performance of the students in the central examination. The model predicts that those students whose parents have a higher level of education are performing better in the exam compared to those students whose parents have a lower level of education. The education level to which the other categories should be compared to is *no education* (i.e. reference category).

C. *Random Coefficients Model*

We included school level effects in the four specifications that were considered in the previous section in order to account for the fact that some schools can be more affected by conflict than others. The school level variance was used in that case to measure the extent to which schools vary in that respect. Yet, we assumed implicitly in those models that the extent to which examination results vary because of the intensity of the conflict is the same across different schools. Nonetheless, there are reasons to think that this is not the case. It is possible for schools in the same municipality to have a different level of affectation because of conflict. This could possibly lead to a different effect on the mathematics and language scores. To account for that, we are going to allow the slope parameter to have a variance. In other words, the slope of the regression line can now take a different value across the values of the level 2 variable (school).

Mathematics Examination

As it can be observed in *Table 20*, the results for the year 2001 show a significant relationship between our contemporary measure of intensity of the conflict and the performance in the mathematics examination in three of the four the specifications considered. According to our results there is a significantly positive relationship between the contemporary measure of intensity of the conflict and the performance in the mathematics examination. On average an increase of one standard deviation in the variable that captures the intensity of the conflict causes an increase of approximately 0.18 standard deviations in the score in the mathematics examination in *model 1*, an increase of approximately 0.21 standard deviations in *model 2*, and an increase of approximately 0.19 standard deviations in *model 4*. Following the same line of analysis of the results discussed previously, this could have been the case because the families of those students living in municipalities with a higher intensity of conflict might be making a special effort to keep track of their kids' education in order to keep them away from the influence of the conflict. This special follow up might be contributing to higher levels of performance, and in doing so, it could help us why according to our results the relationship between the contemporary measure of intensity of conflict and the performance in the exit exam is positive. Additionally, as discussed previously, it is also possible that those students who have the highest levels of affectation have left the formal education system (dropped out) before taking this examination.

In line with the results obtained using a random intercept model for the year 2002, the results from the random coefficients model (*Table 21*) show that the variable that captures the intensity of conflict is not significant in any of the specifications considered. This is true for both the contemporary measure of intensity and for the lagged measures of intensity of conflict. At first glance, these results would imply that there is not a relationship between the performance in the mathematics examination administered in 2002 and the intensity of the conflict. However, the relationship between these two variables will be further explored in the next section of this chapter using an instrumental variables approach.

Table 20				
2001-2	Model 1 Mathematics	Model 2 Mathematics	Model 3 Mathematics	Model 4 Mathematics
Intensity 2001	0.044* (0.020)	0.05** (0.025)	0.042 (0.028)	0.046* (0.021)
Intensity 2000		-0.001 (0.003)	0.001 (0.003)	
Intensity 1999			-0.003 (0.016)	
Intensity 1998			0.004 (0.010)	
Intensity 1997			-0.033** (0.016)	
Intensity 1996			0.016* (0.008)	
Average Intensity Previous 5 Years				-0.003 (0.010)

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 21				
2002-2	Model 1 Mathematics	Model 2 Mathematics	Model 3 Mathematics	Model 4 Mathematics
Intensity 2002	-0.025 (0.045)	-0.031 (0.047)	-0.042 (0.048)	-0.034 (0.048)
Intensity 2001		0.001 (0.004)	-0.006 (0.007)	
Intensity 2000			0.002 (0.006)	
Intensity 1999			-0.003 (0.021)	
Intensity 1998			0.021 (0.014)	
Intensity 1997			-0.014 (0.014)	
Average Intensity Previous 5 Years				-0.001 (0.016)

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Language Examination

Table 22 shows the results of a random coefficients model that studies the relationship between the intensity of the conflict and the performance in the language examination in 2001. The variable that captures the contemporary measure of intensity of the conflict is significant in all of the specifications. According to the results, there is a significantly positive relationship between the contemporary measure of intensity of the conflict and the performance in the mathematics examination. On average an increase of one standard deviation in the variable that captures the intensity of the conflict causes an increase of approximately 0.36 standard deviations in the score in the language examination in *model 1*, an increase of approximately 0.35 standard deviations in *model 2*, an increase of approximately 0.32 standard deviations in *model 3*, and an increase of approximately 0.34 standard deviations in *model 4*. On the other hand, all of the variables capturing lagged measures of intensity were not significant in the different specifications that were considered in the analysis.

Table 22				
2001-2	Model 1 Language	Model 2 Language	Model 3 Language	Model 4 Language
Intensity 2001	0.095*** (0.032)	0.094*** (0.032)	0.086*** (0.031)	0.090*** (0.032)
Intensity 2000		0.002 (0.004)	0.004 (0.005)	
Intensity 1999			-0.001 (0.022)	
Intensity 1998			0.011 (0.014)	
Intensity 1997			-0.028 (0.021)	
Intensity 1996			0.009 (0.011)	
Average Intensity Previous 5 Years				0.009 (0.032)

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

The results for the year 2002 in the language examination (*Table 23*) show that the variable that captures the intensity of conflict is not significant in any of the specifications considered. This is true for both the contemporary measure of

intensity and for the lagged measures of intensity of conflict. At first glance, these results would imply that there is not a relationship between the performance in the language examination administered in 2002 and the intensity of the conflict. However, these results could be subject to endogeneity bias. For that reason, the relationship between these two variables will be further explored in the next section using an instrumental variables approach.

Table 23				
2002-2	Model 1 Language	Model 2 Language	Model 3 Language	Model 4 Language
Intensity 2002	0.005 (0.056)	0.005 (0.056)	0.021 (0.056)	0.016 (0.051)
Intensity 2001		0.001 (0.005)	-0.007 (0.008)	
Intensity 2000			0.006 (0.006)	
Intensity 1999			-0.009 (0.024)	
Intensity 1998			-0.010 (0.017)	
Intensity 1997			-0.020 (0.017)	
Average Intensity Previous 5 Years				-0.022 (0.018)

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

6.3 Instrumental Variables Results

Tables 24 and 25 show the results of the instrumental variable approach that was used to quantify the impact of the intensity of the conflict on the results of the language and mathematics exit examinations. We tested two different specifications. In the first specification (*model 1*) we include all the control variables and we instrument the intensity of the conflict in the year under analysis¹⁷ using the lagged measure (t-1) of laboratories dismantle and antinarcotics operations. In the second specification (*model 2*) we include all the control variables and we instrument the intensity of the conflict in the year before the students took the

¹⁷ The year in which the students took the National Examination.

National Examination using the lagged measure (t-1) of laboratories dismantle and antinarcotics operations.

Table 24 presents the results for the year 2001. The results obtained for *model 1* show a positive and significant effect of the intensity of the conflict on the performance of students in the mathematics exit examination. These results show the same trend as the ones found using the hierarchical model (random coefficient model), however the results obtained using the instrumental variables approach show a larger effect. In that sense, the results obtained using the multilevel model seem to show a lower bound of the effect. On average an increase of one standard deviation in the variable that captures the intensity of the conflict causes an increase of approximately 0.47 standard deviations in the score in the mathematics examination. On the other hand, the results obtained for *model 2* are not significant.

As we can see in table 24, the results obtained in the language examination that was taken by students in 2001 show a positive and significant effect of the intensity of the conflict for *model 2* but not for *model 1*. *Model 2* shows that on average an increase of one standard deviation in the variable that captures a lagged measure of the intensity of the conflict causes an increase of approximately 0.53 standard deviations in the score in the language examination. This result is similar to the result obtained using the contemporaneous measure of intensity in the hierarchical model (random coefficient model).

Table 24				
2001-2	Model 1 Language	Model 2 Language	Model 1 Mathematics	Model 2 Mathematics
Intensity 2001	0.014 (0.041)		0.113*** (0.046)	
Intensity 2000		0.140* (0.079)		0.072 (0.078)
Instruments: lagged laboratories dismantle and antinarcotics operations				

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 25 presents the results for the year 2002. The results show that there is a negative and significant effect of the intensity of the conflict on the performance of students in the language examination in both specifications considered. According to the results for *model 1*, on average an increase of one standard deviation in the variable that captures the intensity of the conflict causes a reduction of approximately 0.91 standard deviations in the score in the language examination. In the case of *model 2*, the results show that on average an increase of one standard

deviation in the variable that captures the lagged measure of intensity of the conflict causes a reduction of approximately 0.67 standard deviations in the score in the language examination. It should be highlighted that the results obtained using the hierarchical model (random coefficient model) showed completely different results to those obtained using the instrumental variables approach. The results using the random coefficient model showed that there was not a significant effect of the intensity of the conflict on the results in the language examination. As mentioned previously, this might suggest the existence of bias in our previous estimations of the language specification.

As we can see in *table 25*, the results do not show a significant effect of the intensity of the conflict on the performance in the mathematics examination administered in 2002 in the two specifications that were considered for the analysis. In this case, the results are very similar to the results obtained using the hierarchical model (random coefficient). The results obtained using these two methodologies show that the variable capturing the intensity of the conflict is not significant under any of the specifications considered. Additionally, it should be highlighted that in the results obtained using the two methodologies the coefficients for the intensity variable are very similar.

Table 25				
2002-2	Model 1 Language	Model 2 Language	Model 1 Mathematics	Model 2 Mathematics
Intensity 2002	-0.21* (0.126)		-0.032 (0.095)	
Intensity 2001		-0.155* (0.086)		-0.080 (0.060)
Instruments: lagged laboratories dismantle and antinarcotics operations				

Source: author's calculations using the databases from ICFES, DANE and CERAC. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

7. Discussion of the Results

The results obtained in this chapter show that the relationship between the intensity of the conflict and the performance in exit examinations is not as straightforward as it may seem at first glance. For the year 2001 we found a positive relationship between our contemporaneous measure of the intensity of the conflict and the performance in the mathematics and language examinations, but for the year 2002 we only found a negative relationship between the intensity of

the conflict and the performance in the language examination. The results obtained using an instrumental variables approach and a multilevel model (random coefficient and random intercept) suggest that there is a positive relationship between our contemporaneous measure of the intensity of the conflict and the performance of students in the mathematics examination in the year 2001. In the case of the language examination the results obtained using multilevel models (random coefficient and random intercept models) show a positive relationship between our contemporaneous measure of the intensity of the conflict and the performance of students in the exam administered in 2001. However, for the year 2002 we find completely different results. The results obtained using an instrumental variables approach show a negative relationship between our contemporaneous measure of the intensity of the conflict and the performance of students in the language examination administered in 2002, and no significant relationship in the case of the mathematics examination.

These results can be contextualized using the findings of a number of researchers who have studied the effects of conflict on educational achievement. The contributions from other researchers, including Rodriguez and Sanchez (2012), Sánchez and Díaz (2005), Barrera and Ibañez (2004) and Wharton and Oyelere (2012), are going to help us establish a number of factors through which the intensity of the conflict could affect the performance of students in exit examinations, and in doing so, help us get a better understanding of the results obtained in this chapter. By combining the results obtained in this dissertation and the results from that literature we will be able to get a better picture of how conflict is affecting education, and in doing so, better inform the policy making process.

There are a number of factors that could help us explain the results obtained using data from the year 2001. As mentioned previously, the results obtained for that year show a positive relationship between the contemporaneous measure of the intensity of the conflict and the performance in the mathematics and language examinations. The first factor that can explain this positive relationship is that the parents of those students that stay in conflict zones are most likely to invest as many resources as needed to give their kids a good education that can keep them away from the conflict. The families living in conflict areas are often faced with a very difficult dilemma: stay in their hometown or move to a non-conflict area. Those families that decide to stay are interested in the safety of their kids, and to do so, they will take as many actions to keep the kids away from the conflict and from conflict groups. One of those actions will likely be to follow up as close as possible the educational process of their sons and/or daughters. Such effort could translate in levels of performance that are similar (or possibly higher) to those of kids studying in non-conflict zones. This factor can be linked to the high levels of resilience that some individuals develop when confronted with harsh conditions.

A second factor that can explain this positive relationship is that there might be an auto-selection channel in action. It could be the case that those school-aged individuals that are mostly affected by conflict are the ones that do not participate in the formal education system. Those individuals that do not have the resources (monetary, psychological, family support) to overcome or adapt to the rough conditions that are prevalent in conflict areas are more likely to drop out from the formal educational system. I argue that those individuals that continue their formal education in conflict zones have certain resources that allow them to stay in the system. There is auto-selection because those individuals with resources to overcome or adapt to the harsh conditions are most likely to continue with their studies. As a consequence, I suspect that the levels of performance of those students in conflict zones are not significantly affected. This second channel can be easily linked to previous research efforts that have studied the relationship between conflict, low levels of enrollment and school drop-outs. Rodriguez and Sanchez (2012) have found that armed conflict reduces the average years of schooling in 8.78% for a pooled sample of Colombian children, and it reduces the average years of schooling in 17.03% for children between sixteen and seventeen years old. Barrera and Ibañez (2004) have also found that after controlling for individual, household and other context variables, violence influences negatively school enrollment. Furthermore, Sánchez and Diaz (2005) show that municipalities that were affected by the activities of illegally armed groups, had slower growth rates of primary and secondary school enrolment compared to those in municipalities that were not affected by illegal activities using a matching mechanism. These two researchers show that the dropout rate in conflict-affected municipalities is around two per cent higher than in peaceful municipalities (control group). Wharton and Oyelere (2012) also show that children who live in a municipality with high levels of conflict have a gap in enrolment and accumulation. In line with Wharton and Oyelere's (2012) work, Dueñas and Sanchez (2007) use a duration model to show that the presence of illegal armed groups increases dropout rates. All of these results imply that for many students in conflict zones "the reductions in utility, households' income and returns from education stemming from violence outweigh increments in private returns caused by shortages in supply of educated individuals" (Barrera and Ibañez 2004). In that sense, only those students who highly value education are going to continue their studies. Consequently, the auto-selection is very possibly driving our results.

As mentioned above, the results obtained using an instrumental variables approach show a negative relationship between our contemporaneous measure of the intensity of the conflict and the performance of students in the exam administered in 2002. The historical evolution of the conflict can play an important role in the interpretation of these results. In that sense, it is helpful to incorporate

into the analysis a historical perspective of the evolution of the Colombian conflict in order to get a more complete understanding of the results of this chapter.

Restrepo et al. (2006) have called the period between the years 1996 and 2002, the escalation of violence period given the increase in the intensity of the conflict and the frequency of the combats and the attacks. In fact, according to these authors the year 2002 was the year in which the armed conflict reached the highest intensity. The recrudescence of violence in Colombia during those years is explained by a number of factors: the unification of the majority of paramilitary groups into one organization (Autodefensas Unidas de Colombia), the implementation of a terror strategy against civilians, and the modernization of the Colombian military forces through the implementation of new technologies in order to have a better response against the attacks of illegal groups (Restrepo et al. 2006). Under such conditions, it could be the case that the resilience hypothesis that was developed to explain the positive relationship between the intensity of the conflict and the levels of performance in the examinations is no longer valid. Additionally, the auto-selection channel that was discussed could have also been at least partially lost. Given the extremely high levels of affectation to civilians, it is more likely that not only those school-aged individuals that are mostly affected by conflict take the decision to drop-out from the formal education system but also those students who stay in the system probably have less resources to adapt or at least overcome the difficulties associated to the conflict.

8. Policy Implications

The evidence obtained in this chapter can be used to guide the formulation of educational policy. Targeted policies should be implemented in order to provide special support to those students who remain in conflict areas so that they can continue their studies. There are a number of policy recommendations that could be implemented to guarantee that those students who stay in conflict zones have the adequate incentives and resources to continue with their studies.

As Winthrop and Kirk (2008) highlight, those students who live in a country affected by a civil conflict are more likely to be heterogeneous in terms of their age, social background, competence, family responsibilities, and previous education opportunities. Given these heterogeneities, and in order to formulate targeted policies, it should be a priority for the Colombian government to conduct a baseline assessment. Local leaders, who have information on population movements, should play a central role in this tracking process (Hillesund et al. 2014).

For education to be effective it is necessary to guarantee the conditions that facilitate learning (Winthrop and Kirk 2008). Governmental institutions, specially the Ministry of Education and the National Police of Colombia, should work together to assure the best climate for teaching and learning within schools. This includes, but it is not limited to, making sure that surrounding areas are safe (e.g. free of landmines). Furthermore, schools and classrooms should be well maintained. Hillesund et al. (2014) have shown that the quality of school facilities affects attendance and learning outcomes, and therefore, having a well maintained infrastructure is of crucial importance.

Winthrop and Kirk (2008) point out that psychosocial intervention in educational settings can help improve the potential of children affected by conflict. In that respect, the Colombian government should offer special psychological support to families and communities to help children cope with the stress and trauma caused by the civil conflict. It is also very important to identify the specialists who can address the psychosocial needs of children and young adults, as these interventions can help students handle the trauma of educational disruption (Chand et al. 2003). Building a database of specialists and consulting them periodically, through activities like in-school workshops, could be strategic in maintaining kids at school (Chand et al. 2003).

Considering the diverse needs and life experiences of children and young adults in countries affected by conflict, it is particularly important to design a variety of learning programs. As Hillesund et al. (2014) have highlighted, alternative education programs can be implemented, either in connection with or parallel to the Colombian regular educational system. For instance, remedial or accelerated learning programs can be an alternative (Hillesund et al. 2014). The remedial (accelerated) programs teach several years of schooling in a shorter time period. The implementation of these programs can play a crucial role in creating incentives to retain and bring students back to school because it helps children and young adults to catch up (Hillesund et al. 2014).

According to Hillesund et al. (2014), school participation levels depend more on reduced cost of access than improved quality of education. Furthermore, the literature has shown that vulnerable households respond well to reduction in access costs (Hillesund et al. 2014). Access costs include school fees and indirect costs for uniforms, equipment and unofficial fees. There are a number of policy instruments that could be implemented in order to reduce direct and indirect costs, and in doing so increase school participation levels in Colombia. One possibility is the implementation of a vouchers system. School vouchers are government-funded scholarships that pay for students to attend either public or private school. These vouchers allow students to attend the school of their preference. A second

mechanism that could be considered is conditional cash transfer (CCT) programs. In a conditional cash transfer program, the government transfers money to vulnerable households that meet certain conditions. In order to reduce to cost of access to education, policy makers in Colombia could establish a CCT program to provide tuition and additional stipends conditional on school attendance.

The implementation of the policies suggested above can play a central role in guaranteeing that those students who stay in conflict zones have the adequate incentives and resources to continue with their studies. This is particularly important because by keeping kids in school it is possible to improve the levels of human capital accumulation in the areas of the country particularly affected by the civil conflict, so that individuals can get better opportunities in the future.

9. Concluding Remarks

This chapter explores the relationship between civil conflict and educational achievement in Colombia through the use of multilevel analysis and an instrumental variables approach. The findings of this chapter contribute to the existing literature that studies the impact of civil conflict on education. It is relevant to explore the particular case of the students that stay in the conflict zones because these students might be adapting (or not) to the rough conditions that are prevalent in these areas, and perhaps they have a different (smaller/larger) educational gap (in terms of magnitude) than policy makers think.

Multilevel analysis was employed in this chapter in order to deal adequately with the hierarchical structure of the data by simultaneously modeling variables at different levels without resorting to aggregation or disaggregation of the dataset. We also used an instrumental variables approach to tackle the possibility of bias associated to endogeneity. It is credible to think that a student takes into consideration the intensity of conflict when deciding whether to study or not. Some students might have special characteristics (i.e. support from parents) that make them more able to face conflict-related violence, and continue with their studies. As a consequence, it is necessary to consider the possibility of endogeneity of the intensity of conflict measure.

The results obtained in this chapter show that the relationship between the intensity of the conflict and the performance in exit examinations is not as straightforward as it may seem at first glance. For the year 2001 we found a positive relationship between our contemporaneous measure of the intensity of the conflict and the performance in the mathematics and language examinations, but for the year 2002 we found a negative relationship between the intensity of the

conflict and the performance in the language examination. These results were contextualized using the contributions from other researchers, including Rodriguez and Sanchez (2012), Sánchez and Diaz (2005), Barrera and Ibañez (2004) and Wharton and Oyelere (2012), to help us establish possible factors through which the intensity of the conflict could affect the performance of students in exit examinations, and in doing so, help us get a better understanding of the results. We also incorporated into the analysis a historical perspective of the evolution of the Colombian conflict in order to get a more complete understanding of the results obtained for the year 2002.

The evidence obtained in this chapter can be used to guide the formulation of educational policy. Policy makers should design policies to provide special support to those students who remain in conflict areas so that they can continue their studies. More governmental resources should be spent in order to meet the special needs of those students who have been affected by conflict (i.e. psychological help, school supplies, special remedial tutorials) and to supervise their progress at educational institutions. As noted before, it is possible that those students who are facing difficulties associated to the armed conflict are developing high levels of resilience that allows them to continue with their studies, but they need a special follow up. We need to design effective policies to guarantee that those students who stay in conflict zones have the adequate incentives and resources to continue with their studies. By doing so we can help improve the levels of human capital accumulation in the areas of the country particularly affected by the civil conflict, so that individuals can get better opportunities in the future.

Appendix A. Tables

Random Intercept Model

Table 26 – Random Intercept Model – Mathematics 2001				
2001-2	Model 1 Mathematics	Model 2 Mathematics	Model 3 Mathematics	Model 4 Mathematics
Intensity 2001	0.043 (0.027)	0.043 (0.027)	0.041 (0.028)	0.044* (0.020)
Intensity 2000		-0.001 (0.003)	0.001 (0.003)	
Intensity 1999			-0.004 (0.016)	
Intensity 1998			0.004 (0.009)	
Intensity 1997			-0.033 (0.016)	
Intensity 1996			0.016 (0.008)	
Average Intensity Previous 5 Years				-0.003 (0.010)
Gender	0.392*** (0.071)	0.392*** (0.071)	0.391*** (0.071)	0.392*** (0.071)
Intensity * Gender	0.008 (0.006)	0.008 (0.006)	0.009 (0.006)	0.008 (0.006)
Age	-0.113*** (0.015)	-0.113*** (0.015)	-0.113*** (0.015)	-0.113*** (0.015)
Intensity * Age	-0.002* (0.001)	-0.002* (0.001)	-0.002* (0.001)	-0.002* (0.001)
Tuition level – Low	-0.201 (0.128)	-0.200 (0.129)	-0.190 (0.129)	-0.200 (0.129)
Tuition level – Mid	0.806*** (0.164)	0.808*** (0.164)	0.808*** (0.164)	0.809*** (0.164)
Tuition level – High	0.054 (0.094)	0.054 (0.094)	0.056 (0.094)	0.055 (0.094)
Mother education – Low level	0.006 (0.117)	0.006 (0.117)	0.006 (0.117)	0.006 (0.117)
Mother education – Mid level	0.132 (0.135)	0.133 (0.135)	0.135 (0.135)	0.133 (0.135)
Mother education – High level	0.637*** (0.165)	0.637*** (0.165)	0.642*** (0.165)	0.636*** (0.165)
Father education – Low level	-0.027 (0.108)	-0.027 (0.108)	-0.027 (0.108)	-0.026 (0.108)
Father education – Mid level	0.077 (0.130)	0.077 (0.130)	0.077 (0.130)	0.077 (0.130)

Father education – High level	0.631*** (0.172)	0.631*** (0.172)	0.631*** (0.172)	0.631*** (0.172)
Low level of income	0.388*** (0.096)	0.388*** (0.096)	0.391*** (0.096)	0.388*** (0.096)
Mid level of Income	0.842*** (0.182)	0.842*** (0.182)	0.845*** (0.182)	0.841*** (0.182)
High level of income	1.519*** (0.397)	1.518*** (0.397)	1.519*** (0.397)	1.518*** (0.397)
Low income * Intensity	0.001 (0.007)	0.001 (0.007)	0.001 (0.007)	0.001 (0.007)
Mid income * intensity	-0.005 (0.009)	-0.005 (0.009)	-0.006 (0.009)	-0.005 (0.009)
High income * intensity	-0.019 (0.017)	-0.019 (0.017)	-0.020 (0.018)	-0.018 (0.017)
Low Father Education * Intensity	0.001 (0.009)	0.001 (0.009)	0.001 (0.009)	0.001 (0.009)
Mid Father Education * Intensity	0.000 (0.010)	0.0001 (0.010)	0.0001 (0.010)	0.0001 (0.010)
High Father Education * Intensity	0.013 (0.011)	0.013 (0.011)	0.014 (0.011)	0.013 (0.011)
# of People living in the household	-0.025 (0.021)	-0.025 (0.021)	-0.025 (0.021)	-0.025 (0.021)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	1.263*** (0.249)	1.262*** (0.249)	1.263*** (0.249)	1.262*** (0.249)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.084 (0.084)	0.084 (0.084)	0.084 (0.084)	0.084 (0.084)
# of people contributing to household income	-0.040 (0.041)	-0.040 (0.041)	-0.039 (0.041)	-0.040 (0.041)
Number of brothers and sisters	0.011 (0.019)	0.011 (0.019)	0.012 (0.019)	0.011 (0.019)
Number of brothers and sisters with a higher education degree	-0.049 (0.032)	-0.049 (0.032)	-0.049 (0.032)	-0.049 (0.032)
Student works (dummy variable; 1 if student works, 0 otherwise)	-0.141 (0.142)	-0.141 (0.142)	-0.142 (0.142)	-0.141 (0.142)
Father occupation – unemployed	0.281 (0.227)	0.281 (0.227)	0.282 (0.227)	0.281 (0.227)
Father occupation – qualified independent	0.052 (0.189)	0.052 (0.189)	0.053 (0.189)	0.052 (0.189)
Father occupation – not qualified independent	0.213** (0.102)	0.213** (0.102)	0.214** (0.102)	0.213*** (0.102)
Father occupation – qualified dependent	0.011 (0.165)	0.011 (0.165)	0.012 (0.165)	0.011 (0.165)
Father occupation – not	0.171* (0.171)	0.171* (0.171)	0.171* (0.171)	0.171* (0.171)

qualified dependent	(0.100)	(0.100)	(0.100)	(0.100)
Population of the municipality (log)	0.077 (0.082)	0.078 (0.082)	0.091 (0.082)	0.079 (0.082)
Degree of rurality	-0.305 (0.258)	-0.309 (0.258)	-0.233 (0.260)	-0.309 (0.258)
Resources assigned by the central government to education (log)	0.134* (0.071)	0.135* (0.071)	0.125* (0.071)	0.135* (0.071)
Unsatisfied basic needs index	-0.002 (0.004)	-0.001 (0.004)	-0.002 (0.004)	-0.001 (0.004)
Area of the municipality in KM ² (log)	0.004 (0.055)	0.004 (0.055)	0.005 (0.055)	0.004 (0.055)
Distance to the capital city (log)	-0.003 (0.061)	-0.003 (0.061)	-0.020 (0.062)	-0.002 (0.061)
Altitude of the municipality (log)	0.199*** (0.037)	0.200*** (0.037)	0.197*** (0.038)	0.199*** (0.037)
Number of teachers available per student	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	-0.091 (0.121)	-0.088 (0.122)	-0.094 (0.122)	-0.088 (0.122)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	0.146 (0.140)	0.145 (0.140)	0.127 (0.141)	0.147 (0.140)
Classroom space per student	0.013 (0.021)	0.013 (0.021)	0.013 (0.021)	0.013 (0.021)
Lab space per student	0.068 (0.045)	0.067 (0.045)	0.068 (0.045)	0.068 (0.045)
School meeting time (day/night) – Dummy variable: 1 daytime; 0 otherwise	0.379* (0.218)	0.379* (0.218)	0.373* (0.219)	0.378* (0.218)
-2*loglikelihood (deviance)	144196.098	144196.026	144191.44	144195.998
LR test vs. Linear Regression (p-value)	Chi2(2) = 317.73 (0.0000)	Chi2(2) = 317.58 (0.0000)	Chi2(2) = 311.17 (0.0000)	Chi2(2) = 317.81 (0.0000)

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 27 – Random Intercept Model – Mathematics 2002				
2002-2	Model 1 Mathematics	Model 2 Mathematics	Model 3 Mathematics	Model 4 Mathematics
Intensity 2002	-0.025 (0.045)	-0.023 (0.046)	-0.033 (0.046)	-0.029 (0.046)
Intensity 2001		-0.003 (0.006)	-0.004 (0.007)	
Intensity 2000			0.003 (0.005)	
Intensity 1999			-0.008 (0.019)	
Intensity 1998			0.019 (0.013)	
Intensity 1997			-0.015 (0.014)	
Average Intensity Previous 5 Years				0.006 (0.013)
Gender	0.967*** (0.079)	0.967*** (0.079)	0.968*** (0.079)	0.967*** (0.079)
Intensity * Gender	0.020 (0.012)	0.020 (0.012)	0.020 (0.012)	0.020 (0.012)
Age	-0.122*** (0.015)	-0.122*** (0.015)	-0.122*** (0.015)	-0.121*** (0.015)
Intensity * Age	0.0001 (0.002)	0.0001 (0.002)	0.0001 (0.002)	0.0001 (0.002)
Tuition level - Low	-0.282* (0.145)	-0.282* (0.145)	-0.280* (0.145)	-0.282* (0.145)
Tuition level - Mid	0.577*** (0.179)	0.578*** (0.179)	0.574*** (0.179)	0.574*** (0.179)
Tuition level – High	0.274*** (0.103)	0.276*** (0.103)	0.276*** (0.103)	0.273*** (0.103)
Mother education - Low level	0.215 (0.134)	0.215 (0.134)	0.215 (0.134)	0.214 (0.134)
Mother education - Mid level	0.594*** (0.150)	0.594*** (0.150)	0.594*** (0.150)	0.593*** (0.150)
Mother education - High level	1.395*** (0.177)	1.395*** (0.177)	1.395*** (0.177)	1.395*** (0.177)
Father education - Low level	0.063 (0.125)	0.062 (0.125)	0.063 (0.125)	0.063 (0.125)
Father education - Mid level	-0.010 (0.146)	-0.010 (0.146)	-0.009 (0.146)	-0.010 (0.146)
Father education - High level	0.926*** (0.185)	0.926*** (0.185)	0.927*** (0.185)	0.926*** (0.185)
Low level of income	0.774*** (0.106)	0.774*** (0.106)	0.774*** (0.106)	0.775*** (0.106)
Mid level of Income	1.690*** (0.198)	1.689*** (0.198)	1.690*** (0.198)	1.691*** (0.198)

High level of income	2.687*** (0.397)	2.684*** (0.397)	2.684*** (0.397)	2.690*** (0.397)
Low income * Intensity	0.013 (0.015)	0.013 (0.015)	0.013 (0.015)	0.013 (0.015)
Mid income * intensity	-0.036 (0.030)	-0.036 (0.030)	-0.037 (0.030)	-0.037 (0.030)
High income * intensity	0.008 (0.068)	0.009 (0.068)	0.009 (0.068)	0.007 (0.068)
Low Father Education * Intensity	0.020 (0.017)	0.020 (0.017)	0.020 (0.017)	0.020 (0.017)
Mid Father Education * Intensity	0.037** (0.019)	0.037** (0.019)	0.036** (0.019)	0.036** (0.019)
High Father Education * Intensity	-0.017 (0.024)	-0.017 (0.024)	-0.017 (0.024)	-0.017 (0.024)
# of People living in the household	-0.017 (0.026)	-0.017 (0.026)	-0.017 (0.026)	-0.017 (0.026)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	1.189*** (0.290)	1.190*** (0.290)	1.192*** (0.290)	1.190*** (0.290)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.615*** (0.089)	0.614*** (0.089)	0.614*** (0.089)	0.615*** (0.089)
# of people contributing to household income	-0.165*** (0.050)	-0.165*** (0.050)	-0.165*** (0.050)	-0.165*** (0.050)
Number of brothers and sisters	-0.016 (0.024)	-0.016 (0.024)	-0.016 (0.024)	-0.016 (0.024)
Number of brothers and sisters with a higher education degree	-0.043 (0.037)	-0.043 (0.037)	-0.043 (0.038)	-0.043 (0.037)
Student works (dummy variable; 1 if student works, 0 otherwise)	-0.174 (0.147)	-0.174 (0.147)	-0.175 (0.147)	-0.173 (0.147)
Father occupation – unemployed	0.522** (0.231)	0.522** (0.231)	0.522** (0.231)	0.522** (0.231)
Father occupation – qualified independent	0.186 (0.201)	0.186 (0.201)	0.187 (0.201)	0.187 (0.201)
Father occupation – not qualified independent	0.083 (0.112)	0.082 (0.112)	0.083 (0.112)	0.083 (0.112)
Father occupation – qualified dependent	-0.050 (0.173)	-0.051 (0.173)	-0.050 (0.173)	-0.050 (0.173)
Father occupation – not qualified	-0.054 (0.110)	-0.054 (0.110)	-0.054 (0.110)	-0.054 (0.110)

dependent				
Population of the municipality (log)	0.141 (0.121)	0.144 (0.121)	0.144 (0.121)	0.138 (0.121)
Degree of rurality	-0.003 (0.339)	0.009 (0.340)	0.026 (0.340)	0.0001 (0.339)
Resources assigned by the central government to education (log)	0.069 (0.069)	0.071 (0.069)	0.068 (0.069)	0.066 (0.069)
Unsatisfied basic needs index	-0.020*** (0.005)	-0.020*** (0.005)	-0.020*** (0.005)	-0.020*** (0.005)
Area of the municipality in KM ² (log)	-0.133* (0.072)	-0.131* (0.072)	-0.133* (0.072)	-0.135* (0.072)
Distance to the capital city (log)	-0.058 (0.081)	-0.056 (0.081)	-0.062 (0.081)	-0.061 (0.081)
Altitude of the municipality (log)	0.197*** (0.048)	0.197*** (0.048)	0.198*** (0.048)	0.196*** (0.048)
Number of teachers available per student	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	0.335** (0.163)	0.334** (0.163)	0.331** (0.164)	0.332** (0.163)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	0.634*** (0.184)	0.634*** (0.184)	0.638*** (0.184)	0.636*** (0.184)
Classroom space per student	0.158*** (0.033)	0.158*** (0.033)	0.158*** (0.033)	0.158*** (0.033)
Lab space per student	-0.066 (0.047)	-0.065 (0.047)	-0.065 (0.047)	-0.066 (0.047)
-2*loglikelihood (deviance)	177841.398	177841.162	177840.634	177841.18
LR test vs. Linear Regression (p-value)	Chi2(2) = 982.51 (0.0000)	Chi2(2) = 979.70 (0.0000)	Chi2(2) = 972.32 (0.0000)	Chi2(2) = 974.46 (0.0000)

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 28 – Random Intercept Model – Language 2001				
2001-2	Model 1 Language	Model 2 Language	Model 3 Language	Model 4 Language
Intensity 2001	0.095*** (0.031)	0.094*** (0.031)	0.088*** (0.031)	0.091*** (0.031)
Intensity 2000		0.002 (0.004)	0.004 (0.004)	
Intensity 1999			-0.004 (0.021)	
Intensity 1998			0.13 (0.013)	
Intensity 1997			-0.028 (0.021)	
Intensity 1996			0.009 (0.011)	
Average Intensity Previous 5 Years				0.009 (0.014)
Gender	-0.164** (0.076)	-0.163** (0.076)	-0.163** (0.076)	-0.163*** (0.076)
Intensity * Gender	-0.003 (0.006)	-0.003 (0.006)	-0.003 (0.006)	-0.003 (0.006)
Age	-0.249*** (0.016)	-0.249*** (0.016)	-0.249*** (0.016)	-0.249*** (0.016)
Intensity * Age	-0.005** (0.002)	-0.005*** (0.002)	-0.004*** (0.002)	-0.005*** (0.002)
Tuition level - Low	-0.230 (0.145)	-0.235 (0.145)	-0.223 (0.145)	-0.234 (0.145)
Tuition level - Mid	0.832*** (0.190)	0.828*** (0.191)	0.825*** (0.191)	0.826*** (0.191)
Tuition level – High	0.222** (0.106)	0.221** (0.106)	0.225** (0.106)	0.221** (0.106)
Mother education - Low level	0.225* (0.125)	0.225* (0.125)	0.224* (0.125)	0.225* (0.125)
Mother education - Mid level	0.380*** (0.144)	0.379*** (0.144)	0.379*** (0.144)	0.379*** (0.144)
Mother education - High level	1.154*** (0.176)	1.154*** (0.176)	1.155*** (0.176)	1.154*** (0.176)
Father education - Low level	-0.099 (0.115)	-0.100 (0.115)	-0.100 (0.115)	-0.100 (0.115)
Father education - Mid level	0.094 (0.139)	0.094 (0.139)	0.094 (0.139)	0.094 (0.139)
Father education - High level	1.468*** (0.183)	1.468*** (0.183)	1.468*** (0.183)	1.469*** (0.183)
Low level of income	0.957*** (0.103)	0.957*** (0.103)	0.960*** (0.103)	0.957*** (0.103)
Mid level of Income	1.352*** (0.196)	1.352*** (0.196)	1.356*** (0.196)	1.352*** (0.196)

High level of income	2.103*** (0.430)	2.104*** (0.430)	2.106*** (0.430)	2.105*** (0.430)
Low income * Intensity	0.0001 (0.007)	0.0001 (0.007)	0.0001 (0.007)	0.0001 (0.007)
Mid income * intensity	0.004 (0.010)	0.004 (0.010)	0.003 (0.010)	0.004 (0.010)
High income * intensity	0.015 (0.019)	0.015 (0.019)	0.014 (0.019)	0.015 (0.019)
Low Father Education * Intensity	-0.007 (0.010)	-0.007 (0.010)	-0.007 (0.010)	-0.007 (0.010)
Mid Father Education * Intensity	-0.013 (0.010)	-0.013 (0.010)	-0.014 (0.010)	-0.013 (0.010)
High Father Education * Intensity	-0.020* (0.011)	-0.020* (0.011)	-0.021* (0.011)	-0.021* (0.011)
# of People living in the household	-0.104*** (0.022)	-0.104*** (0.022)	-0.104*** (0.022)	-0.104*** (0.022)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	2.346*** (0.266)	2.347*** (0.266)	2.347*** (0.266)	2.347*** (0.266)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.310*** (0.090)	0.310*** (0.090)	0.310*** (0.090)	0.309*** (0.090)
# of people contributing to household income	-0.145*** (0.044)	-0.145*** (0.044)	-0.145*** (0.044)	-0.145*** (0.044)
Number of brothers and sisters	-0.036* (0.020)	-0.036* (0.020)	-0.035* (0.020)	-0.036* (0.020)
Number of brothers and sisters with a higher education degree	-0.098*** (0.034)	-0.098*** (0.034)	-0.098*** (0.034)	-0.098*** (0.034)
Student works (dummy variable; 1 if student works, 0 otherwise)	-0.239 (0.153)	-0.239 (0.153)	-0.239 (0.153)	-0.239 (0.153)
Father occupation – unemployed	0.663*** (0.242)	0.663*** (0.242)	0.663*** (0.242)	0.663*** (0.242)
Father occupation – qualified independent	-0.167 (0.201)	-0.168 (0.201)	-0.167 (0.201)	-0.167 (0.201)
Father occupation – not qualified independent	0.290*** (0.109)	0.291*** (0.109)	0.291*** (0.109)	0.291*** (0.109)
Father occupation – qualified dependent	-0.467*** (0.176)	-0.467*** (0.176)	-0.467*** (0.176)	-0.468*** (0.176)

Father occupation – not qualified dependent	0.175 (0.107)	0.175 (0.107)	0.175 (0.107)	0.175 (0.107)
Population of the municipality (log)	0.137 (0.107)	0.134 (0.107)	0.146 (0.107)	0.133 (0.107)
Degree of rurality	-0.166 (0.334)	-0.155 (0.334)	-0.087 (0.336)	-0.156 (0.334)
Resources assigned by the central government to education (log)	0.124 (0.095)	0.121 (0.095)	0.114 (0.095)	0.120 (0.095)
Unsatisfied basic needs index	-0.014*** (0.005)	-0.015*** (0.005)	-0.016*** (0.005)	-0.015*** (0.005)
Area of the municipality in KM ² (log)	-0.025 (0.071)	-0.023 (0.071)	-0.028 (0.071)	-0.025 (0.071)
Distance to the capital city (log)	0.020 (0.080)	0.020 (0.080)	0.014 (0.082)	0.018 (0.080)
Altitude of the municipality (log)	0.248*** (0.049)	0.246*** (0.049)	0.238*** (0.050)	0.248*** (0.049)
Number of teachers available per student	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	-0.162 (0.148)	-0.172 (0.148)	-0.176 (0.148)	-0.172 (0.148)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	0.323* (0.172)	0.325* (0.173)	0.309* (0.174)	0.321* (0.173)
Classroom space per student	0.015 (0.026)	0.015 (0.026)	0.014 (0.026)	0.015 (0.026)
Lab space per student	0.065 (0.056)	0.065 (0.056)	0.066 (0.056)	0.065 (0.056)
School meeting time (day/night) – Dummy var.: 1 day; 0 otherwise	0.811*** (0.253)	0.811*** (0.253)	0.814*** (0.254)	0.813*** (0.253)
-2*loglikelihood (deviance)	147394.666	147394.304	147391.696	147394.248
LR test vs. Linear Regression (p-value)	Chi2(2) = 681.57 (0.0000)	Chi2(2) = 681.78 (0.0000)	Chi2(2) = 677.92 (0.0000)	Chi2(2) = 681.98 (0.0000)

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 29 – Random Intercept Model – Language 2002				
2002-2	Model 1 Language	Model 2 Language	Model 3 Language	Model 4 Language
Intensity 2002	0.007 (0.048)	0.012 (0.048)	0.022 (0.049)	0.012 (0.048)
Intensity 2001		-0.008 (0.007)	-0.005 (0.007)	
Intensity 2000			0.006 (0.005)	
Intensity 1999			-0.012 (0.021)	
Intensity 1998			-0.006 (0.014)	
Intensity 1997			-0.024 (0.016)	
Average Intensity Previous 5 Years				-0.013 (0.013)
Gender	-0.176** (0.081)	-0.176** (0.081)	-0.175** (0.081)	-0.176** (0.081)
Intensity * Gender	0.009 (0.013)	0.009 (0.013)	0.009 (0.013)	0.009 (0.013)
Age	-0.178*** (0.016)	-0.178*** (0.016)	-0.178*** (0.016)	-0.178*** (0.016)
Intensity * Age	-0.001 (0.002)	-0.001 (0.002)	-0.002 (0.002)	-0.002 (0.002)
Tuition level - Low	-0.188 (0.154)	-0.188 (0.154)	-0.191 (0.154)	-0.187 (0.154)
Tuition level - Mid	0.946*** (0.193)	0.950*** (0.193)	0.963*** (0.193)	0.956*** (0.193)
Tuition level – High	0.677*** (0.109)	0.681*** (0.109)	0.686*** (0.109)	0.679*** (0.109)
Mother education - Low level	0.336** (0.138)	0.337** (0.138)	0.337** (0.138)	0.337** (0.138)
Mother education - Mid level	0.493*** (0.154)	0.493*** (0.154)	0.494*** (0.154)	0.493*** (0.154)
Mother education - High level	1.476*** (0.182)	1.476*** (0.182)	1.476*** (0.182)	1.476*** (0.182)
Father education - Low level	0.093 (0.128)	0.093 (0.128)	0.092 (0.128)	0.093 (0.128)
Father education - Mid level	0.337** (0.150)	0.337** (0.150)	0.335** (0.150)	0.337** (0.150)
Father education - High level	1.438*** (0.190)	1.439*** (0.190)	1.437*** (0.190)	1.438*** (0.190)
Low level of income	1.033*** (0.109)	1.032*** (0.109)	1.032*** (0.109)	1.032*** (0.109)
Mid level of Income	1.656***	1.654***	1.658***	1.654***

	(0.204)	(0.204)	(0.204)	(0.204)
High level of income	2.552*** (0.413)	2.547*** (0.413)	2.565*** (0.413)	2.551*** (0.413)
Low income * Intensity	-0.012 (0.015)	-0.012 (0.015)	-0.011 (0.015)	-0.011 (0.015)
Mid income * intensity	0.017 (0.031)	0.018 (0.031)	0.018 (0.031)	0.018 (0.031)
High income * intensity	-0.068 (0.071)	-0.065 (0.071)	-0.070 (0.071)	-0.066 (0.071)
Low Father Education * Intensity	0.027 (0.017)	0.027 (0.017)	0.027 (0.017)	0.027 (0.017)
Mid Father Education * Intensity	0.014 (0.019)	0.015 (0.019)	0.015 (0.019)	0.015 (0.019)
High Father Education * Intensity	-0.020 (0.025)	-0.020 (0.025)	-0.020 (0.025)	-0.020 (0.025)
# of People living in the household	-0.165*** (0.026)	-0.165*** (0.026)	-0.165*** (0.026)	-0.165*** (0.026)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	2.478*** (0.298)	2.479*** (0.298)	2.481*** (0.298)	2.479*** (0.298)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.572*** (0.091)	0.571*** (0.091)	0.572*** (0.091)	0.572*** (0.091)
# of people contributing to household income	-0.123** (0.052)	-0.123** (0.052)	-0.124** (0.052)	-0.123** (0.052)
Number of brothers and sisters	-0.097*** (0.024)	-0.097*** (0.024)	-0.097*** (0.024)	-0.097*** (0.024)
Number of brothers and sisters with a higher education degree	-0.019 (0.039)	-0.019 (0.039)	-0.020 (0.039)	-0.019 (0.039)
Student works (dummy variable; 1 if student works, 0 otherwise)	0.030 (0.152)	0.029 (0.152)	0.031 (0.152)	0.030 (0.152)
Father occupation – unemployed	1.367*** (0.237)	1.367*** (0.237)	1.366*** (0.237)	1.367*** (0.237)
Father occupation – qualified independent	0.336 (0.206)	0.335 (0.206)	0.333 (0.206)	0.335 (0.206)
Father occupation – not qualified independent	0.452*** (0.115)	0.451*** (0.115)	0.451*** (0.115)	0.451*** (0.115)
Father occupation – qualified dependent	0.321* (0.177)	0.320* (0.177)	0.319* (0.177)	0.320* (0.177)
Father occupation – not qualified dependent	0.253** (0.114)	0.254** (0.114)	0.256** (0.114)	0.254** (0.114)
Population of the	0.241* (0.114)	0.248* (0.114)	0.246* (0.114)	0.246* (0.114)

municipality (log)	(0.133)	(0.134)	(0.134)	(0.134)
Degree of rurality	-0.076 (0.378)	-0.046 (0.379)	-0.004 (0.382)	-0.077 (0.378)
Resources assigned by the central government to education (log)	0.020 (0.075)	0.026 (0.075)	0.038 (0.076)	0.028 (0.075)
Unsatisfied basic needs index	-0.026*** (0.005)	-0.026*** (0.005)	-0.028*** (0.005)	-0.026*** (0.005)
Area of the municipality in KM ² (log)	-0.220*** (0.079)	-0.215*** (0.079)	-0.205*** (0.079)	-0.216*** (0.079)
Distance to the capital city (log)	-0.001 (0.088)	0.006 (0.088)	0.022 (0.088)	0.009 (0.088)
Altitude of the municipality (log)	0.230*** (0.053)	0.231*** (0.053)	0.215*** (0.054)	0.230*** (0.053)
Number of teachers available per student	-0.001 (0.0001)	-0.001 (0.0001)	-0.001 (0.0001)	-0.001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	0.398** (0.184)	0.394** (0.184)	0.385** (0.185)	0.407** (0.184)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	1.228*** (0.208)	1.225*** (0.208)	1.232*** (0.207)	1.224*** (0.208)
Classroom space per student	0.176*** (0.038)	0.176*** (0.038)	0.175*** (0.038)	0.176*** (0.038)
Lab space per student	0.010 (0.053)	0.012 (0.053)	0.011 (0.053)	0.011 (0.053)
-2*loglikelihood (deviance)	179498.902	179497.672	179496.412	179497.914
LR test vs. Linear Regression (p-value)	Chi2(2) = 1272.35 (0.0000)	Chi2(2) = 1273.36 (0.0000)	Chi2(2) = 1273.48 (0.0000)	Chi2(2) = 1272.56 (0.0000)

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Random Coefficients Model

Table 30 – Random Coefficients Model – Mathematics 2001				
2001-2	Model 1 Mathematics	Model 2 Mathematics	Model 3 Mathematics	Model 4 Mathematics
Intensity 2001	0.044* (0.020)	0.05** (0.025)	0.042 (0.028)	0.046* (0.021)
Intensity 2000		-0.001 (0.003)	0.001 (0.003)	
Intensity 1999			-0.003 (0.016)	
Intensity 1998			0.004 (0.010)	
Intensity 1997			-0.033** (0.016)	
Intensity 1996			0.016* (0.008)	
Average Intensity Previous 5 Years				-0.003 (0.010)
Gender	0.392*** (0.071)	0.392*** (0.071)	0.392*** (0.071)	0.392*** (0.071)
Intensity * Gender	0.008 (0.006)	0.008 (0.006)	0.008 (0.006)	0.008 (0.006)
Age	-0.113*** (0.015)	-0.113*** (0.015)	-0.113*** (0.015)	-0.113*** (0.015)
Intensity * Age	-0.002* (0.001)	-0.002* (0.001)	-0.002* (0.001)	-0.003* (0.001)
Tuition level - Low	-0.200 (0.128)	-0.198 (0.129)	-0.189 (0.129)	-0.198 (0.129)
Tuition level - Mid	0.807*** (0.164)	0.808*** (0.164)	0.809*** (0.164)	0.809*** (0.164)
Tuition level – High	0.054 (0.094)	0.055 (0.094)	0.056 (0.094)	0.055 (0.094)
Mother education - Low level	0.006 (0.117)	0.006 (0.117)	0.007 (0.117)	0.006 (0.117)
Mother education - Mid level	0.133 (0.135)	0.133 (0.135)	0.135 (0.135)	0.133 (0.135)
Mother education - High level	0.637*** (0.165)	0.637*** (0.165)	0.642*** (0.165)	0.637*** (0.165)
Father education - Low level	-0.027 (0.108)	-0.026 (0.108)	-0.027 (0.108)	-0.026 (0.108)
Father education - Mid level	0.078 (0.130)	0.078 (0.130)	0.077 (0.130)	0.078 (0.130)
Father education - High level	0.632*** (0.172)	0.632*** (0.172)	0.632*** (0.172)	0.632*** (0.172)
Low level of income	0.390***	0.390***	0.392***	0.390***

	(0.096)	(0.096)	(0.096)	(0.096)
Mid level of Income	0.844*** (0.182)	0.844*** (0.182)	0.848*** (0.182)	0.844*** (0.182)
High level of income	1.525*** (0.397)	1.525*** (0.397)	1.525*** (0.397)	1.524*** (0.397)
Low income * Intensity	0.001 (0.007)	0.001 (0.007)	0.0001 (0.007)	0.001 (0.007)
Mid income * intensity	-0.006 (0.009)	-0.006 (0.009)	-0.007 (0.009)	-0.006 (0.009)
High income * intensity	-0.020 (0.018)	-0.020 (0.018)	-0.021 (0.018)	-0.020 (0.018)
Low Father Education * Intensity	0.001 (0.009)	0.001 (0.009)	0.001 (0.009)	0.001 (0.009)
Mid Father Education * Intensity	0.0001 (0.010)	0.0001 (0.010)	0.0001 (0.010)	0.0001 (0.010)
High Father Education * Intensity	0.013 (0.011)	0.013 (0.011)	0.013 (0.011)	0.013 (0.011)
# of People living in the household	-0.025 (0.021)	-0.025 (0.021)	-0.025 (0.021)	-0.025 (0.021)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	1.262*** (0.249)	1.261*** (0.249)	1.262*** (0.249)	1.261*** (0.249)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.083 (0.084)	0.083 (0.084)	0.083 (0.084)	0.083 (0.084)
# of people contributing to household income	-0.039 (0.041)	-0.039 (0.041)	-0.039 (0.041)	-0.039 (0.041)
Number of brothers and sisters	0.011 (0.019)	0.011 (0.019)	0.012 (0.019)	0.011 (0.019)
Number of brothers and sisters with a higher education degree	-0.049 (0.032)	-0.049 (0.032)	-0.049 (0.032)	-0.049 (0.032)
Student works (dummy variable; 1 if student works, 0 otherwise)	-0.141 (0.142)	-0.141 (0.142)	-0.142 (0.142)	-0.141 (0.142)
Father occupation – unemployed	0.281 (0.227)	0.281 (0.227)	0.282 (0.227)	0.281 (0.227)
Father occupation – qualified independent	0.054 (0.189)	0.054 (0.189)	0.055 (0.189)	0.054 (0.189)
Father occupation – not qualified independent	0.213** (0.102)	0.213*** (0.102)	0.214** (0.102)	0.213*** (0.102)
Father occupation – qualified dependent	0.011 (0.165)	0.011 (0.165)	0.012 (0.165)	0.012 (0.165)

Father occupation – not qualified dependent	0.170* (0.100)	0.170* (0.100)	0.171* (0.100)	0.170* (0.100)
Population of the municipality (log)	0.077 (0.082)	0.078 (0.082)	0.090 (0.082)	0.078 (0.082)
Degree of rurality	-0.304 (0.258)	-0.308 (0.258)	-0.234 (0.260)	-0.308 (0.258)
Resources assigned by the central government to education (log)	0.134* (0.071)	0.135* (0.071)	0.126* (0.071)	0.135* (0.071)
Unsatisfied basic needs index	-0.002 (0.004)	-0.001 (0.004)	-0.002 (0.004)	-0.001 (0.004)
Area of the municipality in KM ² (log)	0.005 (0.055)	0.004 (0.055)	0.005 (0.055)	0.005 (0.055)
Distance to the capital city (log)	-0.003 (0.061)	-0.003 (0.061)	-0.020 (0.062)	-0.002 (0.061)
Altitude of the municipality (log)	0.199*** (0.037)	0.200*** (0.037)	0.197*** (0.038)	0.199*** (0.037)
Number of teachers available per student	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	-0.093 (0.121)	-0.089 (0.122)	-0.095 (0.121)	-0.089 (0.121)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	0.144 (0.140)	0.143 (0.140)	0.126 (0.141)	0.145 (0.140)
Classroom space per student	0.013 (0.021)	0.013 (0.021)	0.012 (0.021)	0.013 (0.021)
Lab space per student	0.068 (0.045)	0.068 (0.045)	0.069 (0.045)	0.068 (0.045)
School meeting time (day/night) – Dummy variable: 1 daytime; 0 otherwise	0.382* (0.218)	0.382* (0.218)	0.376* (0.219)	0.380* (0.218)
-2*loglikelihood (deviance)	144195.734	144195.66	144191.106	144195.628
LR test vs. Linear Regression (p-value)	Chi2(3) = 318.09 (0.0000)	Chi2(3) = 317.94 (0.0000)	Chi2(6) = 311.51 (0.0000)	Chi2(3) = 318.18 (0.0000)

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 31 – Random Coefficients Model – Mathematics 2002				
2002-2	Model 1 Mathematics	Model 2 Mathematics	Model 3 Mathematics	Model 4 Mathematics
Intensity 2002	-0.025 (0.045)	-0.031 (0.047)	-0.042 (0.048)	-0.034 (0.048)
Intensity 2001		0.001 (0.004)	-0.006 (0.007)	
Intensity 2000			0.002 (0.006)	
Intensity 1999			-0.003 (0.021)	
Intensity 1998			0.021 (0.014)	
Intensity 1997			-0.014 (0.014)	
Average Intensity Previous 5 Years				-0.001 (0.016)
Gender	0.967*** (0.079)	0.967*** (0.079)	0.966*** (0.079)	0.965*** (0.079)
Intensity * Gender	0.020 (0.012)	0.019 (0.013)	0.020 (0.012)	0.020 (0.013)
Age	-0.122*** (0.015)	-0.122*** (0.015)	-0.122*** (0.015)	-0.122*** (0.015)
Intensity * Age	0.0001 (0.002)	0.0001 (0.002)	0.0001 (0.002)	0.0001 (0.002)
Tuition level - Low	-0.282* (0.145)	-0.278* (0.145)	-0.266* (0.144)	-0.280* (0.144)
Tuition level - Mid	0.577*** (0.179)	0.590*** (0.179)	0.591*** (0.179)	0.592*** (0.179)
Tuition level – High	0.274*** (0.103)	0.269*** (0.103)	0.272*** (0.103)	0.267*** (0.103)
Mother education - Low level	0.215 (0.134)	0.216 (0.134)	0.216 (0.134)	0.215 (0.134)
Mother education - Mid level	0.594*** (0.150)	0.598*** (0.150)	0.598*** (0.149)	0.599*** (0.149)
Mother education - High level	1.395*** (0.177)	1.399*** (0.177)	1.400*** (0.177)	1.400*** (0.177)
Father education - Low level	0.063 (0.125)	0.062 (0.125)	0.059 (0.125)	0.063 (0.125)
Father education - Mid level	-0.010 (0.146)	-0.012 (0.146)	-0.014 (0.146)	-0.013 (0.146)
Father education - High level	0.926*** (0.185)	0.926*** (0.185)	0.923*** (0.185)	0.926*** (0.185)
Low level of income	0.774*** (0.106)	0.772*** (0.106)	0.774*** (0.106)	0.774*** (0.106)
Mid level of Income	1.690*** (0.198)	1.694*** (0.198)	1.699*** (0.198)	1.697*** (0.198)

High level of income	2.687*** (0.397)	2.713*** (0.396)	2.737*** (0.396)	2.726*** (0.396)
Low income * Intensity	0.013 (0.015)	0.013 (0.015)	0.011 (0.015)	0.012 (0.015)
Mid income * intensity	-0.036 (0.030)	-0.043 (0.030)	-0.050* (0.030)	-0.046 (0.030)
High income * intensity	0.008 (0.068)	-0.008 (0.069)	-0.027 (0.068)	-0.017 (0.069)
Low Father Education * Intensity	0.020 (0.017)	0.021 (0.017)	0.022 (0.017)	0.021 (0.017)
Mid Father Education * Intensity	0.037* (0.019)	0.037* (0.019)	0.037* (0.019)	0.037* (0.019)
High Father Education * Intensity	-0.017 (0.024)	-0.019 (0.024)	-0.020 (0.024)	-0.020 (0.024)
# of People living in the household	-0.017 (0.026)	-0.017 (0.026)	-0.017 (0.026)	-0.017 (0.026)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	1.189*** (0.290)	1.188*** (0.290)	1.182*** (0.290)	1.185*** (0.290)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.615*** (0.089)	0.616*** (0.089)	0.617*** (0.089)	0.617*** (0.089)
# of people contributing to household income	-0.165*** (0.050)	-0.164*** (0.050)	-0.165*** (0.050)	-0.164*** (0.050)
Number of brothers and sisters	-0.016 (0.024)	-0.016 (0.024)	-0.016 (0.024)	-0.016 (0.024)
Number of brothers and sisters with a higher education degree	-0.043 (0.037)	-0.043 (0.037)	-0.043 (0.037)	-0.043 (0.037)
Student works (dummy variable; 1 if student works, 0 otherwise)	-0.174 (0.147)	-0.173 (0.147)	-0.175 (0.147)	-0.173 (0.147)
Father occupation – unemployed	0.522** (0.231)	0.523** (0.231)	0.525** (0.230)	0.523** (0.230)
Father occupation – qualified independent	0.186 (0.201)	0.187 (0.201)	0.194 (0.201)	0.187 (0.201)
Father occupation – not qualified independent	0.083 (0.112)	0.082 (0.112)	0.082 (0.112)	0.082 (0.112)
Father occupation – qualified dependent	-0.050 (0.173)	-0.048 (0.173)	-0.046 (0.173)	-0.047 (0.173)
Father occupation – not qualified dependent	-0.054 (0.110)	-0.058 (0.110)	-0.056 (0.110)	-0.059 (0.110)
Population of the municipality (log)	0.141 (0.121)	0.143 (0.120)	0.166 (0.120)	0.145 (0.120)
Degree of rurality	-0.003 (0.339)	0.022 (0.337)	0.117 (0.339)	0.026 (0.337)
Resources assigned by	0.069	0.065	0.063	0.067

the central government to education (log)	(0.069)	(0.069)	(0.069)	(0.069)
Unsatisfied basic needs index	-0.020*** (0.005)	-0.020*** (0.005)	-0.021*** (0.005)	-0.020*** (0.005)
Area of the municipality in KM ² (log)	-0.133* (0.072)	-0.135* (0.071)	-0.137* (0.071)	-0.133* (0.071)
Distance to the capital city (log)	-0.058 (0.081)	-0.062 (0.080)	-0.058 (0.080)	-0.061 (0.081)
Altitude of the municipality (log)	0.197*** (0.048)	0.198*** (0.048)	0.193*** (0.048)	0.197*** (0.048)
Number of teachers available per student	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	0.335** (0.163)	0.311* (0.163)	0.280* (0.164)	0.287* (0.164)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	0.634*** (0.184)	0.629*** (0.183)	0.620*** (0.183)	0.617*** (0.183)
Classroom space per student	0.158*** (0.033)	0.156*** (0.033)	0.156*** (0.033)	0.153*** (0.033)
Lab space per student	-0.066 (0.047)	-0.065 (0.046)	-0.063 (0.046)	-0.065 (0.046)
-2*loglikelihood (deviance)	177841.398	177835.976	177825.01	177831.468
LR test vs. Linear Regression (p-value)	Chi2(3) = 982.51 (0.0000)	Chi2(4) = 986.85 (0.0000)	Chi2(6) = 977.63 (0.0000)	Chi2(4) = 984.17 (0.0000)

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 32 – Random Coefficients Model – Language 2001				
2001-2	Model 1 Language	Model 2 Language	Model 3 Language	Model 4 Language
Intensity 2001	0.095*** (0.032)	0.094*** (0.032)	0.086*** (0.031)	0.090*** (0.032)
Intensity 2000		0.002 (0.004)	0.004 (0.005)	
Intensity 1999			-0.001 (0.022)	
Intensity 1998			0.011 (0.014)	
Intensity 1997			-0.028 (0.021)	
Intensity 1996			0.009 (0.011)	
Average Intensity Previous 5 Years				0.009 (0.032)
Gender	-0.166** (0.076)	-0.166** (0.076)	-0.166** (0.076)	-0.165** (0.076)
Intensity * Gender	-0.003 (0.006)	-0.003 (0.006)	-0.002 (0.006)	-0.002 (0.006)
Age	-0.249*** (0.016)	-0.249*** (0.016)	-0.249*** (0.016)	-0.249*** (0.016)
Intensity * Age	-0.005*** (0.002)	-0.005*** (0.002)	-0.004*** (0.002)	-0.005*** (0.002)
Tuition level - Low	-0.226 (0.145)	-0.235 (0.145)	-0.222 (0.145)	-0.230 (0.145)
Tuition level - Mid	0.848*** (0.191)	0.843*** (0.191)	0.848*** (0.191)	0.843*** (0.191)
Tuition level – High	0.218** (0.106)	0.218** (0.106)	0.220** (0.106)	0.217** (0.106)
Mother education - Low level	0.226* (0.125)	0.225* (0.125)	0.224* (0.125)	0.226* (0.125)
Mother education - Mid level	0.381*** (0.144)	0.380*** (0.144)	0.379*** (0.144)	0.380*** (0.144)
Mother education - High level	1.156*** (0.176)	1.156*** (0.176)	1.157*** (0.176)	1.156*** (0.176)
Father education - Low level	-0.100 (0.115)	-0.101 (0.115)	-0.101 (0.115)	-0.101 (0.115)
Father education - Mid level	0.094 (0.139)	0.093 (0.139)	0.093 (0.139)	0.094 (0.139)
Father education - High level	1.469*** (0.183)	1.469*** (0.183)	1.470*** (0.183)	1.469*** (0.183)
Low level of income	0.958*** (0.103)	0.957*** (0.103)	0.960*** (0.103)	0.958*** (0.103)
Mid level of Income	1.355*** (0.196)	1.355*** (0.196)	1.361*** (0.196)	1.355*** (0.196)

High level of income	2.113*** (0.430)	2.115*** (0.430)	2.122*** (0.430)	2.114*** (0.430)
Low income * Intensity	0.0001 (0.007)	0.0001 (0.007)	-0.001 (0.007)	0.0001 (0.007)
Mid income * intensity	0.002 (0.010)	0.002 (0.010)	0.0001 (0.010)	0.002 (0.010)
High income * intensity	0.012 (0.019)	0.012 (0.019)	0.009 (0.019)	0.012 (0.019)
Low Father Education * Intensity	-0.007 (0.010)	-0.007 (0.010)	-0.007 (0.010)	-0.007 (0.010)
Mid Father Education * Intensity	-0.014 (0.010)	-0.014 (0.010)	-0.014 (0.010)	-0.014 (0.010)
High Father Education * Intensity	-0.021* (0.011)	-0.022 (0.011)	-0.022 (0.011)	-0.022 (0.011)
# of People living in the household	-0.104*** (0.022)	-0.104*** (0.022)	-0.104*** (0.022)	-0.104*** (0.022)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	2.346*** (0.266)	2.345*** (0.266)	2.348*** (0.266)	2.347*** (0.266)
Family doesn't have a mortgage (dummy var.; 1: yes, 0: no)	0.309*** (0.090)	0.309*** (0.090)	0.309*** (0.090)	0.309*** (0.090)
# of people contributing to household income	-0.144*** (0.044)	-0.144*** (0.044)	-0.143*** (0.044)	-0.144*** (0.044)
Number of brothers and sisters	-0.036* (0.020)	-0.036* (0.020)	-0.035* (0.020)	-0.036* (0.020)
Number of brothers and sisters with a higher education degree	-0.098*** (0.034)	-0.099*** (0.034)	-0.099*** (0.034)	-0.099*** (0.034)
Student works (dummy variable; 1 if student works, 0 otherwise)	-0.240 (0.153)	-0.241 (0.153)	-0.242 (0.153)	-0.240 (0.153)
Father occupation – unemployed	0.663*** (0.242)	0.663*** (0.242)	0.665*** (0.242)	0.664*** (0.242)
Father occupation – qualified independent	-0.163 (0.201)	-0.162 (0.201)	-0.159 (0.201)	-0.163 (0.201)
Father occupation – not qualified independent	0.291*** (0.109)	0.292*** (0.109)	0.291*** (0.109)	0.291*** (0.109)
Father occupation – qualified dependent	-0.466*** (0.176)	-0.466*** (0.176)	-0.467*** (0.176)	-0.466*** (0.176)
Father occupation – not qualified dependent	0.174 (0.107)	0.175 (0.107)	0.174 (0.107)	0.174 (0.107)
Population of the municipality (log)	0.135 (0.107)	0.132 (0.107)	0.142 (0.108)	0.131 (0.107)
Degree of rurality	-0.161 (0.333)	-0.151 (0.334)	-0.077 (0.336)	-0.150 (0.334)
Resources assigned by the central government to	0.124 (0.095)	0.122 (0.096)	0.115 (0.096)	0.120 (0.096)

education (log)				
Unsatisfied basic needs index	-0.014*** (0.005)	-0.015*** (0.005)	-0.016*** (0.005)	-0.015*** (0.005)
Area of the municipality in KM ² (log)	-0.024 (0.070)	-0.022 (0.071)	-0.027 (0.071)	-0.024 (0.071)
Distance to the capital city (log)	0.021 (0.080)	0.021 (0.080)	0.016 (0.082)	0.018 (0.081)
Altitude of the municipality (log)	0.247*** (0.049)	0.245*** (0.049)	0.235*** (0.050)	0.247*** (0.049)
Number of teachers available per student	0.0001 (0.00001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	-0.164 (0.147)	-0.189 (0.149)	-0.191 (0.148)	-0.173 (0.148)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	0.326* (0.172)	0.323* (0.172)	0.305* (0.173)	0.324* (0.172)
Classroom space per student	0.014 (0.025)	0.014 (0.025)	0.014 (0.025)	0.015 (0.025)
Lab space per student	0.064 (0.056)	0.065 (0.056)	0.065 (0.055)	0.064 (0.056)
School meeting time (day/night) – Dummy var.: 1 day; 0 otherwise	0.815*** (0.252)	0.816*** (0.252)	0.819*** (0.253)	0.817*** (0.252)
-2*loglikelihood (deviance)	147393.286	147392.556	147388.902	147392.854
LR test vs. Linear Regression (p-value)	Chi2(3) = 682.95 (0.0000)	Chi2(4) = 683.52 (0.0000)	Chi2(6) = 682.68 (0.0000)	Chi2(3) = 683.38 (0.0000)

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 33 – Random Coefficients Model – Language 2002				
2002-2	Model 1 Language	Model 2 Language	Model 3 Language	Model 4 Language
Intensity 2002	0.005 (0.056)	0.005 (0.056)	0.021 (0.056)	0.016 (0.051)
Intensity 2001		0.001 (0.005)	-0.007 (0.008)	
Intensity 2000			0.006 (0.006)	
Intensity 1999			-0.009 (0.024)	
Intensity 1998			-0.010 (0.017)	
Intensity 1997			-0.020 (0.017)	
Average Intensity Previous 5 Years				-0.022 (0.018)
Gender	-0.180** (0.081)	-0.180** (0.081)	-0.180** (0.081)	-0.178** (0.081)
Intensity * Gender	0.008 (0.013)	0.008 (0.013)	0.009 (0.013)	0.009 (0.013)
Age	-0.179*** (0.016)	-0.179*** (0.016)	-0.179*** (0.016)	-0.178*** (0.016)
Intensity * Age	-0.001 (0.003)	-0.001 (0.003)	-0.001 (0.003)	-0.001 (0.002)
Tuition level - Low	-0.186 (0.154)	-0.187 (0.154)	-0.189 (0.154)	-0.194 (0.154)
Tuition level - Mid	0.938*** (0.193)	0.937*** (0.193)	0.956*** (0.193)	0.965*** (0.193)
Tuition level – High	0.684*** (0.109)	0.684*** (0.109)	0.689*** (0.109)	0.679*** (0.109)
Mother education - Low level	0.338** (0.138)	0.338** (0.138)	0.339** (0.138)	0.338** (0.138)
Mother education - Mid level	0.497*** (0.154)	0.497*** (0.154)	0.498*** (0.154)	0.497*** (0.154)
Mother education - High level	1.482*** (0.182)	1.482*** (0.182)	1.482*** (0.182)	1.480*** (0.182)
Father education - Low level	0.096 (0.128)	0.096 (0.128)	0.096 (0.128)	0.095 (0.128)
Father education - Mid level	0.343** (0.150)	0.342** (0.150)	0.341** (0.150)	0.337** (0.150)
Father education - High level	1.452*** (0.190)	1.452*** (0.190)	1.450*** (0.190)	1.443*** (0.190)
Low level of income	1.040*** (0.109)	1.040*** (0.109)	1.039*** (0.109)	1.034*** (0.109)
Mid level of Income	1.677*** (0.204)	1.678*** (0.204)	1.677*** (0.204)	1.662*** (0.204)

High level of income	2.551*** (0.413)	2.552*** (0.413)	2.565*** (0.413)	2.569*** (0.412)
Low income * Intensity	-0.015 (0.016)	-0.015 (0.016)	-0.014 (0.015)	-0.012 (0.015)
Mid income * intensity	0.007 (0.031)	0.007 (0.031)	0.009 (0.031)	0.010 (0.031)
High income * intensity	-0.054 (0.073)	-0.054 (0.073)	-0.058 (0.073)	-0.074 (0.072)
Low Father Education * Intensity	0.024 (0.018)	0.024 (0.018)	0.024 (0.018)	0.027 (0.017)
Mid Father Education * Intensity	0.010 (0.020)	0.010 (0.020)	0.011 (0.020)	0.014 (0.020)
High Father Education * Intensity	-0.030 (0.025)	-0.031 (0.025)	-0.029 (0.025)	-0.022 (0.025)
# of People living in the household	-0.164*** (0.026)	-0.164*** (0.026)	-0.164*** (0.026)	-0.165*** (0.026)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	2.475*** (0.298)	2.475*** (0.298)	2.479*** (0.298)	2.479*** (0.298)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.574*** (0.091)	0.574*** (0.091)	0.574*** (0.091)	0.573*** (0.091)
# of people contributing to household income	-0.123** (0.052)	-0.123** (0.052)	-0.123** (0.052)	-0.123** (0.052)
Number of brothers and sisters	-0.099*** (0.024)	-0.099*** (0.024)	-0.098*** (0.024)	-0.098*** (0.024)
Number of brothers and sisters with a higher education degree	-0.017 (0.039)	-0.017 (0.039)	-0.018 (0.039)	-0.018 (0.039)
Student works (dummy variable; 1 if student works, 0 otherwise)	0.027 (0.152)	0.027 (0.152)	0.028 (0.152)	0.028 (0.152)
Father occupation – unemployed	1.374*** (0.237)	1.374*** (0.237)	1.372*** (0.237)	1.370*** (0.237)
Father occupation – qualified independent	0.332 (0.206)	0.332 (0.206)	0.330 (0.206)	0.333 (0.206)
Father occupation – not qualified independent	0.452*** (0.115)	0.452*** (0.115)	0.450*** (0.115)	0.450*** (0.115)
Father occupation – qualified dependent	0.323* (0.177)	0.323* (0.177)	0.321* (0.177)	0.320* (0.177)
Father occupation – not qualified dependent	0.253** (0.114)	0.253** (0.114)	0.254** (0.114)	0.251** (0.114)
Population of the municipality (log)	0.238* (0.133)	0.236* (0.133)	0.245* (0.133)	0.245* (0.133)
Degree of rurality	-0.022 (0.379)	-0.016 (0.380)	0.028 (0.381)	-0.031 (0.377)
Resources assigned by the	0.026	0.026	0.043	0.034

central government to education (log)	(0.076)	(0.076)	(0.076)	(0.076)
Unsatisfied basic needs index	-0.026*** (0.005)	-0.026*** (0.005)	-0.028*** (0.005)	-0.027*** (0.005)
Area of the municipality in KM ² (log)	-0.226*** (0.079)	-0.225*** (0.079)	-0.209*** (0.079)	-0.211*** (0.079)
Distance to the capital city (log)	0.008 (0.088)	0.007 (0.088)	0.030 (0.089)	0.012 (0.088)
Altitude of the municipality (log)	0.230*** (0.054)	0.229*** (0.054)	0.221*** (0.054)	0.227*** (0.053)
Number of teachers available per student	-0.001 (0.0001)	-0.001 (0.0001)	-0.001 (0.0001)	-0.001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	0.397** (0.183)	0.393** (0.183)	0.380** (0.184)	0.359** (0.185)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	1.220*** (0.205)	1.222*** (0.205)	1.220*** (0.206)	1.224*** (0.207)
Classroom space per student	0.171*** (0.037)	0.171*** (0.037)	0.169*** (0.037)	0.173*** (0.037)
Lab space per student	0.009 (0.052)	0.009 (0.052)	0.011 (0.052)	0.010 (0.053)
-2*loglikelihood (deviance)	179494.398	179494.328	179490.07	179490.844
LR test vs. Linear Regression (p-value)	Chi2(3) = 1276.85 (0.0000)	Chi2(4) = 1276.89 (0.0000)	Chi2(5) = 1271.77 (0.0000)	Chi2(4) = 1280.13 (0.0000)

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Instrumental Variable Approach

Table 34 – Instrumental Variable Approach 2001				
2001-2	Model 1 Language	Model 2 Language	Model 1 Mathematics	Model 2 Mathematics
Intensity 2001	0.014 (0.041)		0.113** (0.046)	
Intensity 2000		0.140* (0.079)		0.072 (0.078)
Gender	-0.233** (0.094)	-0.085 (0.141)	0.425*** (0.089)	0.519*** (0.125)
Intensity * Gender	-0.032 (0.025)	0.005 (0.043)	0.029 (0.036)	0.010 (0.038)
Age	-0.269*** (0.020)	-0.248*** (0.023)	-0.125*** (0.018)	-0.111*** (0.020)
Intensity * Age	0.0001 (0.002)	-0.003 (0.002)	-0.003 (0.002)	-0.001 (0.002)
Tuition level - Low	-0.320* (0.176)	-0.928** (0.426)	-0.352* (0.205)	-0.585 (0.410)
Tuition level - Mid	0.834*** (0.283)	0.033 (0.588)	0.274 (0.344)	0.301 (0.561)
Tuition level – High	0.077 (0.134)	-0.031 (0.174)	-0.223* (0.133)	-0.094 (0.144)
Mother education - Low level	0.251* (0.131)	0.227* (0.136)	-0.052 (0.118)	-0.038 (0.120)
Mother education - Mid level	0.434*** (0.156)	0.298 (0.183)	0.057 (0.137)	0.037 (0.161)
Mother education - High level	1.246*** (0.194)	1.345*** (0.205)	0.703*** (0.175)	0.716*** (0.186)
Father education - Low level	-0.174 (0.113)	-0.227* (0.122)	-0.041 (0.103)	-0.064 (0.110)
Father education - Mid level	0.022 (0.136)	-0.066 (0.152)	0.055 (0.132)	0.021 (0.148)
Father education - High level	1.349*** (0.176)	1.400*** (0.188)	0.737*** (0.167)	0.789*** (0.167)
Low level of income	1.135*** (0.129)	1.192*** (0.139)	0.589*** (0.123)	0.540*** (0.122)
Mid level of Income	2.051*** (0.268)	2.377*** (0.322)	1.485*** (0.303)	1.383*** (0.317)
High level of income	3.462*** (0.599)	3.861*** (0.619)	3.103*** (0.994)	2.940*** (0.947)
Low income * Intensity	0.006 (0.039)	-0.072 (0.055)	-0.082* (0.048)	-0.048 (0.056)
Mid income * intensity	-0.042 (0.092)	-0.257* (0.138)	-0.228* (0.121)	-0.124 (0.142)
High income * intensity	-0.007 (0.118)	-0.205 (0.152)	-0.312 (0.195)	-0.194 (0.166)

# of People living in the household	-0.119*** (0.022)	-0.088*** (0.027)	-0.022 (0.021)	-0.014 (0.026)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	2.597*** (0.246)	2.898*** (0.345)	1.361*** (0.220)	1.574*** (0.290)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.415*** (0.099)	0.325*** (0.111)	0.229** (0.090)	0.151 (0.100)
# of people contributing to household income	-0.174*** (0.047)	-0.196*** (0.055)	-0.051 (0.041)	-0.069 (0.046)
Number of brothers and sisters	-0.056*** (0.020)	-0.045** (0.023)	-0.007 (0.019)	0.004 (0.021)
Number of brothers and sisters with a higher education degree	-0.055 (0.037)	-0.079** (0.040)	-0.038 (0.034)	-0.040 (0.035)
Student works (dummy variable; 1 if student works, 0 otherwise)	-0.341** (0.152)	-0.283* (0.159)	-0.165 (0.141)	-0.150 (0.144)
Father occupation – unemployed	0.884*** (0.244)	0.998*** (0.251)	0.252 (0.242)	0.376 (0.253)
Father occupation – qualified independent	0.079 (0.212)	0.225 (0.238)	0.227 (0.217)	0.159 (0.240)
Father occupation – not qualified independent	0.488*** (0.106)	0.639*** (0.140)	0.281*** (0.098)	0.335*** (0.136)
Father occupation – not qualified dependent	0.309*** (0.110)	0.327*** (0.109)	0.116 (0.105)	0.212** (0.099)
Population (log)	0.146 (0.110)	-0.056 (0.162)	-0.013 (0.112)	0.020 (0.144)
Degree of rurality	-0.035 (0.317)	0.736 (0.543)	-0.226 (0.232)	0.268 (0.498)
Resources assigned by the central government to education (log)	0.149 (0.116)	-0.153 (0.204)	-0.090 (0.125)	0.014 (0.194)
Unsatisfied basic needs index	-0.013* (0.007)	-0.048** (0.020)	0.014* (0.008)	-0.021 (0.020)
Area of the municipality in KM ² (log)	-0.027 (0.072)	0.043 (0.098)	-0.139* (0.076)	0.019 (0.078)
Distance to the capital city (log)	-0.103 (0.123)	-0.233* (0.133)	-0.292** (0.125)	-0.099 (0.111)
Altitude of the municipality (log)	0.236*** (0.049)	0.081 (0.109)	0.196*** (0.049)	0.124 (0.101)
Sewage service	0.293* (0.116)	0.980** (0.381)	0.235 (0.125)	0.558 (0.290)

(Dummy variable: 1 service available; 0 otherwise)	(0.176)	(0.453)	(0.155)	(0.429)
Classroom space per student	0.015 (0.047)	0.038 (0.049)	0.054 (0.053)	0.028 (0.052)
Lab space per student	0.012 (0.085)	-0.029 (0.089)	-0.005 (0.076)	0.018 (0.073)
School meeting time (day/night) – Dummy variable: 1 daytime; 0 otherwise	0.703*** (0.235)	0.773*** (0.225)	0.208 (0.238)	0.355 (0.227)
Observations	23843	23843	23843	23843
First Stage F-Test	4.05	3.41	4.05	3.41
Prob > F	0.0177	0.0333	0.0177	0.0333
Hansen's J Statistic	2.202	2.002	1.248	0.136
P-value	0.1378	0.1672	0.2640	0.7122
Instruments: lagged laboratories dismantle and antinarcotics operations				

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 35 – Instrumental Variable Approach 2002				
2002-2	Model 1 Language	Model 2 Language	Model 1 Mathematics	Model 2 Mathematics
Intensity 2001	-0.21* (0.126)		-0.032 (0.095)	
Intensity 2000		-0.155* (0.086)		-0.080 (0.060)
Gender	-0.481*** (0.094)	-0.505*** (0.106)	1.037*** (0.096)	1.031*** (0.099)
Age	-0.190*** (0.018)	-0.200*** (0.021)	-0.139*** (0.014)	-0.148*** (0.015)
Tuition level - Low	-0.730*** (0.174)	-0.562** (0.235)	-0.449*** (0.163)	-0.389** (0.188)
Tuition level - Mid	1.171*** (0.222)	1.404*** (0.361)	0.663*** (0.198)	0.837*** (0.268)
Tuition level – High	0.420*** (0.122)	0.666*** (0.224)	0.179* (0.107)	0.317** (0.161)
Mother education - Low level	0.248* (0.137)	0.284** (0.127)	0.255** (0.108)	0.247** (0.106)
Mother education - Mid level	0.579*** (0.137)	0.626*** (0.136)	0.538*** (0.114)	0.548*** (0.118)
Mother education - High level	1.798*** (0.167)	1.793*** (0.169)	1.627*** (0.145)	1.591*** (0.149)
Father education - Low level	-2.059 (1.302)	-1.062* (0.646)	-0.354 (0.986)	-0.607 (0.457)
Father education - Mid level	-2.134 (1.524)	-0.934 (0.743)	-0.329 (1.152)	-0.599 (0.523)
Father education - High level	-1.203 (1.709)	0.144 (0.847)	0.550 (1.302)	0.246 (0.608)
Low level of income	1.623*** (0.106)	1.597*** (0.107)	1.203*** (0.089)	1.214*** (0.089)
Mid level of Income	2.970*** (0.175)	3.074*** (0.207)	2.729*** (0.195)	2.792*** (0.208)
High level of income	4.388*** (0.311)	4.485*** (0.328)	4.157*** (0.347)	4.283*** (0.360)
# of People living in the household	-0.204*** (0.022)	-0.185*** (0.023)	-0.049** (0.020)	-0.042** (0.021)
Family owns their own house (dummy variable; 1: yes, 0: otherwise)	3.174*** (0.231)	3.081*** (0.224)	1.944*** (0.192)	1.923*** (0.192)
Family does not have a house mortgage (dummy variable; 1: yes, 0: otherwise)	0.719*** (0.087)	0.592*** (0.133)	0.705*** (0.078)	0.615*** (0.104)
# of people contributing to household income	-0.256*** (0.042)	-0.331*** (0.061)	-0.309*** (0.039)	-0.350*** (0.050)
Number of brothers and	-0.211***	-0.193***	-0.078***	-0.076***

sisters	(0.025)	(0.022)	(0.021)	(0.019)
Father occupation – unemployed	1.126*** (0.192)	1.194*** (0.202)	0.835*** (0.177)	0.846*** (0.182)
Father occupation – qualified independent	0.594** (0.161)	0.429** (0.207)	0.559*** (0.149)	0.451*** (0.170)
Father occupation – not qualified independent	0.217* (0.125)	0.209 (0.130)	0.316*** (0.103)	0.260*** (0.101)
Father occupation – qualified dependent	0.226* (0.135)	0.200 (0.145)	0.344*** (0.129)	0.329** (0.132)
Father occupation – not qualified dependent	-0.007 (0.123)	0.158 (0.097)	0.056 (0.101)	0.090 (0.088)
Population (log)	0.919 (0.561)	0.652* (0.371)	0.078 (0.422)	0.267 (0.254)
Degree of rurality	0.910 (0.548)	1.674** (0.843)	-0.126 (0.406)	0.506 (0.574)
Resources assigned by the central government to education (log)	-0.051 (0.106)	0.001 (0.086)	0.079 (0.082)	0.057 (0.063)
Unsatisfied basic needs index	-0.011 (0.010)	-0.021*** (0.007)	-0.015* (0.008)	-0.015*** (0.005)
Area of the municipality in KM ² (log)	-0.140** (0.063)	-0.106 (0.078)	-0.085 (0.055)	-0.084 (0.060)
Altitude of the municipality (log)	0.499*** (0.154)	0.378*** (0.088)	0.300*** (0.115)	0.328*** (0.064)
Number of teachers available per student	0.001 (0.001)	0.0001 (0.001)	0.0001 (0.0001)	0.0001 (0.0001)
Clean water service (Dummy variable: 1 service available; 0 otherwise)	0.293 (0.185)	0.251 (0.234)	0.252 (0.190)	0.227 (0.212)
Sewage service (Dummy variable: 1 service available; 0 otherwise)	0.912*** (0.193)	0.807*** (0.233)	0.512*** (0.157)	0.491 (0.176)
Classroom space per student	0.097*** (0.035)	0.082** (0.042)	0.075** (0.036)	0.062* (0.036)
Lab space per student	0.030 (0.030)	0.033 (0.025)	-0.001 (0.039)	0.001 (0.032)
Observations	54279	54279	54279	54279
First Stage F-Test Prob > F	5.43 0.0044	6.75 0.0012	5.43 0.0044	6.75 0.0012
Hansen's J Statistic P-value	2.555 0.1099	0.111 0.7393	1.497 0.2212	0.032 0.8571
Instruments: lagged laboratories dismantle and antinarcotics operations				

Source: author's calculations using the databases from ICFES, DANE and CERAC.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Appendix B. Measuring Violent Conflict – CERAC’s Colombian Conflict Database

As mentioned by Restrepo et al (2004), there are very few detailed databases of internal conflicts available around the world. Measuring and analyzing violence due to conflict either directly or indirectly is particularly difficult, and many times can also be life-threatening (Restrepo et al 2009). There are several issues that should be addressed in the process of measurement and analysis. The first issue that has to be tackled is related to the methodology that should be employed in this process. Specifically, this methodological issue is related to the procedure behind determining what to measure and how to measure it (Restrepo et al 2009). The specific objective of CERAC’s quantification effort is to measure violence associated to conflict. To do so, as mentioned by Restrepo et al (2009), it is necessary to acknowledge that there is violence in a given context, and given the definition of conflict quantify the violent events that are associated to it. The events associated to this type of violence are defined as those actions carried out by a conflict group which seeks to obtain its preferred social option through intentional harmful actions to people (civil society and combatants), communities and infrastructure (Restrepo et al 2004) (Restrepo et al 2009). According to Restrepo et al (2009), the use of this definition assumes that violence involves effective intentional damage. This entails that it does not include potential damage or fear. Intentional damage is associated to the destructive violent activities of an armed group to obtain its preferred social option. These actions are against the members and the infrastructure of the other groups that take part in the conflict (Restrepo et al 2009).

The methodological criteria described above allows for a flexible approach because it is possible to include violent events between any of the conflict actors, and it doesn’t impose any restrictions in terms of the type of event, the weapons being used, or the number of people affected (Restrepo et al 2009). Additionally, the methodology does not impose restrictions in terms of the nature of the violent groups. However, it excludes violent events that are not associated to the groups that take part in the Colombian armed conflict. As mentioned above, conflict associated to violence can only take place if the armed groups carry out actions that lead to violent events or these armed groups are victims of violent actions (Restrepo et al 2009). CERAC points out that the objective behind such definition is to omit violent events that are associated to organized crime or individual criminals.

An additional methodological issue that is addressed by the research team at CERAC is related to the direct and indirect effects of violence. Specifically, a

distinction between the direct and the indirect impacts of violence associated to conflict is established. Such distinction is essential to determine the precision level of the measurements done by CERAC (Restrepo et al 2009).

After sorting out the methodological issues associated to the definition of what should be considered violent conflict events, a strategy to collect and process the information was established by CERAC. As it will become clear in what follows, establishing this strategy was not an easy task either. As highlighted by Restrepo et al (2009), all the methodologies available to collect and process information are subject to bias. Specifically, defining what to measure might lead to selection bias because it is possible to exclude information that should have been included and include information that should have been excluded. In the statistical analysis literature these errors are known as type I and type II errors (Restrepo et al 2009). Additionally, managing the information that has been collected requires important attention in order to avoid errors while processing and organizing the data through the use of a database.

CERAC's Database

A database is a collection of records, which contain one or more fields about a given entity (Halpin and Morgan 2010). There are several types of database models, including the flat and the relational models. Each database model has its own structure and a given number of operations that can be performed within it. The simplest type of database is a flat database, which is a very useful tool for small scale and simple applications (Halpin and Morgan 2010). An example of this database model is a single Excel table containing a list of groups as well as information about these groups, e.g. their members and some of their activities. However, the computational capacity of flat databases is very weak. As a result of the advance in computer science in the last decades, new database models were developed to deal more efficiently with larger and more complex information. In the seventies, E. Codd, a researcher at IBM, developed the relational database to manage information more efficiently. Data is organized in a relational database in such a way that it appears to be stored in a series of interrelated tables (Halpin and Morgan 2010). A relational database is particularly attractive because of two important features. The first feature is that relational databases limit redundancy, or in other words, the duplication of data (Codd 2001). All the information about a given object is stored together, and this information is then linked to related objects so that the original object is not stored in more than one place (Codd 2001). The second feature is that in relational databases data inconsistencies are avoided. This is due to the fact that information is stored in only one place, and therefore the likelihood of inconsistencies is minimized (Codd 2001).

The construction of CERAC's database started in the year 2002. The objective of this database was initially to collect, document, register and analyze information about violent actions related to Colombian armed conflict. In this initial database, the information was codified under a simple database structure or flat database (Restrepo et al 2009). However, as mentioned by Restrepo et al (2009), and also as discussed above, such structure had important limitations in terms of precision and quality. As a result, the model supporting the dataset, the objective of the dataset and the codification strategy evolved and had important improvements (Restrepo et al 2009). Specifically the objective was now to collect information about the violent actions, the event in which these actions occur, the victims of the event, the attackers and the group to which they belong. A transition period was necessary to apply these changes. This transition took place between 2006 and 2007.

The information system implemented in 2006, known as System of Analysis and Registry of Conflict Actions (*Sistema de Análisis y Registro de Acciones de Conflicto – SARAC*), registers documental information in a relational database (Restrepo et al 2009). CERAC's relational database has a collection of two dimensional tables. Each table represents an event about which information is being collected. One of the advantages for CERAC of using a relational database is that it stores the data in separated tables so that the data can be administered dynamically. Specifically, the database stores information that has been codified following certain analytic categories that are related. This implies that each analytical category can also be expressed using tables that are related to each other. Another important feature of CERAC's relational database is that its structure facilitates the normalization of the data. Normalization incorporates several procedures that are designed to remove non-simple domains and the duplication of information. Normalization is an important procedure because it helps to prevent the manipulation of the information collected, the loss of data integrity and it reduces data anomalies (Restrepo et al 2009).

Sources of Information Used in CERAC's Database

The information system uses information from multiple sources. All the sources that are used by CERAC have public access. These sources are categorized as follows: press (regional, national, international), local non-governmental organizations (NGOs), international non-governmental organizations (NGOs), international organizations, governmental institutions and non-governmental groups (Restrepo et al 2009). CERAC keeps an up-to-date record of the different events related to the civil conflict based on events listed in different public sources, including, Colombian newspapers and the annexes to the periodicals Justicia y Paz and Noche y Niebla, which are published quarterly by the Colombian NGOs

CINEP and the Comisión Intercongregacional de Justicia y Paz (Restrepo et al 2006b). It should be highlighted that CERAC's database does not include sources that have a legal reserve, and/or testimonials from primary sources (Restrepo et al 2009).

As mentioned by Restrepo et al (2009), the system through the use of multiple sources has the potential to solve divergences without dropping the original information from each source about the event. At a daily basis, the system receives documental information about violent events that are visible in the context of the Colombian armed conflict (Restrepo et al 2009). The events that are included are those that are directly associated with armed conflict violence.

It should be highlighted that the System of Analysis and Registry of Conflict Actions (SARAC) incorporates geographical information at the municipal level, which is the unit of analysis most commonly used to collect statistical information in Colombia. This tool is particularly helpful for the analysis of information because it is possible to georeferenciate and geocodify conflict events (Restrepo et al 2009).

Quality Control

CERAC follows a very strict process to assure that the data contained in the database is representative and has a high quality. As part of this process, a large number of events are randomly sampled. These events are then compared against the original source of information to assure that the coding of the information was done properly. Additionally, the information from a different randomly selected sample is double-checked using press archives. This is done to make sure that the information taken from the annexes to the periodicals Justicia y Paz and Noche y Niebla is of high quality, and confirm that these events should have been included in the annexes (Restrepo et al 2004) (Restrepo et al 2006b).

Potential Difficulties and Limitations

As it will be discussed in what follows, there are many potential difficulties in the collection and processing of information. Such difficulties should motivate researchers to carry out more careful processes so that the quantification is more precise and the limitations associated to biases and other problems are reduced as much as possible.

An important limitation that the methodology followed by CERAC has to face is related to the availability and continuity of the sources of information. For instance, as mentioned by Restrepo et al (2009), governmental sources did not

report any events associated to paramilitary groups after august 2006. During that month the paramilitary groups signed an agreement with the government to stop participating in violent events. However, other sources of information continued reporting participation of these armed groups in violent events. This entails a methodological challenge to try to accommodate for unavailability of information from a number of sources (Restrepo et al 2009).

It is also important to highlight that the information contained in CERAC's database cannot be considered by any means official or perfect. As Restrepo et al (2009) point out, the database is not intended to be a census of violent events directly associated to the armed conflict in Colombia. This entails that the quantitative information collected by CERAC can only be used to analyze tendencies and relative levels between variables associated to the Colombian conflict. Its objective is not on absolute values (Restrepo et al 2009).

Appendix C. Multilevel Modeling

Given the structure of the dataset employed in this chapter, a three level multilevel model is used for our analysis. Following Raudenbush and Bryk (2002), the first step to implement a multilevel model is to run a fully unconditional model. A fully unconditional model has no specified predictors at any of the three levels and it represents how variation in an outcome measure is allocated across the three different levels (Raudenbush and Bryk 2002). After estimating the unconditional model, the predictors at the different levels are incorporated. For that purpose, a general model at each level is presented following the formulation proposed by Raudenbush and Bryk (2002). The following sub-sections provide a brief description of this process in order to clearly expose the procedure followed in the estimation.

A. Fully Unconditional Model

The fully unconditional model is the simplest model considered in multilevel analysis. The specification of this model does not have predictor variables from any of the levels considered (Raudenbush and Bryk 2002). According to Raudenbush and Bryk (2002), the fully unconditional model can be employed to calculate an estimate of the variance explained at each level explored in the model. Following Raudenbush and Bryk (2002), in this dissertation the fully unconditional model is given by the following three equations:

Level 1

$$T_{ijk} = \gamma_{0jk} + e_{ijk} \quad (\text{Equation 1})$$

Where:

- T_{ijk} is the score in the Colombian High School exit exam for student i in school j and municipality k .
- γ_{0jk} is the mean achievement of school j in municipality k .
- e_{ijk} is a random within student effect normally distributed with a mean of zero and a variance of σ^2 . In other words, e_{ijk} captures the deviation of a student's ijk score from the school mean.

Following Raudenbush and Bryk (2002), the sub-indices i, j, k denote students, schools and municipalities:

10. $i = 1, 2, \dots, n_{jk}$ students within school j in municipality k .

11. $j = 1, 2, \dots, J_k$ schools within municipality k .
12. $k = 1, 2, \dots, K$ municipalities.

Level 2

$$\gamma_{0jk} = \beta_{00k} + \varphi_{0jk} \quad \text{(Equation 2)}$$

Where:

- β_{00k} is the mean score in the High School exit exam in municipality k .
- φ_{0jk} is a random “school effect” normally distributed with a mean of zero and a variance of τ . In other words, φ_{0jk} captures the deviation of school jk ’s mean from the municipality mean.

Level 3

$$\beta_{00k} = \vartheta_{000} + \mu_{00k} \quad \text{(Equation 3)}$$

Where:

- ϑ_{000} is the grand mean in the High School exit exam across municipalities.
- μ_{00k} is a random “municipality effect” normally distributed with a mean of zero and a variance of ψ . In other words, μ_{00k} captures the deviation of municipality k ’s mean from grand mean.

As mentioned previously, the fully unconditional model can be employed to calculate an estimate of the variance explained at each level explored in the model. In other words, the total variability in the outcome can be partitioned into three components. Following Raudenbush and Bryk (2002), the proportion of the variance at each of the levels is given by the following three equations:

- Proportion variance due to differences between students = intra-student correlation = $\frac{\sigma^2}{\sigma^2 + \tau + \psi}$
- Proportion variance due to differences between schools = intra-school correlation = $\frac{\tau}{\sigma^2 + \tau + \psi}$

- Proportion variance due to differences between municipalities = intra-municipal correlation = $\frac{\psi}{\sigma^2 + \tau + \psi}$

It should be highlighted that the random effects at the different levels are assumed to be uncorrelated.

B. Conditional Model

As mentioned previously, the fully unconditional model allows us to calculate an estimate of the variance explained at each level explored in the model. However, as Raudenbush and Bryk (2002) point out, part of the variability observed at each level could be potentially explained by variables that are measured at each of the levels. For example in this case, socioeconomic characteristics of the students (level 1), the teaching experience of the teachers in each school (level 2) and the intensity of the conflict in each municipality (level 3) can be predictors at each level. Additionally, as Raudenbush and Bryk (2002) highlight, some relationships existent at the school and municipal level may possibly vary randomly among the units considered. For those reasons it is necessary to present a general model at each level.

Level 1

At the level 1 (general model), the student achievement is modeled as a function of student level predictors and a random student level error. As a result, following Raudenbush and Bryk (2002), we have the following equation:

$$T_{ijk} = \gamma_{0jk} + \gamma_{1jk}a_{1ijk} + \gamma_{2jk}a_{2ijk} + \dots + \gamma_{pjk}a_{pijk} + e_{ijk} \quad \text{(Equation 4)}$$

Where:

- T_{ijk} is the achievement of student i in school j and municipality k .
- γ_{0jk} is the intercept for school j in municipality k .
- a_{1ijk} are $p = 1, \dots, P$ student characteristics that predict achievement in standardized exams.
- γ_{pjk} are the level 1 coefficients that show the direction and the strength of the relation between the student characteristic a_p and the outcome in school jk .
- e_{ijk} is a level 1 random effect representing the deviation of each student ijk score from the predicted score based on the student level model. It is assumed to be normally distributed with mean 0 and variance σ^2 .

Level 2

For the case of the level 2 (general model), the level 1 coefficients are modeled by school level (level 2) characteristics. The regression coefficients at the student level can be assumed to be fixed, non-randomly varying or random depending on the specification. Therefore, the following general specification of the model for variation among schools within municipalities results:

$$\gamma_{pjk} = \beta_{p0k} + \sum_{q=1}^{Q_p} \beta_{pqk} X_{qjk} + \varphi_{pjk} \quad \text{(Equation 5)}$$

Where:

- β_{p0k} is the intercept for municipality k in modeling the school effect γ_{pjk}
- X_{qjk} is a school characteristic that is used as a predictor of the school effect γ_{pjk}
- β_{pqk} are the level 2 coefficients that show the direction and the strength of the relation between the school characteristic X_{qjk} and γ_{pjk} .
- φ_{pjk} is a level 2 random effect representing the deviation of each school jk level 1 coefficient (γ_{pjk}) from the predicted value based on the school level model. It is assumed to be multivariate normally distributed.

The level 2 model has $P + 1$ equations because there is one equation for each of the level 1 coefficients. It should be noted that it is assumed that the random effects in the level 2 equations are correlated.

Level 3

The level 3 modeling process is similar to the one that was done before for the level 2 model. In this case, each level 3 outcome is predicted by a municipality level characteristic. Therefore, the following general specification of the level 3 model results:

$$\beta_{pqk} = \vartheta_{pq0} + \sum_{s=1}^{S_{pq}} \vartheta_{pqs} W_{sk} + \mu_{pqk} \quad \text{(Equation 6)}$$

Where:

- ϑ_{pq0} is the intercept in the municipality level model for β_{pqk}
- W_{sk} is a municipality characteristic that is used as a predictor of the municipality effect β_{pqk}
- ϑ_{pqs} are the level 3 coefficients that show the direction and the strength of the relation between the municipality characteristic W_{sk} and β_{pqk} .

- μ_{pqk} is a level 3 random effect representing the deviation of the municipality k coefficient (β_{pqk}) from the predicted value based on the municipality level model. It is assumed to be multivariate normally distributed.
- The level 3 model has $\sum_{p=0}^P Q_p + 1$ equations for each municipality.

WORKS CITED

- (CERAC), Centro de Recursos para el Análisis del Conflicto (2013). *Conflict data*.
- (DANE), Departamento Administrativo Nacional de Estadística (2008). Metodología de la Investigación de Educación Formal. In Departamento de Metodología y Producción Estadística (Ed.), (Vol. 2, pp. 35). Bogotá: Departamento Administrativo Nacional de Estadística (DANE).
- (DANE), Departamento Administrativo Nacional de Estadística (2010). Ficha Metodológica de la Investigación de Educación Formal. In D. d. M. y. P. Estadística (Ed.), (Vol. 6, pp. 10). Bogotá: Departamento Administrativo Nacional de Estadística (DANE).
- (FCM), Federación Colombiana de Municipios (2014). *Municipalities Dataset*. Available in: <http://www.fcm.org.co/> Date accessed: June 20, 2014
- (ICFES), Instituto Colombiano para la Evaluación de la Educación (2010). Orientaciones para el Examen de Estado de la Educación Media (pp. 30). Bogota: Instituto Colombiano para la Evaluación de la Educación (ICFES).
- (ICFES), Instituto Colombiano para la Evaluación de la Educación (2013). FTP Databases.
- (IGAC), Instituto Geográfico Agustín Codazzi (2014). *Geographic Databases*. Available in: <http://www.igac.gov.co/igac> Date accessed: June 20, 2014
- Akresh, R., and de Walque, D. (2008). Armed conflict and schooling: evidence from the 1994 Rwandan genocide.
- Annan, J., Blattman, C., Mazurana, D., and Carlson, K. (2011). Civil war, reintegration, and gender in Northern Uganda. *Journal of conflict resolution*, 55(6), 877-908.
- Barnett, R. R., Glass, J. C., Snowdon, R. I., and Stringer, K. S. (2002). Size, performance and effectiveness: cost-constrained measures of best-practice performance and secondary-school size. *Education Economics*, 10(3), 291-311.
- Barrera, F., and Ibáñez, A. M. (2004). Does Violence Reduce Investment in Education?: A Theoretical and Empirical Approach: UNIVERSIDAD DE LOS ANDES-CEDE.

- Belley, P., and Lochner, L. (2007). *The changing role of family income and ability in determining educational achievement* (No. w13527). National Bureau of Economic Research.
- Bellows, J., and Miguel, E. (2006). War and institutions: New evidence from Sierra Leone. *The American economic review*, 96(2), 394-399.
- Blake, J. (1981). Family size and the quality of children. *Demography*, 18(4), 421-442.
- Blake, J. (1989). *Family size and achievement* (Vol. 3): University of California Pr.
- Bundervoet, T. (2012). War, Health, and Educational Attainment: A Panel of Children during Burundi's Civil War: Households in Conflict Network.
- Camacho, A., and Rodriguez, C. (2013). Firm exit and armed conflict in colombia. *Journal of Conflict Resolution*, 57(1), 89-116.
- Cameron, A. C., and Miller, D. L. (2013). A Practitioner's Guide to Cluster-Robust Inference. *Forthcoming in Journal of Human Resources*.
- Chand, V. S., Joshi, S., & Dabhi, R. (2003). Emergency Education': The Missing Dimension in Education Policy. *Educational Research for Policy and Practice*, 2(3), 223-235
- Codd, E. F. (2001). A relational model of data for large shared data banks *Pioneers and Their Contributions to Software Engineering* (pp. 61-98): Springer.
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of family psychology*, 19(2), 294.
- de Leeuw, J., and Meijer, E. (2008). *Handbook of multilevel analysis*: Springer.
- Debalen, A., and Paul, S. (2012). Estimating the causal effects of conflict on education in Cote d'Ivoire. *World Bank Policy Research Working Paper* (6077).
- Downey, D. B. (1995). When bigger is not better: Family size, parental resources, and children's educational performance. *American Sociological Review*, 746-761.

- Earthman, G. I. (2002). School facility conditions and student academic achievement.
- Earthman, G., and Lemasters, L. (1996). Review of Research on the Relationship between School Buildings, Student Achievement, and Student Behavior.
- Earthman, G. I., and Lemasters, L. (1998). Where Children Learn: A Discussion of How a Facility Affects Learning.
- Gamoran, A., and Long, D. A. (2007). Equality of Educational Opportunity A 40 Year Retrospective *International studies in educational inequality, theory and policy* (pp. 23-47): Springer.
- Garson, G. D. (2012). *Hierarchical linear modeling: Guide and applications*: Sage.
- Goldstein, H. (2011). *Multilevel statistical models* (Vol. 922): Wiley. com.
- Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttall, D., and Thomas, S. (1993). A multilevel analysis of school examination results [1]. *Oxford review of education*, 19(4), 425-433.
- Häkkinen, I., Kirjavainen, T., and Uusitalo, R. (2003). School resources and student achievement revisited: new evidence from panel data. *Economics of education review*, 22(3), 329-335.
- Halpin, T., and Morgan, T. (2010). *Information modeling and relational databases*: Morgan Kaufmann.
- Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources*, 14(3), 351-388.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 24(3), 1141-1177.
- Hanushek, E. A., and Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics of education review*, 22(5), 481-502.
- Harbison, R. W., and Hanushek, E. A. (1992). *Educational performance of the poor: lessons from rural Northeast Brazil*: Oxford University Press.
- Haveman, R., and Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of economic literature*, 33(4), 1829-1878.

- Hillesund, S., Kebede, T. A., & Pedersen, J. (2014). CAERUS.
- Justino, P. (2010). War and Poverty. *Handbook of the Economics of Peace and Security*.
- Lai, B. a. C. T. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, 44(3), 277-292. doi: 10.1177/0022343307076631
- Merrouche, O. (2006). The human capital cost of landmine contamination in Cambodia: Households in Conflict Network.
- Ministerio de Educación Nacional (2013). Examen de Estado para el Ingreso a la Educación Superior. Retrieved August 21, 2013, from <http://www.colombiaaprende.edu.co/html/home/1592/article-156080.html>
- Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*: SAGE Publications.
- Restrepo, J., Spagat, M., and Vargas, J. F. (2004). The Dynamics of the Colombian Civil Conflict: A New Data Set.
- Restrepo, J. A., Spagat, M., and Vargas, J. F. (2006). El conflicto en Colombia: ¿quién hizo qué a quién? Un enfoque cuantitativo *Nuestra guerra sin nombre: Transformaciones del conflicto en Colombia*: Editorial Norma.
- Restrepo, J. A., Spagat, M., and Vargas, J. F. (2006b). The Severity of the Colombian Conflict: Cross-Country Datasets versus New Micro-Data. *Journal of Peace Research*, 43(1), 99-115. doi: 10.1177/0022343306059924
- Restrepo, J., and Aponte, D. (2009). *Guerra y violencias en Colombia: herramientas e interpretaciones*: CERAC.
- Rodríguez, C., and Sánchez, F. (2010). Books and guns: the quality of schools in conflict zones. Documento CEDE (38).
- Rodriguez, C., and Sanchez, F. (2012). Armed conflict exposure, human capital investments, and child labor: Evidence from Colombia. *Defence and peace economics*, 23(2), 161-184.
- Sánchez, F., and Díaz, A. M. (2005). Los efectos del conflicto armado en el desarrollo social colombiano, 1990-2002. *CEDE, Centro de Estudios sobre Desarrollo Económico*, Facultad de Economía, Universidad de los Andes.

- Schneider, M. (2002). Do School Facilities Affect Academic Outcomes?. ERIC.
- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, 95(2), 186-200. doi: 10.1016/j.jdeveco.2010.05.002
- Somers, M. A., McEwan, P. J., and Willms, J. D. (2004). How effective are private schools in Latin America?. *Comparative Education Review*, 48(1), 48-69.
- Steele, F. (2008). Module 5: Introduction to multilevel modelling concepts. *LEMMA (Learning Environment for Multilevel Methodology and Applications)*, Centre for Multilevel Modelling, University of Bristol.
- Swee, E. L. (2009). On war and schooling attainment: The case of Bosnia and Herzegovina. *Households in Conflict Network Working Paper*, 57.
- Todd, P. E., and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement*. *The Economic Journal*, 113(485), F3-F33.
- Vandenberghe, V., and Robin, S. (2004). Evaluating the effectiveness of private education across countries: a comparison of methods. *Labour economics*, 11(4), 487-506.
- Wharton, K. and Oyelere, R. (2012). *Conflict and Its Impact on Educational Accumulation and Enrollment in Colombia: What We Can Learn from Recent IDPs*. IZA Discussion Paper. IZA. Retrieved from <http://ssrn.com/abstract=1922011>
- Winthrop, R., and Kirk, J. (2008). Learning for a bright future: Schooling, armed conflict, and children's well-being. *Comparative Education Review*, 52(4), 639-661.
- Wößmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2), 117-170.

EDUCATIONAL ACHIEVEMENT OF SCHOOLS: ASSESSING THE EFFECT OF THE CIVIL CONFLICT USING A PSEUDO-PANEL OF SCHOOLS

Education plays a very important role on the wellbeing of individuals and the economic development of nations. As a consequence, the academic literature that studies education has focused on understanding the factors that can possibly influence schooling outcomes amongst students. For instance, researchers have studied the effect of personal, family and school characteristics on the education achievement of students. Yet, academic researchers have put little attention to other factors that could also affect achievement. More specifically, academic literature has not put enough attention to the effects of civil conflict on educational achievement. The high prevalence of civil conflicts over the last decades further justifies the need to get a better understanding of the relationship between civil conflict and educational outcomes. According to Fearon and Laitin (2014), “a surprisingly high percentage of independent states have experienced at least one civil war in the years since World War II - a bit more than two out of every five at the 1000 total deaths threshold”. These researchers further highlight that “almost one in ten countries with a population of at least half a million in 2000 had civil war occurring in at least half of its years since 1945 or independence”. Understanding the relationship between civil conflict and academic achievement is indeed a key to the adequate design of policies geared towards those individuals who have been affected by violence. The purpose of this chapter is to explore the relationship between civil conflict and educational achievement in Colombia through the use of pseudo-panel analysis¹⁸.

In Colombia there are no panel datasets available to study the educational achievement of students enrolled in the formal educational system, but there is a very complete series of independent cross-sections managed by the Colombian Institute for the Evaluation of Education (ICFES). The availability of a rich cross-sectional dataset has its advantages and drawbacks. As Verbeek (2008) highlights, the major limitation of repeated cross-sectional data is that the same individuals are not followed over time, and therefore, individual stories are not available to transform a model to first-differences or in deviations from individual means. Nevertheless, cross-sectional data also has an advantage because it is less prone to

** The contents of this chapter correspond to the second publishable paper of this PhD dissertation.

¹⁸ Deaton (1985)

typical panel data problems like attrition and nonresponse (Verbeek 2008). Due to the lack of longitudinal data, in this chapter we construct a pseudo-panel of schools from the Saber11 examination cross-sectional data sets. These datasets contain the individual level results from the years 2000, 2002, 2005 and 2007. Those specific years were chosen for the analysis because the National Department of Statistics only released information about school characteristics for those given years. According to Deaton (1985), a pseudo-panel can be constructed in order to track ‘cohorts’, i.e. a group with fixed membership, just as if panel data were available. In this case, the individual results are grouped in school ‘cohorts’.

This chapter contributes to the literature by examining the effects of civil war on a very important indicator of social wellness: education. Researchers have not paid enough attention to a number of factors that cause internal distress in countries affected by a civil conflict (Benefield and Tomlinson 2005) (Bruck et al. 2010). One of these factors is educational achievement. There are a number of reasons for the limited availability of academic research. One possible reason is the lack of reliable data of the conflict and other control variables. We also believe that there is a tendency to underestimate the need for an academic analysis because of the general assumption that the negative effects of a civil conflict are “obvious”. Such tendency is clearly limiting our understanding of the effects of civil conflict. Certainly this is problematic because educational outcomes are not only an end in themselves, but they are also closely related to other indicators of social wellbeing. Having a better understanding of the impact of civil conflict on educational outcomes, and therefore of the specific needs of the schools affected by it, has very important implications for policy analysis and policy formulation.

This chapter differs significantly from previous efforts¹⁹ because we construct a pseudo panel of schools in order to get a better understanding of the relationship between civil conflict and education. In fact, to the best of my knowledge this is the first effort to try to understand the relationship between conflict and achievement in standardized exams through the use of a pseudo panel of schools.

The rest of this chapter is organized as follows: *section 2* presents a brief description of the Colombian conflict during the period under analysis; *section 3* presents a review of the literature that uses a pseudo-panel approach to study educational outcomes; *section 4* presents a theoretical framework that serves as backbone for the econometric exercise; *section 5* describes the data and the variables that are being used in the construction of the pseudo panel; *section 6* describes the methodology employed for the analysis and discusses the

¹⁹ Amongst those efforts we can highlight the work of Swee (2008), Akresh and de Walque (2008), Merrouche (2006), Lai and Thyne (2007), Shemyakina (2011), Debalen and Paul (2012), Bundervoet (2012), and Bellows and Miguel (2006).

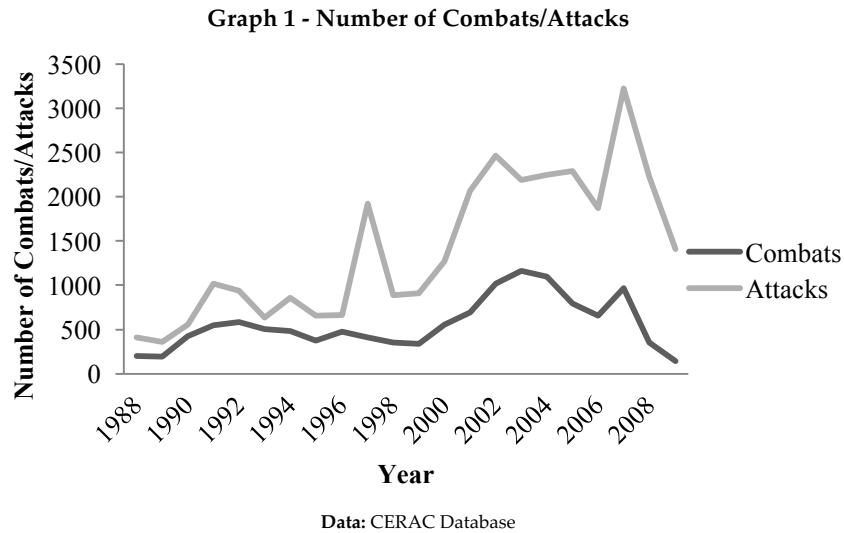
specification of the model; *section 7* reports the main results; *section 8* presents a discussion of the results; and *section 9* presents conclusions and policy implications.

2. Colombian Conflict During The Period Under Analysis

Guerrilla groups went through a period of reformation between the years 1991 and 1994 because of a major cut back in available resources due to the end of the Cold War. As part of these reforms, a number of changes in their operative and financial structure were introduced. In fact, during the first half of the nineties the guerrillas expanded their territorial presence throughout the country. Nevertheless, according to Restrepo et al. (2006) during these years violence did not increase. The reforms that were introduced during the first half of the nineties had a lagged effect on the conflict dynamics. It was only in the year 1996 that the dynamic of conflict started to change significantly as an important increment in the intensity of conflict, and also in the frequency of combats and attacks became evident. These changes gave a new air to the guerrilla groups, allowing them to grow and gain power during the nineties. In fact, Restrepo et al. (2006) called this period the *escalation of violence period* given the increase in the intensity of the conflict and the frequency of the combats and the attacks. According to these authors, the *escalation period* took place between the years 1996 and 2002, year in which the armed conflict reached the highest intensity. The recrudescence of violence in Colombia in the period between the years 1996 and 2000 could be explained by a number of factors (Restrepo et al. 2006). Some of these factors include: the unification of the majority of paramilitary groups into one organization (Autodefensas Unidas de Colombia), the implementation of a terror strategy against civilians, and the modernization of the Colombian military forces through the implementation of new technologies in order to have a better response against the attacks of illegal groups. During *the escalation period*, the illegal groups implemented a strategy that focused in attacking civilians as an instrument to terrorize the society. In fact, the highest number of civilian fatalities during the conflict (global maximum) was reported in the year 2001 (*Graph 2*). By the end of the escalation period in 2002 the number of civilian fatalities started to decrease, although the number of combats and attacks remained high (*Graphs 1 and 2*).

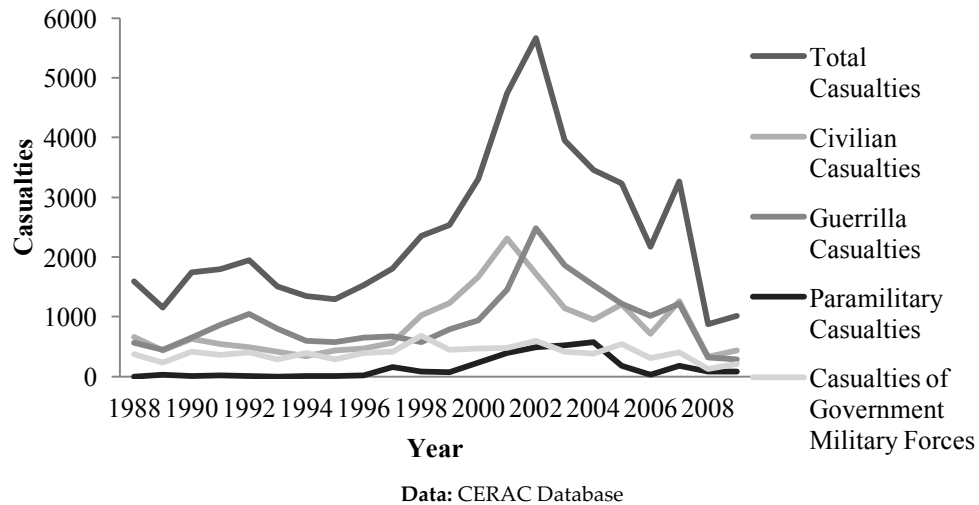
In the period that goes from the years 2003 to 2008, several changes were introduced by the Colombian government in order to deal with the escalation of violence and the problems associated with it. In fact, Restrepo et al. (2006) have called this period *the reorganization period*. During this time period not only did the government introduce important reforms, but also the illegal military groups were

forced to go through a re-organization period in order to react to the increasing pressure from the governmental military forces (Restrepo et al. 2006). Nevertheless, the high levels of violence related to conflict persisted. This remark becomes evident if we analyze the total number of conflict events reported in the period between the years 2003 and 2007. In fact, as we can see in Graph 1, conflict events reach a maximum at the end of this “period of re-organization” (Restrepo et al. 2006).



Restrepo et al. (2006) points out that during the restructuring period, the professionalization and expansion of the Colombian military forces led to important changes in the strategic scenario of the anti-guerrilla war. In effect, several organizational and operative changes introduced by the governmental military forces led to a substantial difference in the warfare capacity between the governmental forces and the insurgent forces (Restrepo et al. 2006). At the beginning of the restructuring period, President Uribe’s administration promised to introduce changes in the internal security and defense strategies. A tributary reform was introduced at that time to collect tax money in order to finance the reforms that were necessary to improve the country’s security levels (Restrepo et al. 2006). As Restrepo et al. (2006) highlight, the additional financial resources that were obtained from tax money and the extension of the Plan Colombia allowed the government to renovate the obsolete armaments and purchase new technologies in order to improve the capacity of the governmental military forces.

Graph 2 - Number of Casualties



Between the years 2003 and 2005, we can see that the number of conflict events remains relatively unchanged, but going from mid-2005 to mid-2006 the number of events decreases significantly. This decrease in conflict events could be explained by the emergence of the Plan Patriota, a governmental military plan that was designed to try to seize control over the guerrillas in the southern areas of Colombia. However, this plan did not achieve the expected goals and in 2007 it was cancelled. After that year, conflict events started to increase and in 2008 the highest number of conflict events was registered in the series (Restrepo et al. 2006). Even though, the number of attacks registered during this period is the highest in the series, the intensity of the conflict decreased. This decrease could be explained by two aspects. First, the number of civilians dying because of the conflict decreased starting in the year 2001 due to a reduction in the number of victims caused by paramilitary attacks (Graph 2). In fact, the demobilization of a growing number of paramilitaries from the year 2003 to the year 2007 contributed to the expansion of governmental control over areas of the Colombian territory, and therefore helped in the improvement of overall security conditions (DeShazo et al. 2007). It should be noted that the paramilitary groups were responsible for many of the civilian casualties in the Colombian conflict during the nineties and beginning of the 21st century (Restrepo et al. 2006). Second, starting in the year 2003 there was a decline in the number of casualties of combatants. Such reduction can be explained by the guerrilla's operational re-organization, given their weaker position in comparison to the governmental forces, which forced them to move to remote areas in the highlands and the jungle (Restrepo et al. 2006).

As Restrepo et al. (2006) highlight, even though there was a reduction in the victimization levels, the number of victims was still very high, reaching levels above those registered in the mid-nineties. There are two aspects that could explain the persistence of high levels of violence associated to the civil conflict during this time period. First, the neo-paramilitary forces, which emerged after the demobilization of the paramilitary forces, kept the paramilitary strategies that mainly targeted civilians (Restrepo et al. 2006). Second, the guerrilla war against the FARC was still causing many civilian casualties (Restrepo et al. 2006).

In this chapter, the relationship between civil conflict and educational achievement will be explored concentrating on the time period identified above. As mentioned previously, a civil war has important economic, social and political implications that potentially interfere with the development process of the countries affected by it. More specifically, a civil conflict could potentially affect the human capital accumulation process via education. Taking advantage of a unique database about civil conflict events in Colombia administered by a Colombian think tank (Resource Centre for Conflict Analysis - CERAC), a government kept database containing the results of standardized tests at different levels of schooling, and additional information about schools and teachers from the National Department of Statistics of Colombia (DANE), the objective of this chapter is to get a better understanding of the relationship between civil conflict and educational achievement, and quantify the magnitude in which civil conflict affects the accumulation of human capital measured by cognitive exams.

3. Literature Review

The literature that studies conflict and education was reviewed in detail in Chapter 4 of this dissertation. Nevertheless, it is also relevant to review the literature that has employed the methodology applied in this chapter, pseudo panel analysis, to study educational outcomes. A number of researchers have contributed to the literature by applying a pseudo-panel methodology to study the determinants of educational achievement and human capital accumulation in countries that do not have longitudinal data. Amongst the academic research that has contributed to this literature we can find the work of Narita (2008), Emery et al. (2012), Simone (2012), Bleakley (2013), Braga et al. (2013), Warunsiri and McNown (2010), Pal and Saha (2014) and Brunello and Comi (2004). It is notable that to the best of my knowledge no effort has been made to try to understand and quantify the impact of civil conflict on education achievement measured by standardized exams through the use of a pseudo panel methodology. In fact, to the best of my knowledge the only effort to try to understand the relationship between civil war and educational outcomes, using a pseudo-panel, was published by Bleakley (2013). However, this

article significantly differs from this chapter because Bleakley (2013) studies the effect on school attendance instead of the effect on examination results, as we do.

It is important to recognize that the review of the literature presented in this chapter concentrates on the economic literature that studies the relationship between civil conflict and education. Nevertheless, there are other areas of study that have also explored this relationship. The reason why we have decided to concentrate our attention on economic literature is that this literature has a stronger quantitative tradition. In what follows, the existing research that has employed pseudo panel analysis to study educational achievement and human capital accumulation will be reviewed in order to contextualize the application of this methodology using educational data and justify the relevance of this chapter.

Some authors, including Simone (2012), Pal and Saha (2014), Narita (2008) and Braga et al. (2013), have studied the determinants of educational achievement using a pseudo-panel data methodology. Simone (2012) studies the determinants of scores in mathematics and science at the end of the lower secondary school (8th grade) using a pseudo panel approach. This author takes into consideration the cumulative effects of education by controlling for the estimated achievement at the 4th grade. The pseudo-panel approach is used in this paper in order to link achievements of the same cohort of Italian students over two waves of the Trends in International Mathematics and Science Study (TIMSS). The author is able to distinguish the importance in generating learning gaps of primary and lower secondary education by incorporating the estimated achievement at the 4th grade as part of the pseudo panel approach. Lacking longitudinal data on students' achievement, information collected on the same cohort of students (repeated cross sections) is connected with a pseudo-panel approach to conduct a system level evaluation of the dynamics of cognitive achievement. Simone (2012) finds that the gender gap in math observed at the grade 8 should be attributed to primary education, while the gap in science should be attributed to the two school levels. In the same line of analysis of Simone (2012), Pal and Saha (2014) study the determinants of school choice and school efficiency, in terms of secondary school's exit exam test scores, by constructing a pseudo-panel using student-level data from Nepal. Given the unavailability of longitudinal data, these authors rearranged a cross-section of student-level data for 2004 to generate subject-level scores for six compulsory subjects of the exit examination for each student, which can be treated as a pseudo panel. By constructing the pseudo-panel, the authors are able to exploit the inter-subject variation in the examination scores, and in doing so, eliminate household/student-level time-invariant omitted factors. Pal and Saha (2014) find that socially motivated schools outperform both profit-motivated schools and government schools. Based on these results, the authors argue that pro-poor philanthropic activity can have a strong efficiency dimension

(Pal and Saha 2014).

Following the same line of Simone (2012) and Pal and Saha (2014), Narita (2008) uses a sample of women, aged 23 to 30 years, from the Brazilian Household Surveys (PNAD 1992-2004) to find the causal effects of teenage women fertility to their later educational achievement, labor market participation and earnings. The author uses information at the state of birth level to keep track of the past social and economical conditions of individuals at the state where they were born. Narita (2008) constructs a panel with averages by state of birth (27 states) and cohorts (1961-1981), so that state of birth and cohort specific components are interpreted as individual specific components, and there are data overtime to account for variation at the state of birth and cohort level. The findings of Narita (2008) do not reject the hypothesis that anticipated motherhood is detrimental for high school completion and labor force participation at women subsequent ages. Braga et al. (2013) also analyze the determinants of educational achievement using pseudo-panel data. These authors study the effects of educational reforms on school attainment in order to get a better understanding of the impact of educational reforms on the distribution of educational attainments in Europe. Braga et al. (2013) use a dataset that contains information for 24 countries on various reforms that affected the school design over the 1929 to 2000 period. They create a pseudo panel using birth cohorts as the time, pooling all surveys' waves together in order to maximise the number of observations available per year of birth. The results obtained by Braga et al. (2013) show that despite market incentives, educational choices can be shaped by policy makers, who may affect both the mean and the variance of the distribution of attainments.

Exploring yet a different determinant of educational achievement, Emery et al. (2012) analyses the long-term effects on human capital formation of natural resource booms using a pseudo panel that follows birth cohorts over time. These authors focus on education outcomes in the province of Alberta during Alberta's 1973 to 1981 oil boom. It should be noted that the case of Alberta was chosen because this province has the large majority of Canada's oil reserves. Emery et al. (2012) conclude that economic booms may change the timing of schooling, but there are no significant long-term negative effects on the total accumulation of human capital associated with them.

Other authors, including Warunsiri and McNown (2010) and Brunello and Comi (2004), have constructed a pseudo-panel to study the returns to education. Warunsiri and McNown (2010) use a pseudo-panel approach to estimate the returns to education in Thailand. These authors construct a pseudo-panel using the age of each individual at the time of the survey to establish the birth cohort to whom the worker belongs for those born between 1946 and 1967. The application

of a pseudo-panel approach enables Warunsiri and McNown (2010) to control for unobservable individual characteristics, such as ability or motivation, which may bias the estimated rate of return to education. Based on the pseudo-panel estimations, the authors find that the overall rate of return to education in Thailand is between 14% and 16%. According to Warunsiri and McNown (2010), those results are considerably higher than those estimated in previous studies that have used different methodologies. On the other side, Brunello and Comi (2004) study the relationship between experience–earnings profiles and education based on a pseudo panel of two cohorts of individuals from 11 European countries from the 1980s and the 1990s. These authors found evidence supporting the view that earnings grow with experience significantly faster for the more educated.

To the best of my knowledge, the only effort to try to understand the relationship between civil war and educational outcomes, using a pseudo-panel, was published by Bleakley (2013). Bleakley (2013) studies the decline of school attendance among white children in the Southern US after the Civil War. This author measures the path to literacy as a function of time in school by constructing a pseudo-panel from the cross sectional data of school ages in 1860 and 1880. This pseudo-panel was constructed under the assumption that the cross-section in a given year is reasonably representative of the behavior of the panel at that point in time (Bleakley 2013). Bleakley (2013) compares how quickly literacy and years of schooling increase with age in those two years, and interprets the relationship between literacy and years of schooling as the marginal benefit to literacy from time in school. Bleakley (2012) argues that if school quality had declined, we would expect this relationship to be less strong over time. The results show the same relationship before and after the war between literacy and time in school, suggesting that Southern students in 1880 were on track to achieve the 1860 level of literacy if they would have spent the same amount of time in school.

The existing literature that employs pseudo panel analysis to study educational achievement and human capital accumulation is very limited. Given the lack of longitudinal data to explore the determinants of educational achievement in many countries around the world, the possibility to construct a pseudo-panel from repeated cross-sectional data becomes particularly attractive for educational researchers. One of the contributions of this chapter is precisely related to the application of this methodology. To the best of my knowledge, this is the first attempt to employ this methodology to try to understand and quantify the impact of civil conflict on educational achievement measured by standardized exams. We are contributing in the search for new and better ways to understand the relationship between civil conflict and educational achievement.

4. Theoretical Framework

The education production function is used in this chapter as the framework for estimating the effect of civil conflict on educational attainment at the school level. In his seminal papers, Hanushek (1979 and 1986) presents education as a production process that uses school and other inputs to produce educational outputs. The economics of education literature has traditionally used an educational production function to model schooling decisions (quantity and quality of education). Following Todd and Wolpin's (2003) notation, let T_{it} be a measure of achievement for student i in period t . Inputs are given by the choices that are being made by parents and schools. Following Todd and Wolpin's (2003) notation, the vector of inputs supplied by the parents at a given age is represented by F_i , and the vector of school-supplied inputs is represented by S_i . D_i denotes external factors (in this case the intensity of the conflict) affecting student i . The student's endowed ability is denoted as μ_{i0} , and a measurement error in test scores as ε_i . Given those inputs, the production function is represented as:

$$T_i = T(F_i, S_i, D_i, \mu_{i0}, \varepsilon_i) \quad (1)$$

In this chapter, the output of interest is the performance of school cohorts at the Colombian High School Exit Examination. We are going to focus on examination results because of data availability and because it is well known that the performance of students in these exams is widely used by students, parents, universities and policy makers as a measure of scholastic performance.

As discussed in Chapter 5, it should be acknowledged that summarizing the learning processes in a function could be subject to many critics due to its complexity. Several issues should be taken into consideration when using the educational production function as the theoretical backbone for analysis. The first issue is related to the decision about which variable should be used to measure educational output. As mentioned previously, the variable that we chose as a proxy for educational output is the result in the Colombian National Exit Examination. The decision to choose that variable and not a different proxy, is subject to debate in the education literature. The literature has not reached consensus regarding the best method to measure the outcomes of schooling. In fact, even though an important number of studies have used standardized achievement test scores, there are some studies that use other indicators of

educational output such as school attendance rates and dropout rates²⁰. Nevertheless, in our case the availability of data on examination results and the general acceptance amongst students, parents, universities and policy makers in Colombia and internationally of the examination results as a measure of scholastic performance, guided our choice of this proxy for educational output.

A second issue to be considered is related to the choice of variables to be included in the production function. As discussed in Chapter 5, a complete interpretation of the determinants of educational attainment would have to include social and cultural factors, and perhaps more importantly, the innate endowment of the students. However, some of these factors are not observable²¹ and therefore are not quantifiable. In that sense, it is impossible to include those factors in a production function that could be estimated empirically. For that reason, an important issue in the estimation of an educational production is the possible lack of accuracy of measurement. Nevertheless, it should be highlighted that some of these issues are more relevant when using disaggregated data to model individual student performance than in the case of a school level analysis of educational performance. For instance, some non-observable student inputs, like innate endowments, should average out at the same level across schools in a school level analysis, like the one done in this chapter. In that sense, a school level analysis is less demanding in terms of data than a student level analysis. However, the results should still be analyzed carefully acknowledging this possible source of bias.

5. Data and Variables

The main data source for this chapter is the Instituto Colombiano para la Evaluación de la Educación (ICFES). Cross-sectional datasets, containing the individual-level nation-wide results of the Colombian High School exit examination for the years 2000, 2002, 2005 and 2007, are employed in the construction of the pseudo-panel that is used for the analysis. In order to be able to follow the same schools, we only use the results of the examinations administered in the second semester of each year. Additionally, a dataset containing the characteristics of the schools and the teachers who work at these schools was obtained from the Colombian National Department of Statistics (*Departamento Administrativo Nacional de Estadística - DANE*). The conflict data was obtained from the Resource Centre for Conflict Analysis (*Centro de Recursos para el Análisis del Conflicto – CERAC*).

²⁰ A number of papers that use these indicators were included in the literature review which can be found in Chapter 4.

²¹ See Chapter 5 for a more complete discussion about this topic.

5.1 Sources of Information

Colombian High School Exit Examination – ICFES Saber 11

In Colombia there are no panel datasets available to study the educational achievement of students enrolled in the formal educational system, but there is a very complete series of independent cross-sections managed by the Colombian Institute for the Evaluation of Education (ICFES). The Colombian nation-wide exit examination (ICFES Exam) was first created more than forty years ago in order to provide support to universities and other higher education institutions in their admissions processes²². Databases containing the nation-wide results of the Colombian high school exit exams (Saber11 exams) have been made public by ICFES for research purposes. In this chapter, ICFES cross-sectional data from the years 2000, 2002, 2005 and 2007 is employed.

The C-600 Census – National Department of Statistics

With the support of the Ministry of Education, the Colombian National Department of Statistics (DANE) has been collecting information about schools and teachers since 1978. This information has been collected through the C-600 statistical census. The C-600 is a statistical census with an annual periodicity that collects information about enrolled students and teachers from all formal education schools in Colombia. Databases containing the nation-wide information for the C600 census for the years 2000, 2002, 2005 and 2007 were made available by the Colombian National Department of Statistics (DANE) for this chapter.

CERAC's Colombian Conflict Dataset

Information about the intensity of the Colombian conflict was obtained from the Resource Centre for Conflict Analysis (CERAC). CERAC is a think tank based in Bogotá (Colombia) with three main lines of research: violent conflict, political violence, and violence associated to drug trafficking. Following these lines of research, an important area of study at CERAC is the quantification of violence directly associated to the internal armed conflict in Colombia. To quantify violent armed conflict, the research team at CERAC developed a methodology to collect, document, code and register information about this type of violence. The construction of CERAC's database started in the year 2002. This database contains information about the violent actions, the event in which these actions occur, the victims of the event, the attackers and the group to which they belong. More

²² More information about this examination and the ICFES dataset was provided in Chapter 6 of this dissertation.

information about the methodology used by CERAC to construct the database is available in *Appendix B of Chapter 7*.

5.2 Variables

Table 1 summarizes all the variables considered in the analysis. A description of these variables is provided in what follows.

The Dependent Variables

Our analysis focuses on two outcome variables: the score in the mathematics exam and the score in the language exam. Even though the results of the examinations in other subjects are available, we have decided to restrict the analysis to mathematics and language because these two subjects capture two very important competencies: quantitative ability and verbal ability. However, this quantitative analysis could potentially be extended to study other subjects that are also tested in the Colombian High School exit examination.

Table 2 presents the statistics of the average scores obtained by the different schools in the mathematics and language High School exit examinations. The results of the Saber11 exam in each of the categories evaluated are reported in a 0 to 100 scale for the years under analysis. The number of observations reported correspond to the number of students in the sample considered for each year.

Table 1 Variable Selection	
Dependent Variable	Score in mathematics and language high school exit exam (high school exit exam)
Independent Variables	Average age of the students taking the exam, percentage of male students, percentage of students who study during nighttime, number of teachers per student, cost of tuition, number of students taking the examination, total expenditure of the municipality in education (log), presence of conflict ²³ (number of conflict events) and intensity of the conflict ²⁴ (number of civilian deaths).

Source: author

²³ In the specifications that were estimated for this empirical exercise we also included lagged measures of this variable.

²⁴ In the specifications that were estimated for this empirical exercise we also included lagged measures of this variable.

Table 2 – Language and Math Scores in the High School Exit Examination (mean and standard deviation)								
	2000-2		2002-2		2005-2		2007-2	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Math Score	42.73	4.94	43.15	6.14	45.03	7.06	45.91	8.43
Language Score	46.90	6.17	48.54	6.73	47.16	8.13	46.81	6.98
Observations (students)	351170		364404		357533		419307	

Source: author's calculations

The Independent Variables

According to the literature, individual background variables play a very important role on educational performance (Häkkinen et al. 2003, Haveman and Wolfe 1995, Woßmann 2003). For that reason, a variable capturing the average age of the students taking the exam at each of the schools and a variable capturing the percentage of students who work were included in the analysis. Additionally, we included a control variable to capture the percentage of male students taking the examination in each of the schools. The academic literature has also found evidence showing that the characteristics of the schools and teachers have a crucial role in determining educational achievement. In fact, Gamoran and Long (2006) have found that the characteristics of the school have an important impact on academic performance in developing countries. Therefore, it is very important to control for the characteristics of the schools and the teachers that are considered to be relevant for the sample under study. Five different variables are included to account for school characteristics: number of students per teacher, cost of tuition, and number of students enrolled. Tables 3 and 4 summarize the descriptive statistics of the variables considered in the analysis, including the minimum and maximum values and the mean and standard deviation.

There are also good reasons to include the number of teachers per student available in the school as an explanatory variable. If a large group of students receives class from only one teacher, most likely the attention received will be less and this might have a negative effect on the achievement of students in evaluations. Additionally, the low number of teachers per student may be a factor discouraging school attendance. Nevertheless, the literature has not shown conclusive results when this variable is included as a determinant. Harbison and Hanushek (1992) reviewed the literature of the impact of teacher-pupil ratios and found that out of sixteen studies with statistically significant effects, eight were

positive and eight were negative. These results cast doubt about the reliability of previous studies, but serves as an additional motivation to include this variable as part of this research.

Table 3 – Explanatory Variables								
	2000-2				2002-2			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Number of teachers per student	0.37	0.38	0.015	2.93	0.33	0.49	0.014	3
Number of students taking the examination per school	53.52	48.98	10	911	48.39	44.02	11	1038

Source: author's calculations

Table 4 – Explanatory Variables								
	2005-2				2007-2			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Number of teachers per student	0.19	0.48	0.005	2.33	0.21	0.64	0.003	2.86
Number of students taking the examination per school	46.13	45.62	10	1110	50.05	46.07	12	1056

Source: author's calculations

There could be significant differences in terms of the tuition that students have to pay at each of the schools considered in our analysis. Such differences could help us capture dissimilarities in the amount of resources available at each school. For that reason, it is important to control for the cost of the tuition at each school. To do so, we have included an explanatory variable to capture the cost of attending each school. For the econometric exercise presented in this chapter, we have decided to divide the cost of tuition variable into three categories comprising the different levels of tuition paid by the students at each school in the sample. The three categories are the following: low tuition (students pay less than \$25 dollars

per month), mid tuition (students pay between \$25 dollars and \$100 dollars per month) and high tuition (students pay more than \$100 dollars per month). Table 5 presents information about the tuition that has to be paid in the different schools that are included in the dataset.

Table 5 – Cost of tuition				
	2000-2	2002-2	2005-2	2007-2
	%	%	%	%
Low Tuition	79	74	87	88
Mid Tuition	12	13	10	9
High Tuition	9	13	3	3

Source: author's calculations

Central government resources are also an important source of income for schools. Every year each municipality receives resources that are assigned by the General System of Participation (Sistema General de Participaciones)²⁵. In order to capture that fact, and also to account for financial differences between municipalities, a variable that captures the resources that are transferred by the central government to the territorial entities to finance education expenditure is included in the model. This variable was transformed to logarithms in order to be included in the specification that was used for the estimation.

International literature has shown that the economic, political and social consequences of civil war are significant (Justino 2010). However, the connection between conflict and education is an under-analysed area of study. This is partly due to the difficulties involved in the systematic collection of evidence on why, where and how attacks occur. As mentioned by Justino (2010), high quality monitoring data and systematic reporting of events are very limited. Additionally, in many places affected by conflict censorship of information is also common. The availability of a rich database about conflict events in Colombia, maintained by the Resource Centre for Conflict Analysis (CERAC), opens new possibilities for the analysis of the relationship between conflict and education. Using information from CERAC, we have included variables that capture the intensity of the conflict and the presence of conflict as the central components of this empirical analysis. The intensity of the conflict is quantified by the number of civilian deaths in each municipality. However, in some of the specifications considered we also included

²⁵ Refer back to Chapter 3 for more information about the General System of Participation.

an additional measure of intensity: a dummy variable that takes a value of 1 if the municipality recorded an intensity of conflict above the national average. The presence of conflict is quantified by the number of conflict events in each municipality. In this case we also included an additional measure of presence of conflict: a dummy variable that takes a value of 1 if the municipality recorded conflict events above the national average.

6. Methodology

A very important methodological issue that is addressed in this chapter is related to the structure of the ICFES datasets. These datasets are only cross-sectional. Nevertheless, we would ideally like to have access to a longitudinal database of students or schools in order to trace their educational performance through time. A major limitation encountered in any empirical analysis that uses repeated cross-sectional data is that the same individuals (students) are not followed over time, and therefore, the individuals' (students') histories cannot be included in the model that is being specified (Verbeek 2008). However, it is still possible to identify and estimate parameters in a panel data model constructed from pooled cross-sections (Imbens and Wooldridge 2007). A model with those characteristics is known in the literature as a pseudo-panel model (Deaton 1985).

As mentioned in a previous section, the economics of education literature has traditionally used an educational production function to model schooling decisions (quantity and quality of education). Following Todd and Wolpin's (2003) notation, let T_{it} be a measure of achievement for student i in period t . Inputs are given by the choices that are being made by parents and schools. Following Todd and Wolpin's (2003) notation, the vector of inputs supplied by the parents at a given age is represented by F_i , and the vector of school-supplied inputs is represented by S_i . D_i denotes external factors (in this case the intensity of the conflict) affecting student i . The student's endowed ability is denoted as μ_{i0} , and a measurement error in test scores as ε_i . Given those inputs, the production function is represented as:

$$T_i = T(F_i, S_i, D_i, \mu_{i0}, \varepsilon_i)$$

Given the data availability, it seems feasible to estimate the following lineal model of pooled cross-sections:

$$EXAM_{it} = \beta X_{it} + \alpha_i + \varepsilon_{it}$$

Where β is the coefficient vector of interest and α_i is an individual fixed effect.

Assuming that the individual effects were not correlated with the explanatory variables, it would be possible to estimate a linear model of pooled cross-sections using ordinary least squares (OLS) estimation. However, using a linear model of pooled cross-sections is not the best alternative. Some of the explanatory variables are individual and household characteristics, and therefore, it is likely that these variables are correlated with the individual effects. If the explanatory variables are correlated with the individual effects, the estimates for the coefficients might be biased.

Having a panel dataset would be the best alternative in this case. Using a panel dataset it is possible to overcome the issue discussed above by employing a fixed effects estimator controlling for all individual specific and time invariant heterogeneity. However, as we mentioned previously the ICFES datasets are only cross-sectional, and therefore given the lack of genuine panel data it is not possible to implement this solution. Nevertheless, following the methodology proposed by Deaton (1985), it is possible to construct a pseudo-panel with the information that is available.

According to Verbeek (2008), the construction of a pseudo-panel is based on grouping individual data into cohorts. Consequently, the definition of cohorts is a critical step (Verbeek 2008). Given that cohorts are followed over time in pseudo-panel analysis, the construction of the panel should be based on characteristics that are time invariant. For instance, this implies that if a student is drawn into the sample in two of the cross-sections, this student would always end up in the same cohort. In this case, the time invariant characteristic is the affiliation to a given school. Therefore, we construct our pseudo-panel dataset by calculating cohort (school) means in each available cross-section. For the construction of a pseudo-panel, it is also necessary to have sufficient within pseudo-individual variation so that the variables of interest can be identified.

6.1 Pseudo Panel Estimation

As mentioned previously, a very important methodological issue that is addressed in this chapter is related to the structure of the ICFES datasets. These datasets are only cross-sectional. Nevertheless, we would ideally like to have access to a longitudinal database of students or schools in order to trace their educational performance through time. In order to overcome this difficulty, we constructed a pseudo panel of schools following Devereux (2007) and Verbeek (2008).

To estimate the pseudo panel of schools it is necessary to use an errors in variables estimator. This estimator corrects for errors of measurement in the cohort means and it also accounts for potential correlation between the errors of measurement in the dependent and the explanatory variables. The literature has shown that if the dependent variable and the independent variables are imperfect measures of the true value, and the errors of measurement are correlated, then the estimated coefficient will capture the correlation of the errors.

In what follows we will provide a brief description of the procedure that we follow in order to estimate the pseudo panel constructed for this chapter.

The starting point is a simple linear model with individual effects:

$$y_{it} = x'_{it}\beta + \alpha_i + u_{it} \quad (1)$$

where $t = 1, \dots, T$ is used to denote time, i is used to denote individuals, x_{it} is a K dimensional vector of explanatory variables, α_i is an individual effect and β is the parameter vector of interest. We also assume that the error term is not correlated with α_i and x_{it} . For simplicity, we also assume that $E\{x_{it}u_{it}\} = 0$ for each t .

According to Verbeek (2008), if the individual effects α_i are uncorrelated with the explanatory variables in x_{it} , model (1) can be estimated consistently from repeated cross-sections by pooling all observations and performing ordinary least squares treating $\alpha_i + u_{it}$ as a composite error term. However, the individual effects can be correlated with some or all of the explanatory variables.

Deaton (1985) has shown that it is possible to use cohorts to obtain consistent estimators for β in model (1) when repeated cross-sections are available, even if α_i is correlated with one or more of the explanatory variables. The cohorts C are groups of individuals sharing characteristics in common. The groups are defined so that each individual is a member of exactly one cohort, which is the same for all periods.

Aggregating²⁶ all the observations to cohort level:

$$\bar{y}_{ct} = \bar{x}'_{ct}\beta + \bar{\alpha}_{ct} + \bar{u}_{ct} \quad (2)$$

Where $c = 1, \dots, C$ is used to denote the cohorts, t is used to denote time, \bar{y}_{ct} is the average value of all observed y_{it} 's in cohort c in period t , and \bar{x}_{ct} is the average

²⁶ In this context aggregating is equivalent to taking means by cohort.

value of all observed x_{it} 's in cohort c in period t . The dataset used in this model is a pseudo panel with repeated observations over T periods and C cohorts.

According to Devereux (2007), the standard approach to estimate *equation (2)* is to replace $\bar{\alpha}_{ct}$ with cohort dummies and use ordinary least squares (OLS) or weighted least squares if the number of observations in each group is different. This approach is known in the literature as the efficient Wald estimator²⁷. Devereux (2007) points out that this estimator is consistent if the number of observations goes to infinity even if α_i and x_{it} are correlated. If the number of observations is finite, the efficient Wald estimator would show biased estimates because $\bar{\alpha}_{ct}$ (the cohort effect) is not constant over time, i.e., $cov(\bar{\alpha}_{ct} - \alpha_c, \bar{x}_{ct}) \neq 0$, where α_c is the true cohort effect (Devereux 2007). $\bar{\alpha}_{ct}$ is not constant over time because the individuals that are being sampled are different in different time periods.

Following Devereux (2007), if we take expectations of *equation (1)* conditional on cohort and year, we would obtain the following cohort population version of *equation (1)*:

$$y_{ct} = x'_{ct}\beta + \alpha_c + u_{ct} \quad (3)$$

Where $c = 1, \dots, C$ is used to denote the cohorts and t is used to denote time. y_{ct} is the population mean of y and x_{ct} is the population mean of x . Assuming that the population in each cohort is fixed over time, α_c would be constant over time. Therefore, it is possible to replace α_c by cohort dummies. As Devereux (2007) points out, the small sample bias of the efficient Wald estimator could now be interpreted as a measurement error problem. In other words, \bar{y}_{ct} and \bar{x}_{ct} are error ridden measures of y_{ct} and x_{ct} .

In order to define the estimator, Devereux (2007) and Deaton (1985) stack the cohorts into an index, g ²⁸, for cohort-time groups. The cohort dummies are absorbed into the x matrix, such that *equation (2)* becomes:

$$\bar{y}_g = \bar{x}'_g\beta + \bar{u}_g \quad (4)$$

Following Devereux (2007), we assume that the sampling error has the following structure:

$$\begin{pmatrix} \bar{y}_g - y_g \\ \bar{x}_g - x_g \end{pmatrix} \sim \text{iid} \left(0, \frac{1}{n_g} \begin{bmatrix} \sigma_{00} & \sigma' \\ \sigma & \Sigma \end{bmatrix} \right) \quad (5)$$

²⁷ The use of this estimator is discussed in detail in Angrist (1991).

²⁸ Where g can go from 1 to G

Where y_g is the population mean of y in group g , and x_g is the population mean of x in group g .

Given the sampling error structure, Devereux (2007) defines the efficient Wald estimator as:

$$\beta^{WALD} = (\sum_{g=1}^G n_g \bar{x}_g \bar{x}'_g)^{-1} (\sum_{g=1}^G n_g \bar{x}_g \bar{y}_g) \quad (6)$$

Deaton (1985), cited by Devereux (2007), showed that the efficient Wald estimator is inconsistent when the number of groups (G) is taken to infinity with a fixed number of observations per group.

However, if we assume that G goes to infinity, it is possible to consistently estimate β using the following errors in variables estimator suggested by Deaton (1985):

$$\beta^{DEATON} = (\sum_{g=1}^G n_g \bar{x}_g \bar{x}'_g - G\hat{\Sigma})^{-1} (\sum_{g=1}^G n_g \bar{x}_g \bar{y}_g - G\hat{\sigma}) \quad (7)$$

Where $\hat{\Sigma}$ and $\hat{\sigma}$ are sample estimates of the relevant population parameters. That is:

$$\begin{aligned} \hat{\Sigma} &= \frac{1}{G} \sum_{g=1}^G \hat{\Sigma}_g \\ \hat{\Sigma}_g &= \frac{1}{n_g - 1} \sum_{i=1}^{n_g} (x_{gi} - \bar{x}_g)(x_{gi} - \bar{x}_g)' \\ \hat{\sigma} &= \frac{1}{G} \sum_{g=1}^G \hat{\sigma}_g \\ \hat{\sigma}_g &= \frac{1}{n_g - 1} \sum_{i=1}^{n_g} (x_{gi} - \bar{x}_g)(y_{gi} - \bar{y}_g)' \end{aligned}$$

However, if we don't have enough observations per group to satisfy the asymptotic properties of Deaton's estimator we could potentially obtain biased estimations because of sampling error. In order to account for that possibility, Devereux (2007) proposed an alternative errors in variables estimator that is approximately unbiased to order $\frac{1}{N}$:

$$\beta^{DEVEREUX} = (\sum_{g=1}^G n_g \bar{x}_g \bar{x}'_g - (G - K - 1)\hat{\Sigma})^{-1} (\sum_{g=1}^G n_g \bar{x}_g \bar{y}_g - (G - K - 1)\hat{\sigma}) \quad (8)$$

Where K is the number of dependent variables and $\sum_{g=1}^G n_g = N$.

In the next section, we will apply the errors in variables estimators introduced in this section to estimate the effect of civil conflict on educational achievement at schools.

7. Results

7.1 Ignoring Measurement Error Problem

As pointed out in *Section 6*, the literature has proposed several errors in variables estimators to take into account the measurement error problem that might arise from aggregating data in Pseudo Panel analysis. Nevertheless, some researchers have used estimation methods designed for genuine panel data (standard within estimator) to estimate pseudo panels assuming that the cohorts are sufficiently large so that it is not necessary to correct for measurement errors²⁹. Given this practice of applied researchers, we also decided to estimate our parameters of interest using methods that do not correct for the pseudo panel structure in order to compare the results obtained using the two estimation approaches, i.e. genuine panel data estimation vs. pseudo panel estimation.

In this section, fixed-effect and random-effect models are estimated to study the relationship between civil conflict and the performance of schools in the Colombian High School exit examination. The Hausman test is used to identify which model is appropriate for this particular case (Yaffee, 2003).

Fixed Effects Model and Random Effects Model

A fixed effect model assumes differences in intercepts across groups or time periods. According to Yaffee (2003), in this model the slope is constant but intercepts differ according to the cross-section, in this case, the country. Although there are no significant temporal effects, there are significant differences among schools in this type of model. While the intercept in this case differs from school to school, it may or may not differ over time. According to Torres (2007), a fixed-effect model assumes that something within the school may impact or bias the predictor or outcome variables. Thus, it becomes necessary to control for this. The

²⁹ It should be noted that there is no consensus in the literature about what constitutes a sufficiently large cohort size. In fact, Devereux (2007) has shown that there is a possibility of substantial bias if pseudo panels with small cohorts (100 to 200 individuals) are estimated without correcting for data grouping.

model removes the effect of time-invariant characteristics from the predictor variables so that it is possible to assess the net effect of the predictor.

On the other hand, a random-effect model explores differences in error variances. The rationale behind a random-effect model is that, unlike the fixed-effect model, the variation across schools is assumed to be random and uncorrelated with the predictor or independent variables (Torres, 2007). This model should be used if there are reasons to believe that differences across schools have some influence on our measures of conflict. The random outcome is a function of a mean value plus a random error.

We were able to construct an unbalanced panel of 6847 schools using the available information. It should be acknowledged that the analysis of an unbalanced panel might raise a number of concerns. If the panel is unbalanced due to reasons that are not totally random, then it might be problematic. With that concern in mind and to further test the robustness of the results, we also estimated an additional specification that only considers those schools for which we have information for all the years. The results obtained using this balanced panel do not differ significantly from the results obtained using the unbalanced panel. These results are reported in *Appendix 2* of this chapter.

The Hausman specification test (random vs. fixed effects models) was calculated and reported for each specification considered in this empirical exercise³⁰. According to the results obtained in the specification test, all the specifications should be estimated using a fixed effects model³¹. *Table 6* presents the results of the fixed effects regression for the intensity of the conflict specification using the unbalanced panel. Section A presents the results using the score in the language exam as dependent variable, and section B presents the results using the score in the mathematics exam as dependent variable. As mentioned in *Section 5* of this chapter, two different measures of intensity were considered: the number of civilian deaths in each municipality and a dummy variable that takes a value of 1 if the municipality recorded an intensity of conflict above the national average. *Column 1* reports the results of the specification that includes the first measure of intensity of conflict (number of civilian deaths) and *Column 2* reports the results of the specification that includes the second measure of intensity of conflict (dummy

³⁰ The objective of the Hausman specification test is to check whether there is a significant correlation between the unobserved school-specific random effects and the regressors. If there is no correlation, then the random effects model is the appropriate model. If there is correlation, the random effects model would be inconsistently estimated and therefore the fixed effects model would be the appropriate model.

³¹ The results for the random effects specification are also available upon request.

variable). The results reported in both columns include all the additional controls discussed in *Section 5*.

Table 6. Fixed Effects Regression (Complete Sample Unbalanced Panel) – Intensity of Conflict				
	Section A. Dependent variable: language		Section B. Dependent variable: mathematics	
	(1)	(2)	(1)	(2)
Intensity of conflict (# of civilian deaths)	0.003 (0.010)		-0.029*** (0.002)	
Municipality with intensity of conflict above national average (dummy = 1 if municipality's intensity is above average)		0.008 (0.060)		-0.369*** (0.080)
Age	-0.191*** (0.023)	-0.223*** (0.016)	-0.206*** (0.032)	-0.268*** (0.021)
Gender	0.175 (0.205)	0.306** (0.140)	0.200 (0.287)	0.483** (0.188)
Number of teachers per student	0.003 (0.002)	0.006*** (0.002)	0.068*** (0.003)	0.106*** (0.002)
Cost of Tuition – Mid level	0.101 (0.076)	0.065 (0.064)	0.687*** (0.107)	0.559*** (0.085)
Cost of Tuition – High level	0.199 (0.145)	0.223* (0.127)	0.825*** (0.203)	1.127*** (0.170)
Total expenditure of the municipality in education (log)	0.327*** (0.076)	0.030 (0.047)	1.761*** (0.104)	0.581*** (0.063)
R-squared (within)	0.1164	0.1396	0.2841	0.2276
R-squared (between)	0.1714	0.949	0.0874	0.0666
R-squared (overall)	0.1490	0.0928	0.1084	0.1087
Hausman Test (Prob>chi2)	1170.25 (0.000)	1891.41 (0.000)	794.56 (0.000)	4515.44 (0.000)
Number of groups	6487	6487	6487	6487

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

The results for mathematics show a negative relation between the intensity of the conflict (both measures) and the level of performance in the examination. The coefficient on the variable that uses the number of civilian deaths to capture the intensity of the conflict is -0.029, while the coefficient on the variable that uses the dummy variable to capture municipalities with levels of conflict above average is -0.369. The size of these coefficients is particularly small. Even though we find significant results with the expected negative coefficient sign, the magnitude of the effect is much smaller than expected. On the other hand, the results obtained in the specification that considers the relationship between the intensity of conflict (both specifications) and the performance in the language examination do not show a

significant relationship between these variables. The results obtained using the unbalanced panel are very similar to those obtained using the balanced panel that only considers those schools for which we have information for all the years³².

Table 7. Fixed Effects Regression (Complete Sample Unbalanced Panel) – Presence of Conflict				
	Section A. Dependent variable: language		Section B. Dependent variable: mathematics	
	(1)	(2)	(1)	(2)
Presence of conflict (# of events)	0.002 (0.002)		0.001 (0.001)	
Municipality with presence of conflict above national average (dummy = 1 if municipality's presence is above average)		-0.188*** (0.045)		-0.231*** (0.058)
Age	-0.199*** (0.019)	-0.217*** (0.016)	-0.194*** (0.025)	-0.234*** (0.021)
Gender	0.062 (0.165)	0.314** (0.140)	0.146 (0.219)	0.416** (0.181)
Number of teachers per student	0.004 (0.002)	0.004 (0.002)	0.072*** (0.003)	0.086*** (0.002)
Cost of Tuition – Mid level	0.082 (0.069)	0.086 (0.064)	0.451*** (0.092)	0.397*** (0.083)
Cost of Tuition – High level	0.233* (0.136)	0.212* (0.126)	1.005*** (0.181)	1.012*** (0.164)
Total expenditure of the municipality in education (log)	0.062 (0.061)	0.019 (0.048)	0.592*** (0.080)	0.118* (0.063)
R-squared (within)	0.1239	0.1413	0.2980	0.2823
R-squared (between)	0.1329	0.0856	0.0785	0.0756
R-squared (overall)	0.1143	0.0845	0.1171	0.1252
Hausman Test (Prob>chi2)	1562.44 (0.000)	1898.49 (0.000)	722.68 (0.000)	5347.53 (0.000)
Number of groups	6487	6487	6487	6487

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

The results for the specification that considers our second measure of presence (dummy variable that takes a value of 1 if the municipality recorded a number of conflict events above the national average) show a negative relation between the presence of conflict and the levels of performance in the mathematics and

³² The results obtained for the balanced panel that only considers those schools for which we have information for all the years are reported in the appendix.

language examinations. On the other hand, the results obtained in the specifications that consider the relationship between our first measure of the presence of conflict (number of conflict events) and the performance in the mathematics and language examinations show no significant relationship. Just like in the case of the specifications that considered the intensity of the conflict measures, the results obtained using the unbalanced panel are very similar to those obtained using the balanced panel that only considers those schools for which we have information for all the years.

7.2 Pseudo Panel Data Estimation

Deaton (1985), cited by Verbeek (2008), has shown that it is possible to use cohorts to obtain consistent estimators when repeated cross-sections are available, even if α_i is correlated with one or more of the explanatory variables. Following the methodology explained in *Section 6*, in this section we are going to apply the errors in variables estimators proposed by Angrist (1991) and Devereux (2007) to estimate the effect of civil conflict on educational performance in schools.

Table 8 presents the results of the pseudo panel regressions using the errors in variables estimators for the specifications that use our first measure of intensity of conflict (number of civilian deaths) as variable of interest. Section A presents the results using the score in the language exam as dependent variable, and section B presents the results using the score in the mathematics exam as dependent variable. Each of the columns in these tables present the results for the different errors in variables estimators considered. *Column 1* reports the results of the specification that was estimated using an EWALD estimator (Angrist 1991) and *Column 2* reports the results of the specification that was estimated using the UEVE estimator (Devereux 2007). The results reported in both sections include all the additional controls discussed in *Section 5*. The results show a negative and significant relation between the intensity of the conflict and the levels of performance in the mathematics and language examinations. The coefficient for the variable that captured the intensity of the conflict obtained using the two estimators that were considered for this econometric exercise is very similar, if not identical. In the case of the mathematics examination, the pseudo panel regressions show that the coefficient on the variable that uses the number of civilian deaths to capture the intensity of the conflict is -0.029 for the two estimators considered. It is worth noting that the results obtained using the errors in variables estimators are identical to the results obtained using the genuine panel data estimator. On the other hand, the results for the language examination specification show that the coefficient of interest is negative and significant, ranging from -0.014 (EWALD) to -

0.015 (UEVE). It is important to highlight that in stark contrast to the results obtained using the errors in variables estimators, the results obtained using the genuine panel data estimator did not show significant results.

Table 9 reports the results of the pseudo panel regressions using the errors in variables estimators for the specifications that use our second measure of intensity of conflict (dummy variable). The results are presented in the same format as *Table 9*.

The results for the specification that considers our second measure of the intensity of conflict (dummy variable that takes a value of 1 if the municipality recorded civilian deaths above the national average) show a negative and significant relation between the intensity of conflict and the levels of performance in the mathematics and language examinations. The pseudo panel regressions results for the specification that uses the score obtained in the mathematics examination as the dependent variable, show that the coefficient on the variable that captures our second measure of intensity of conflict (dummy variable) ranges from -0.783 (EWALD) to -0.778 (UEVE), depending on the errors in variables estimator. In this case our preferred estimator is the UEVE estimator given the less restrictive assumptions that need to be satisfied in order for it to be unbiased³³. On the other hand, in the case of the language examination the results show that the coefficient of interest is also negative and significant, ranging from -0.393 (EWALD) to -0.382 (UEVE). It should be noted that the results obtained using the two errors in variables estimators differ significantly from the results obtained using the genuine panel estimator. Such differences suggest that the results obtained using the genuine panel estimator are likely biased. As we explained in *Section 6*, the errors in variables estimators that are employed in this section correct for errors of measurement in the cohort means. The use of these estimators is also particularly important because they account for potential correlation between the errors of measurement in the dependent and the explanatory variables.

³³ Refer to *Section 6* of this chapter for more details about the errors in variables estimators.

Table 8. Intensity of Conflict Pseudo Panel Regression – Errors in Variables Estimators				
	Section A. Dependent variable: language		Section B. Dependent variable: mathematics	
	(1) EWALD Estimator	(2) UEVE Estimator	(1) EWALD Estimator	(2) UEVE Estimator
Intensity of conflict (# of civilian deaths)	-0.014*** (0.005)	-0.015*** (0.005)	-0.029*** (0.005)	-0.029*** (0.005)
Age	-0.735*** (0.026)	-0.759*** (0.028)	-0.294*** (0.025)	-0.303*** (0.026)
Percentage of students who study during nighttime	-0.616*** (0.154)	-0.679*** (0.157)	-1.365*** (0.144)	-1.339*** (0.147)
Gender	-1.134*** (0.138)	-1.170*** (0.142)	0.989*** (0.129)	0.994*** (0.133)
Number of teachers per student	0.028*** (0.002)	0.028*** (0.002)	0.013*** (0.002)	0.013*** (0.002)
Cost of Tuition – Mid level	2.878*** (0.105)	2.865*** (0.105)	1.520*** (0.098)	1.516*** (0.098)
Cost of Tuition – High level	7.934*** (0.140)	7.932*** (0.140)	6.365*** (0.131)	6.365*** (0.131)
Total expenditure of municipality in education (log)	0.266*** (0.037)	0.264*** (0.037)	0.016 (0.034)	0.016 (0.035)

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

Even though our focus in this econometrical exercise is to understand the relationship between the intensity of the conflict and the performance in standardized testing, it is worth discussing briefly the results obtained for the other control variables included in the analysis. Most of the control variables, including the variables controlling for the school meeting time, gender, age, number of teachers per student, expenditure of the municipality in education and tuition, were significant in all the specifications considered. According to the findings of the model, there is a negative relationship between being enrolled in a nighttime school and the performance in the exit examination. The cost of tuition has a significant relationship with the performance in the mathematics and language examination. It should be highlighted that the variable that captures the cost of tuition is a categorical variable that was divided into separate dummy variables to make it suitable for the econometric exercise³⁴. The tuition category to which the other categories should be compared to is low tuition (i.e. reference category). The schools with a mid-level tuition have higher levels of performance in the examination compared to those schools with low tuition. Such finding would indicate that those schools with the lowest tuition, or those offering a free

³⁴ These categories are discussed in detail in *Section 5* of this chapter.

education, are the ones that are comparatively doing worst in the examination. The variable that captures the gender composition is significant under all the specifications considered. However, the sign of the coefficient changes depending on the dependent variable that is being considered. In the case of the language examination the sign of the coefficient is negative while and for the mathematics examination the sign of the coefficient is positive. These results indicate that those schools with a higher male composition tend to perform relatively better in the mathematics examination, and relatively worst in the language examination. The variable that captures the total expenditure of the municipality in education is only significant for the language examination specifications. The coefficients obtained using the two errors-in-variables estimators show a positive and significant relation between expenditure in education and performance in the language examination.

Table 9. Intensity of Conflict Pseudo Panel Regression – Errors in Variables Estimators				
	Section A. Dependent variable: language		Section B. Dependent variable: mathematics	
	(1) EWALD Estimator	(2) UEVE Estimator	(1) EWALD Estimator	(2) UEVE Estimator
Municipality with intensity of conflict above national average (dummy = 1 if municipality's intensity is above average)	-0.393** (0.171)	-0.382** (0.172)	-0.783*** (0.153)	-0.778*** (0.155)
Age	-0.756*** (0.024)	-0.784*** (0.025)	-0.293*** (0.022)	-0.305*** (0.023)
Percentage of students who study during nighttime	-0.342*** (0.155)	-0.417*** (0.160)	-1.360*** (0.139)	-1.319*** (0.144)
Gender	-1.234*** (0.126)	-1.275*** (0.130)	0.968*** (0.113)	0.973*** (0.117)
Number of teachers per student	0.028*** (0.002)	0.028*** (0.002)	0.013*** (0.001)	0.013*** (0.001)
Cost of Tuition – Mid level	2.970*** (0.095)	2.950*** (0.096)	1.576*** (0.086)	1.569*** (0.086)
Cost of Tuition – High level	7.927*** (0.135)	7.922*** (0.135)	6.323*** (0.121)	6.323*** (0.121)
Total expenditure of the municipality in education (log)	0.226*** (0.029)	0.225*** (0.029)	0.048* (0.026)	0.047* (0.026)
Number of groups	6442	6442	6442	6442

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

Table 10 reports the results of the pseudo panel regressions using the errors in variables estimators for the specifications that use our first measure of presence of conflict (number of conflict events) as variable of interest. Section A presents the results using the score in the language exam as dependent variable, and section B presents the results using the score in the mathematics exam as dependent variable. Each of the columns in these tables present the results for the different errors in variables estimators considered. *Column 1* reports the results of the specification that was estimated using an EWALD estimator (Angrist 1991) and *Column 2* reports the results of the specification that was estimated using the UEVE estimator (Devereux 2007). The results reported in both sections include all the additional controls discussed in Section 5.

The results illustrated in Table 10 show no significant relationship between the presence of conflict (number of conflict events) and the levels of performance in the mathematics and language examinations. This result resembles the result obtained in the initial econometrical exercises that used the genuine panel data estimator.

Table 10. Presence of Conflict Pseudo Panel Regression – Errors in Variables Estimators				
	Section A. Dependent variable: language		Section B. Dependent variable: mathematics	
	(1) EWALD Estimator	(2) UEVE Estimator	(1) EWALD Estimator	(2) UEVE Estimator
Presence of conflict (# of events)	0.018 (0.010)	0.019 (0.010)	-0.010 (0.004)	-0.010 (0.005)
Age	-0.776*** (0.024)	-0.805*** (0.026)	-0.307*** (0.023)	-0.318*** (0.024)
Percentage of students who study during nighttime	-0.770*** (0.146)	-0.851*** (0.149)	-1.377*** (0.135)	-1.342*** (0.139)
Gender	-1.090*** (0.129)	-1.125*** (0.133)	1.071*** (0.120)	1.079*** (0.124)
Number of teachers per student	0.029*** (0.002)	0.029*** (0.002)	0.014*** (0.002)	0.014*** (0.002)
Cost of Tuition – Mid level	2.871*** (0.098)	2.852*** (0.098)	1.581*** (0.090)	1.575*** (0.091)
Cost of Tuition – High level	7.887*** (0.135)	7.878*** (0.135)	6.401*** (0.125)	6.400*** (0.125)
Total expenditure of municipality in education (log)	0.178*** (0.035)	0.175*** (0.035)	0.044 (0.032)	0.044 (0.032)
Number of groups	6442	6442	6442	6442

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

Table 11 reports the results of the pseudo panel regressions using the errors in variables estimators for the specifications that use our second measure of presence of conflict (dummy variable that takes a value of 1 if the municipality recorded a number of conflict events above the national average). The results are presented in the same format as Table 10 and also include all the additional controls discussed in *Section 5*. The results show a negative and significant relation between the presence of conflict and the levels of performance in the mathematics and language examinations. In the case of the mathematics examination, the pseudo panel regressions show that the coefficient on the variable that uses our second measure of presence of conflict (dummy variable) ranges from -0.260 (UEVE) to -0.262 (EWALD) for the two estimators considered. The results obtained using the errors in variables estimators are similar to the result obtained using the genuine panel data estimator. The estimated coefficient on the variable that captures our second measure of presence of conflict calculated using the genuine panel data estimator is -0.260. In line with the results obtained for the mathematics examination, the results for the language examination specification show that the coefficient of interest is negative and significant, ranging from -0.362 (EWALD) to -0.365 (UEVE). It is important to highlight that the results obtained using the two errors in variables estimators differ significantly from the results obtained the genuine panel estimator in the case of the language examination specification. The coefficient that was estimated using the genuine panel data estimator (-0.188) amounts to about half the value of the coefficient estimated using the errors in variables estimators (ranges from -0.362 to -0.365). As we mentioned previously, such differences suggest that the results obtained using the genuine panel estimator are likely biased. It is important to remember that the errors in variables estimators that are employed in this section correct for errors of measurement in the cohort means, and additionally they account for potential correlation between the errors of measurement in the dependent and the explanatory variables.

Several control variables were included as part of the language and mathematics specifications. As it can be seen in Tables 10 and 11, most of the control variables included in the models were significant under the specifications considered. The results show that the time of the day in which classes take place has a positive relationship with the performance in the language and mathematics examinations. Those schools that meet during nighttime have a lower level of performance in both examinations. The cost of tuition also has a significant relationship with the performance in the language and mathematics examinations. The results show that schools with a lower tuition have an inferior performance in the examination compared to those schools with a mid-level and high-level tuition. This indicates that those schools with the lowest tuition, or those offering a free education, are

the ones that are comparatively doing worst in the examination. The variable that captures the gender composition is significant under all the specifications considered. However, the sign of the coefficient changes depending on the dependent variable that is being considered. The positive sign on the gender variable shows that male students are significantly doing better than female students in the mathematics examination. On the other hand, in the case of the language examination the negative sign on the gender variable shows that male students are significantly doing worst. The variable that captures the total expenditure of the municipality in education is only significant for the language examination specifications. The coefficients obtained using the two errors-in-variables estimators show a positive and significant relation between expenditure in education and performance in the language examination.

Table 11. Presence of Conflict Pseudo Panel Regression – Errors in Variables Estimators					
	Section A. Dependent variable: language		Section B. Dependent variable: mathematics		
	(1) EWALD Estimator	(2) UEVE Estimator	(1) EWALD Estimator	(2) UEVE Estimator	
Municipality with presence of conflict above national average (dummy = 1 if municipality's presence is above average)	-0.362*** (0.107)	-0.365*** (0.107)	-0.262** (0.096)	-0.260** (0.097)	
Age	-0.770*** (0.024)	-0.799*** (0.025)	-0.316*** (0.022)	-0.329*** (0.023)	
Percentage of students who study during nighttime	-0.471*** (0.156)	-0.554*** (0.161)	-1.179*** (0.140)	-1.130*** (0.145)	
Gender	-1.239*** (0.126)	-1.280*** (0.130)	0.967*** (0.113)	0.972*** (0.117)	
Number of teachers per student	0.029*** (0.002)	0.028*** (0.002)	0.013*** (0.001)	0.013*** (0.001)	
Cost of Tuition – Mid level	2.968*** (0.095)	2.948*** (0.096)	1.573*** (0.086)	1.566*** (0.086)	
Cost of Tuition – High level	7.943*** (0.134)	7.937*** (0.134)	6.331*** (0.127)	6.330*** (0.121)	
Total expenditure of the municipality in education (log)	0.197*** (0.032)	0.193*** (0.032)	0.012 (0.028)	0.013 (0.029)	
Number of groups	6442	6442	6442	6442	

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

8. Discussion of the Results

The emergence of a civil conflict can affect schools at multiple levels: attacking individuals, communities, and the educational systems (Williams 2004). Indeed, schools are an easy and convenient target for the rebel groups involved in a civil conflict. The results obtained in this chapter provide evidence of the negative repercussions that a civil conflict can have on the achievement of schools. Our results show a negative and significant relation between our two measures of the intensity of the conflict and the levels of performance in the mathematics and language examinations. The results also show a negative and significant relation between our second measure of the presence of conflict (dummy variable) and the levels of performance in the mathematics and language examinations.

Schools in Colombia have been affected in a number of direct and indirect ways by the armed conflict. As O'Malley (2007) highlights, in some cases schools are the only visible symbol of government rule making them easy targets for rebel groups. It is very important to understand the negative effects that can be imposed by the civil conflict so that schools can be prepared and better equipped to overcome these negative consequences. Furthermore, teachers should be trained to deal with these special circumstances. As Thrupp and Lupton (2006) highlight, policymakers should be vigilant so that the constraints imposed by the civil conflict do not become the excuse for low expectations and inequitable schooling provision.

The literature has shown that there is not a simple combination of factors that can produce an effective school (Sammons 1995). In order to analyze school performance, we have to take into consideration several key aspects. Sammons (1995) has highlighted a number of factors for effective schools, including: shared vision and goals, a learning environment, concentration on teaching and learning, purposeful teaching, high expectations, home-school partnership, and learning organization.

Sammons (1995) points out that schools are more effective when there is consensus on the aims and values of the school, and when these objectives are put into practice consistently and in a collaborative way. However, this is particularly difficult to achieve in schools that face special challenges because of external pressures associated to problematic social and political contexts. In the specific case of Colombia the problematic social and political context is associated to the civil conflict. Schools located in conflict zones are particularly affected by the negative context. According to Thrupp and Lupton (2006), the adaptation of teaching and school leadership practices is often necessary in order to deal with the social and political contexts of schools. These adaptations can also be extended

to other aspects of the organization within a school. For example these adjustments can be applied to the length of classes, class sizes, additional learning support, and extra curricular activities (Thrupp and Lupton 2006). Nevertheless, those adaptations are not viable in all cases resulting in lower levels of achievement.

The learning environment is also particularly important for school efficiency. The available research in this topic has shown that schools are more effective under an orderly atmosphere (Sammons 1995). Sammons (1995) highlights that “successful schools are more likely to be calm rather than chaotic places”. Unfortunately, schools in conflict zones are more likely to experience a poor learning atmosphere. This is often the case because of a charged emotional environment. The poor learning environment has serious implications for teachers’ and administrators’ work. Teachers have to develop strategies for dealing with the emotional needs within the classroom of the students affected by conflict. This is problematic, because as Lupton (2004) highlights, “a charged emotional environment can lead to an unpredictable working environment...incidents could erupt at any time, such that neither lessons nor free time could be relied upon to go according to plan”. These unplanned events can potentially have a great incidence on schools’ outcomes (Lupton 2004). It is indeed very difficult to consistently maintain a good environment for learning and a high quality of teaching under a charged emotional environment associated to civil conflict.

The physical environment of a school can also have an effect on achievement. According to Sammons (1995), if the school infrastructure is adequate and it is properly maintained, this can result in higher standards of academic attainment and behavior. According to this researcher, this is the case because attractive and stimulating working conditions tend to improve morale (Sammons 1995). However, as Novelli (1998) has shown, the infrastructure of schools in Colombia, particularly in rural areas, has been directly affected by the civil conflict. According to this researcher, between the year 1996 and the year 2003, “71 schools suffered attacks by guerillas, paramilitaries and state agents, often during combat between the different groups” (Novelli 1998).

Sammons (1995) has shown that positive expectations of pupil achievement amongst teachers, pupils and parents are also a very important feature of effective schools. However, an environment affected by civil conflict can negatively impact this aspect of good schooling. It is particularly difficult for teachers to maintain high expectations when these expectations are frequently disappointed because of the harsh conditions associated to the civil conflict.

Schools are designed as places in which to teach and learn. Nevertheless, as Sammons (1995) highlights, although these activities appear to be obvious,

research suggests that schools can differ greatly in the extent to which they concentrate on teaching. Achievement is clearly dependent upon classroom teaching and the quality of it, and this is determined, at least in part, by the quality of the teachers in the school. However, as Sammons (1995) points out, "high quality teachers do not always perform to their full potential, and teaching styles and strategies are important factors related to pupil progress." In the case of Colombia, the teachers of schools located in conflict zones are possibly not performing at their full potential because of the use or threat of violence in various ways. According to O'Malley (2007), a number of terror tactics have been used against teachers as part of the Colombian civil conflict, including: "telephone and written threats, harassment at home by armed personnel, public declarations or written messages singling them out as military targets, demands for their resignation from their job with an ultimatum to leave the work place, extortion by illegal armed groups against teachers on the highest salary scales, murdering teachers in front of their families and students, and forced disappearances and kidnappings". Illegal armed groups in Colombia have targeted school teachers because of their leadership role in the defense of human rights on behalf of the community. O'Malley (2007) highlights that teachers, who are often times the most educated members of Colombian towns which have been affected by the civil confrontation, have become in many occasions the spokespeople on issues affecting the communities where they live. According to Thrupp and Lupton (2006), "attempts by teachers and managers to respond to their local context often lead to trade-offs between equally valuable activities". In other words, dealing with problems associated with the conflict environment may reduce the time that teachers can dedicate to adequately prepare classes or the time to plan new educational projects. Thrupp and Lupton (2006) further highlight that "disadvantaged contexts generate additional time implications, both for mainstream teaching staff and in particular for senior staff."

Sammons (1995) has highlighted that effective schools are "learning organizations, with teachers and senior managers continuing to be learners, keeping up to date with their subjects and with advances in understanding about effective practice". However, this role of schools as learning organizations can be lost during a civil conflict. Williams (2004) has pointed out that during a civil conflict the capacity of the education authorities is weakened, and organizational development slows or stops. In that sense, to a greater or lesser extent a civil conflict can interfere with the curriculum development at schools. This can have a direct effect on the levels of achievement of those schools directly affected.

The negative and significant effect that was found is of relevance for the long term development of human capital, and therefore, it should be closely examined by the policy makers in charge of educational policy. However, it should be

acknowledged that the magnitude of the effect that was found is not as substantial as expected. There are some factors, which are unique to the Colombian context that could explain why the magnitude of the effect is not as large as expected. The first factor is associated to the high levels of resilience amongst teachers, school officials, students and parents that can be developed when confronted with harsh conditions over a long period of time. As we explained in *Chapter 2* of this dissertation the Colombian civil conflict is a long duration and low intensity conflict that dates back to the times of *La Violencia* in the 1940s. In fact, the Colombian conflict is considered by some researchers as one of the world's longest running civil conflicts (Restrepo et al. 2004). The different schools considered in this study have been exposed to the civil conflict to different extents, directly or indirectly, for many years. It is possible that those schools located in municipalities with a high intensity and/or a high presence of conflict have developed strategies to cope with this. Such efforts could translate in levels of performance that are similar to those of schools in non-conflict zones. A second factor is associated to the supportive relation between families and schools. Research has shown that supportive relations and co-operation between families and schools can also have positive effects on levels of achievement (Summons 1995). In fact, Summons (1995) highlights that we should put special attention to "the benefits of schools fostering parents' involvement in their children's learning." Those families that decide to stay in conflict zones are interested in the safety of their kids, and to do so, they take as many actions to keep their kids away from the conflict and from conflict groups. One of those actions will likely be to follow up as close as possible the educational process of their sons and/or daughters. Such effort could translate in levels of performance of schools in conflict zones that are similar to those of schools in non-conflict zones.

9. Concluding Remarks

The purpose of this chapter is to explore the relationship between civil conflict and educational achievement in Colombia through the use of pseudo-panel analysis. In Colombia there are no panel datasets available to study the educational achievement of students enrolled in the formal educational system, but there is a very complete series of independent cross-sections managed by the Colombian Institute for the Evaluation of Education (ICFES). Due to the lack of longitudinal data, in this chapter we constructed a pseudo-panel of schools from the Saber11 examination cross-sectional data sets. This chapter contributes to the existing literature that studies the relationship between civil conflict and education, including the contributions of Swee (2008), Akresh and de Walque (2008), Merrouche (2006), Lai and Thyne (2007), Shemyakina (2011), Debalen and Paul (2012), Bundervoet (2012), and Bellows and Miguel (2006). However, this chapter

differs significantly from previous efforts because we construct a pseudo panel of schools in order to get a better understanding of the relationship between civil conflict and education. In fact, to the best of my knowledge this is the first effort to try to understand the relationship between conflict and achievement in standardized exams through the use of a pseudo panel of schools.

The literature has proposed several errors in variables estimators to take into account the measurement error problem that might arise from aggregating data in Pseudo Panel analysis. However, some researchers have used estimation methods designed for genuine panel data (standard within estimator) to estimate pseudo panels assuming that the cohorts are sufficiently large so that it is not necessary to correct for measurement errors. Given this practice of applied researchers, we decided to estimate our parameters of interest using the standard within estimator and using errors in variables estimators in order to compare the results obtained using the two estimation approaches, i.e. genuine panel data estimation vs. pseudo panel estimation.

The results obtained in this chapter provide evidence of the negative repercussions that a civil conflict can have on achievement of schools. Our results show a negative and significant relation between our two measures of the intensity of the conflict and the levels of performance in the mathematics and language examinations at the school level. The results also show a negative and significant relation between our second measure of the presence of conflict (dummy variable) and the levels of performance in the mathematics and language examinations at the school level. However, our results using the errors in variables estimators and the genuine panel data estimator show no significant relationship between the first measure of the presence of conflict (number of conflict events) and the levels of performance in the mathematics and language examinations at the school level.

Several additional control variables were also included as part of the language and mathematics specifications. Most of these variables, including the variables controlling for the school meeting time, gender, age, number of teachers per student, expenditure of the municipality in education and tuition, were significant in all the specifications considered. According to our findings, there is a negative relationship between being enrolled in a nighttime school and the performance in the exit examination at the school level. The results also indicate that the cost of tuition has a significant relationship with the performance in the mathematics and language examination at the school level. We also found that the variable that captures the gender composition is significant under all the specifications considered. However, the sign of the coefficient changes depending on the dependent variable that is being considered.

There are a number of factors that could be taken into consideration to explain the results obtained in the econometric exercise. It is particularly important to establish the factors through which the intensity and presence of the conflict could affect the performance of schools in exit examinations, and in doing so, help us get a better understanding of the results obtained in this chapter. Following Lupton (2004), Williams (2004), O'Malley (2007), Sammons (1995) and Thrupp and Lupton (2006) we discussed some of these factors in this chapter.

The negative and significant effect that was found is of relevance for the long term development of human capital, and therefore, it should be closely examined by the policy makers in charge of educational policy. However, it should be acknowledged that the magnitude of the effect that was found is not as substantial as expected. In this chapter we argued that there might be some factors, which are unique to the Colombian context that could explain why the magnitude of the effect is not as large as expected.

The evidence obtained in this chapter can be used to guide the formulation of educational policy. Special efforts should be made at conflict zones to strengthen the capacity of the education authorities, and to promote organizational development at schools. Additionally, policy makers should design policies to develop the different factors that are conducive to effective schools (i.e. shared vision and goals, a learning environment, concentration on teaching and learning, purposeful teaching, high expectations, home-school partnership, and learning organization). We need to design effective policies to guarantee that those schools in conflict zones have the adequate incentives and resources to continue developing these factors. By doing so we can help improve the levels of achievement of schools in the areas of the country particularly affected by the civil conflict.

Appendix 1. Descriptive Statistics

Table 12. Descriptive Statistics				
Variable	2000		2002	
	Mean	S.D.	Mean	S.D.
Male students (dummy; male = 1, female = 0)	0.46	0.23	0.47	0.24
Age	18.22	2.08	18.48	3.04
Total teachers per student	0.37	0.38	0.33	0.49
Low tuition	0.79	0.36	0.74	0.40
Mid tuition	0.12	0.27	0.13	0.29
High tuition	0.09	0.25	0.13	0.30
Intensity of the conflict (# of deaths)	11.09	16.47	28.08	27.61
Presence of conflict (# of conflict events)	14.60	15.27	36.36	44.79
Number of Observations (schools)	4916		6316	

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

Table 13. Descriptive Statistics				
Variable	2005		2007	
	Mean	S.D.	Mean	S.D.
Male students (dummy; male = 1, female = 0)	0.48	0.21	0.47	0.20
Age	17.70	1.93	17.78	2.19
Total teachers per student	0.19	0.48	0.21	0.64
Low tuition	0.87	0.33	0.88	0.33
Mid tuition	0.10	0.30	0.09	0.29
High tuition	0.03	0.16	0.03	0.17
Intensity of the conflict (# of deaths)	15.25	13.90	6.03	6.73
Presence of conflict (# of conflict events)	26.83	27.46	26.63	26.97
Number of Observations (schools)				

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

Appendix 2. Results

Table 14. Fixed Effects Regression (Restricted Sample Balanced Panel) – Intensity of Conflict				
	Section A. Dependent variable: language		Section B. Dependent variable: mathematics	
	(1)	(2)	(1)	(2)
Intensity of conflict (# of deaths)	0.004 (0.010)		-0.030*** (0.002)	
Municipality with intensity of conflict above national average (dummy = 1 if municipality's intensity is above average)		-0.012 (0.065)		-0.410*** (0.091)
Age	-0.274*** (0.030)	-0.324*** (0.020)	-0.232*** (0.044)	-0.358*** (0.028)
Gender	0.162 (0.240)	0.383** (0.161)	-0.003 (0.352)	0.398* (0.225)
Number of teachers per student	0.002 (0.002)	0.005*** (0.002)	0.066*** (0.003)	0.104*** (0.003)
Total expenditure of the municipality in education (log)	0.356*** (0.077)	0.090* (0.049)	1.854*** (0.111)	0.630*** (0.069)
R-squared (within)	0.1374	0.1656	0.3128	0.2524
R-squared (between)	0.2590	0.2167	0.1157	0.1018
R-squared (overall)	0.1844	0.1503	0.1258	0.1358
Hausman Test (Prob>chi2)	952.33 (0.000)	1562.64 (0.000)	258.73 (0.000)	2951.67 (0.000)
Number of groups	3573	3573	3573	3573

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

Table 15. Fixed Effects Regression (Restricted Sample Balanced Panel) – Presence of Conflict				
	Section A. Dependent variable: language		Section B. Dependent variable: mathematics	
	(1)	(2)	(1)	(2)
Presence of conflict (# of events)	0.002 (0.002)		0.0003 (0.001)	
Municipality with presence of conflict above national average (dummy = 1 if municipality's presence is above average)		-0.170*** (0.049)		-0.254*** (0.065)
Age	-0.326*** (0.025)	-0.317*** (0.020)	-0.265*** (0.035)	-0.311*** (0.027)
Gender	0.072 (0.194)	0.398*** (0.161)	-0.004 (0.269)	0.360* (0.216)
Number of teachers per student	0.003 (0.002)	0.003 (0.002)	0.069*** (0.003)	0.084*** (0.003)
Total expenditure of the municipality in education (log)	0.112* (0.062)	0.044 (0.051)	0.639*** (0.085)	0.144** (0.068)
R-squared (within)	0.1495	0.1664	0.3313	0.3106
R-squared (between)	0.2549	0.2024	0.1104	0.1127
R-squared (overall)	0.1743	0.1378	0.1437	0.1568
Hausman Test (Prob>chi2)	1218.30 (0.000)	1580.99 (0.000)	256.28 (0.000)	3920.17 (0.000)
Number of groups	3573	3573	3573	3573

Source: author's calculations using the databases from ICFES, CERAC, and DANE.

WORKS CITED

- (CERAC), Centro de Recursos para el Análisis del Conflicto (2013). *Conflict data*.
- (DANE), Departamento Administrativo Nacional de Estadísticas (2010). Ficha Metodológica de la Investigación de Educación Formal. Vol. 6, pp. 10. Bogotá: Departamento Administrativo Nacional de Estadística (DANE).
- (ICFES), Instituto Colombiano para la Evaluación de la Educación (2013). FTP Databases. from Instituto Colombiano para la Evaluación de la Educación (ICFES)
- Akresh, R., and de Walque, D. (2008). Armed conflict and schooling: evidence from the 1994 Rwandan genocide.
- Angrist, J., 1991, Grouped-data estimation and testing in simple labor supply models, *Journal of Econometrics* 47, 243-266.
- Barrera, F., and Ibáñez, A. M. (2004). Does Violence Reduce Investment in Education?: A Theoretical and Empirical Approach: UNIVERSIDAD DE LOS ANDES-CEDE.
- Bellows, J., and Miguel, E. (2006). War and institutions: New evidence from Sierra Leone. *The American economic review*, 96(2), 394-399.
- Benefield, P. a. T., K. (2005). Education and Conflict: Research and Research Possibilities: National Foundation for Educational Research (NFER).
- Bleakley, H., and Hong, S. C. (2013). When the Race between Education and Technology Goes Backwards: The Postbellum Decline of White School Attendance in the Southern US. *Unpublished manuscript, University of Chicago. April*.
- Braga, M., Checchi, D., and Meschi, E. (2013). Educational policies in a long-run perspective. *Economic Policy*, 28(73), 45-100.
- Brück, T., Justino, Patricia, Verwimp Philip and Avdeenko, Alexandra. (2010). Identifying Conflict and Violence in Micro-Level Surveys. *IZA Discussion Paper*, (No. 5067). <http://ssrn.com/abstract=1648352>

- Brunello, G., and Comi, S. (2004). Education and earnings growth: evidence from 11 European countries. *Economics of education review*, 23(1), 75-83.
- Bundervoet, T. (2012). War, Health, and Educational Attainment: A Panel of Children during Burundi's Civil War: Households in Conflict Network.
- Camacho, A. (2008). Stress and birth weight: evidence from terrorist attacks. *The American Economic Review*, 98(2), 511-515.
- Debalen, A., and Paul, S. (2012). Estimating the causal effects of conflict on education in Cote d'Ivoire. *World Bank Policy Research Working Paper*(6077).
- Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of econometrics*, 30(1), 109-126.
- DeShazo, P., Primiani, T., and McLean, P. (2007). *Back from the brink: evaluating progress in Colombia, 1999-2007*: CSIS.
- Devereux, P. (2007), Small Sample Bias in Synthetic Cohort Models of Labor Supply, *Journal of Applied Econometrics*, 22, 839-848.
- Emery, J. (2012). Long-term consequences of natural resource booms for human capital accumulation. *Indus. and Lab. Rel. Rev.*, 65, 708.
- Fearon, J. D., and Laitin, D. (2014). Does Contemporary Armed Conflict Have 'Deep Historical Roots'?. *Available at SSRN 1922249*.
- Gamoran, A., and Long, D. A. (2007). Equality of Educational Opportunity A 40 Year Retrospective *International studies in educational inequality, theory and policy* (pp. 23-47): Springer.
- Häkkinen, I., Kirjavainen, T., and Uusitalo, R. (2003). School resources and student achievement revisited: new evidence from panel data. *Economics of Education Review*, 22(3), 329-335.
- Haveman, R., and Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of economic literature*, 1829-1878.
- Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources*, 14(3), 351-388.

- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 24(3), 1141-1177.
- Harbison, R. W., and Hanushek, E. A. (1992). *Educational performance of the poor: lessons from rural Northeast Brazil*: Oxford University Press.
- Hoeffler, A., and Reynal-Querol, M. (2003). *Measuring the costs of conflict*. Washington, DC: World Bank.
- Imbens, G., and Wooldridge, J. (2007). Linear panel data models. *Lecture Notes*.
- Inoue, A. (2008). Efficient Estimation and Inference in Linear Pseudo-Panel Data Models. *Journal of Econometrics*, 142 (1), 449-466.
- Justino, P. (2010). War and Poverty. *Handbook of the Economics of Peace and Security*.
- Kennedy, P. (2008). *A Guide to Econometrics*. Willey-Blackwell Publishers. 6th edition.
- Lai, B. a. C. T. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, 44(3), 277-292. doi: 10.1177/0022343307076631
- Lupton, R. (2004). Schools in disadvantaged areas: recognising context and raising quality. *LSE STICERD Research Paper No. CASE076*.
- Merrouche, O. (2006). The human capital cost of landmine contamination in Cambodia: Households in Conflict Network.
- Narita, R. (2008). *Teenage Motherhood, Education and Labour Market Outcomes of the Mother: Evidence from Brazilian Data*. Paper presented at the Apresentação de Trabalho/Congresso.
- O'Malley, B. (2007). *Education Under Attack: A global study on targeted political and military violence against education staff, students, teachers, union and government officials, and institutions*. UNESCO.
- Pal, S., and Saha, B. (2014). In Trusts We Trust: Socially Motivated Private Schools in Nepal. Available at SSRN 2395944.
- Restrepo, J. A., Spagat, M., and Vargas, J. F. (2006). El conflicto en Colombia: ¿quién hizo qué a quién? Un enfoque cuantitativo *Nuestra guerra sin nombre: Transformaciones del conflicto en Colombia*: Editorial Norma.

- Restrepo, J., Spagat, M., and Vargas, J. F. (2004). The Dynamics of the Colombian Civil Conflict: A New Data Set.
- Rodriguez, C., and Sanchez, F. (2012). Armed conflict exposure, human capital investments, and child labor: Evidence from Colombia. *Defence and peace economics*, 23(2), 161-184.
- Sammons, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research*. B and MBC Distribution Services, 9 Headlands Business Park, Ringwood, Hants BH24 3PB, England, United Kingdom.
- Sánchez, F., and Díaz, A. M. (2005). *Los efectos del conflicto armado en el desarrollo social colombiano, 1990-2002*: CEDE, Centro de Estudios sobre Desarrollo Económico, Facultad de Economía, Universidad de los Andes.
- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, 95(2), 186-200. doi: 10.1016/j.jdeveco.2010.05.002
- Simone, G. (2013). Render unto primary the things which are primary's: Inherited and fresh learning divides in Italian lower secondary education. *Economics of education review*, 35, 12-23.
- Swee, E. L. (2009). On war and schooling attainment: The case of Bosnia and Herzegovina. *Households in Conflict Network Working Paper*, 57.
- Thrupp, M., and Lupton, R. (2006). Taking school contexts more seriously: The social justice challenge. *British Journal of Educational Studies*, 54(3), 308-328.
- Todd, P. E., and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement*. *The Economic Journal*, 113(485), F3-F33.
- Torres, O. (2007). Panel Data Analysis: Fixed and Random Effects. Data and Statistical Services, Princeton University. Available at: <http://www.princeton.edu/~otorres/Panel101.pdf>
- Verbeek, M., and Nijman, T. (1992). Can cohort data be treated as genuine panel data? *Empirical Economics*, 17(1), 9-23.

- Verbeek, M. (2008). Pseudo-panels and repeated cross-sections *The Econometrics of Panel Data* (pp. 369-383): Springer Berlin Heidelberg.
- Warunsiri, S., and McNown, R. (2010). The returns to education in Thailand: A pseudo-panel approach. *World Development*, 38(11), 1616-1625.
- Wharton, K. and Oyelere, R. (2012). *Conflict and Its Impact on Educational Accumulation and Enrollment in Colombia: What We Can Learn from Recent IDPs*. IZA Discussion Paper. IZA. Retrieved from <http://ssrn.com/abstract=1922011>
- Williams, J. H. (2004). Civil conflict, education, and the work of schools: Twelve propositions. *Conflict Resolution Quarterly*, 21(4), 471-481.
- Wößmann, L. (2003). Specifying human capital. *Journal of economic surveys*, 17(3), 239-270.
- Yaffee, R. (2003). A Primer for Panel Data Analysis. Connect - Information Technology at New York University. Available at: http://www.nyu.edu/its/pubs/connect/fall03/pdfs/yaffee_primer.pdf

FORCED DISPLACEMENT AND EDUCATIONAL ACHIEVEMENT IN BOGOTA'S PUBLIC SCHOOLS

The Colombian civil conflict has been going on for many decades affecting civilians directly and indirectly. The exclusion of dissident political forces from the *Frente Nacional*, an agreement to share control of the State signed by the two major political parties of the country, contributed to the emergence of guerrilla groups in the mid-1960s. One of the repercussions of the long-term conflict has been the forced displacement of a large number of civilians from conflict areas to non-conflict areas. In fact, according to the 2012 Global Overview Report published in 2013 by the Norwegian Refugee Council, as of the end of 2012 Colombia was the country with the highest number of people displaced by armed conflict (IDPs) in the world, with a total of between 4.9 and 5.5 million (Albuja et al. 2013).

The purpose of this chapter is to identify the effects of forced displacement on the performance of Colombian students in the national high school exit examination. In this chapter we are focusing in those individuals who have been directly affected by conflict given their internal displacement status, instead of looking at the overall effect of conflict like we did in previous chapters. This chapter contributes to the existing literature that studies the impact of civil conflict on education, including Rodríguez and Sánchez (2010), Akresh and de Walque (2008), Lai and Thyne (2007), Shemyakina (2011), Dabalen and Paul (2012), Bundervoet (2012) and Merrouche (2006)³⁵. This chapter also contributes to the literature that studies the impact of internal displacement on the welfare of civilians. Amongst the research that has contributed to this literature we can find the work of Ibañez and Velez (2008), Ibañez and Moya (2010), Ruiz and Vargas (2013), Calderón and Ibañez (2009), Kirschhoff and Ibañez (2002), Kondylis (2008 and 2010), and Fiala (2009). Nevertheless, what is notable is that to the best of my knowledge this is the first effort to try to understand and quantify the impact of forced displacement on education achievement (measured by standardized exams).

The Secretariat of Education of Bogota, a public entity in charge of managing and executing policies and programs in order to guarantee the rights of individuals to access education in the capital city of Colombia, provided us a detailed database for the years 2011 and 2012, containing information about those students who are attending public schools in Bogota and have been internally displaced because of

^{***} The contents of this chapter correspond to the third publishable paper of this PhD dissertation.

³⁵ This literature was reviewed in detail in Chapter 4.

the Colombian civil conflict. The information provided by the Secretariat of Education of Bogota was a key element in the process of identifying those students who have been displaced. Nevertheless, it should be acknowledged that a possible weakness of this dataset is the fact that the information on forced displacement comes from individuals' self-reports.

In this chapter we are going to use two different methodologies to estimate the effects of forced displacement on the performance of Colombian students in the national high school exit examination. Our first approach uses cluster robust inference to estimate the effect of displacement. In cluster robust inference observations are grouped into clusters, with model errors uncorrelated across clusters but correlated within clusters. However, this is our naïve approach in the sense that we are not considering the possibility of endogeneity. Following Calderon and Ibañez (2009), we are going to use an instrumental variable as our second method of analysis in order to deal with the possibility of endogeneity as a source of bias. In fact, the instrumental variables approach considered in sections 11 and 12 of this chapter is our preferred method of estimation.

The rest of this chapter is organized as follows: section 2 presents a brief description of the civil conflict and forced displacement in Colombia; section 3 reviews the academic literature that has studied the impacts of internal displacement; section 4 presents a conceptual background of the forced displacement phenomena; section 5 presents the theoretical framework and identification strategy; section 6 describes the conditions that should be met for an individual to be considered a victim of displacement in Colombia; section 7 presents the dataset that is going to be employed in this empirical exercise; section 8 presents descriptive statistics; section 9 reviews the cluster robust inference methodology; section 10 reports and analyzes the main results obtained using cluster robust inference; section 11 reviews the instrumental variables methodology; section 12 reports and analyses the results obtained using an instrumental variables approach; section 13 presents the discussion of the results and section 14 presents concluding remarks.

2. The Civil Conflict and Forced Displacement in Colombia

The Colombian civil conflict can be traced back to the time of independence³⁶. The political enmity between the two political parties that emerged after independence, the Partido Conservador Colombiano and the Partido Liberal, has marked Colombian history ever since. An important trait of Colombian politics is the depth

³⁶ For more details please refer to Chapter 2.

of the affiliation of the citizens to these two political parties throughout the history of the country (Legrand 2001). The depth of affiliation has led to a high level of hostility between the members of these two political parties, resulting in numerous insurrections and civil wars throughout the 19th and the 20th century. In the 19th century, the most grave of these wars was the Thousand Days War (*Guerra de los Mil Días*), which was fought from 1899 to 1902 causing over 100,000 casualties.

In April 1948, after a period of relative stability of about forty years, the assassination of the Liberal leader and presidential candidate Jorge Eliecer Gaitán led to another major violent outburst now known as the times of the Violence (*La Violencia*). During *La Violencia*, which went from 1946 to 1965, political tensions between the leaders of the Liberal and Conservative parties led to clashes in the countryside between peasants supporting the two sides (Legrand 2001). Armed groups of peasants supporting the Conservative party and supporting the Liberal party were formed during that time. These groups fought each other affecting principally the rural population and the inhabitants of small cities. LeGrand (2001) has highlighted that *La Violencia* is historically one of the civil conflicts in the Western Hemisphere with the greatest number of casualties with over 300,000 people killed. In an effort to ease the confrontation between the supporters of the two parties, the Conservatives and the Liberals came together to form the National Front (Frente Nacional) in 1958. The National Front was a pact between these two political parties, establishing that they would alternate the presidency of the country (Legrand 2001). This meant that although elections were still held everyone knew who would win: a Liberal candidate, then a Conservative candidate, and so on. Additionally, the presidents had to invite opposition figures to hold cabinet positions as part of the pact (Hudson 2010). This agreement, which lasted until 1974, helped to greatly reduce the violence between the supporters of the two parties in the early sixties. However, the exclusion of dissident political forces from the Frente Nacional contributed to the emergence of guerrilla groups in the mid-1960s, who were also looking for participatory power. These new movements were inspired by the Cuban revolution and their successful guerrilla tactics. In 1964, the Fuerzas Armadas Revolucionarias de Colombia (FARC), which later on became the largest guerrilla group in Colombia, was founded, and in 1965, the Ejército de Liberación Nacional (ELN) and the Ejército de Liberación Popular (EPL) began operations.

Although guerrilla war started in Colombia in the 1960s with the emergence of the left wing guerrilla groups, Ibañez and Moya (2010) highlight that this war intensified during the late seventies and early eighties, when these illegal groups got involved in the illegal production of marijuana and cocaine. In fact, these two authors point out that illicit drug trade has provided massive financial resources to rebel groups that have helped sustain the conflict ever since (Ibañez and Moya

2010). The financial resources obtained from these illegal activities were also fundamental in the establishment of the right wing paramilitary groups, which became notorious in the late eighties and beginning of the nineties, and that in many regions of Colombia, challenged the power of the guerrilla groups. The emergence of paramilitary groups, and the strengthening of the guerrilla groups resulting from the drug trafficking business, were key elements in the major intensification of the Colombian conflict in the nineties.

One of the repercussions of civil conflict is the forced displacement of civilians. According to Moya (2012), an important strategy implemented by the actors of the Colombian civil war is to target the civilian population in order to intimidate them, and in doing so, halt collective action, dismantle social networks, and take control over lands and productive assets. Ibañez and Velez (2003) further highlight that the escalation of crimes against civilians is a low cost and effective strategy that allows illegal groups to develop their activities at ease. Faced by these threats, civilians have to decide whether to remain in their hometown living in constant fear because of the risk of losing their lives, or move to another region (Carrillo 2009).

Violent actions against the civil population have indeed resulted in the forced displacement of a considerable number of civilians within the country. Although the internal armed conflict has caused internal displacement in Colombia over the past six decades, the intensification of the conflict in the last two decades has increased the magnitude of this phenomenon. According to the 2012 Global Overview report, published in 2013 by the Norwegian Refugee Council, as of the end of 2012 Colombia was the country with the highest number of people displaced by armed conflict (IDPs) in the world, with a total of between 4.9 and 5.5 million (Albuja et al. 2013). The Norwegian Refugee Council has also reported that in the year 2012 Colombia's internal armed conflict caused the forced displacement of an estimated 230,000 people (Albuja et al. 2013).

3. Literature Review

The available literature that studies conflict and education was reviewed in detail in Chapter 4 of this dissertation. Nevertheless, it is also important to review the literature that studies the effects of forced displacement in order to have a better understanding of the state of the art in this subject. A number of researchers have contributed to this literature by analysing the cases of countries that have been afflicted by a civil conflict. Amongst the academic research that has contributed to this literature we can find the work of Ibañez and Velez (2008), Ibañez and Moya (2010), Ruiz and Vargas (2013), Calderón and Ibañez (2009), Kirschhoff and Ibañez

(2002), Kondylis (2008 and 2010), and Fiala (2009). However, what is notable is that to the best of my knowledge no effort has been made to try to understand and quantify the impact of displacement on education achievement measured by standardized exams. In what follows, I will review the existing research concerning the impacts of forced displacement in order to contextualize the topic under study and justify the relevance of this chapter.

It should be recognized that this literature review focuses on the economic literature that is currently available. Yet, there are other disciplines that have also done scientific research on the relationship between civil conflict, forced displacement and education, including Sociology and Political Science. Nevertheless, the available economic literature that studies this relationship has a stronger quantitative tradition. For our purposes, and considering our interest in quantitative methods, we have decided to center our attention on the economic literature that is currently available.

A number of researchers have studied the impact of displacement on labor market outcomes. Kondylis (2010) uses longitudinal data to analyze the impact of displacement on labor market outcomes for Bosnians in post-war Bosnia and Herzegovina. The effect of displacement is measured using indicators of work, unemployment, inactivity, wages and hours worked. The author finds that displacement negatively impacted the labor market outcomes of Bosnians (men and women), especially in terms of access to employment. Kondylis (2010) shows that displaced men and women are less likely to have a job by about 15 percentage points in relation to individuals that stayed in their hometown. These results imply a significant and large cost of displacement in terms of labor market outcomes. Similarly, Calderon and Ibañez (2009) study the effects of massive inflows of migrants on wages, employment, unemployment, and labor force participation at the destination sites in Colombia. The results obtained by Calderon and Ibañez (2009) suggest that the large inflows of migrants generate an expansion of the informal economy, and such expansion is accompanied by a considerable decline in wages in the sector. These findings imply that not only are the displaced families being affected, but there is also a considerable impact on the most vulnerable groups of society at the destination site. In fact, Calderon and Ibañez (2009) find that informal female workers face the largest impact. Their results show that a 10 percent increase in the participation of displaced individuals in the working age population leads to a decrease in the real wages of about 3 per cent. Particularly, for females working in the informal sector an increment of 10 percent in the participation of displaced individuals in the working age population generates a reduction of about 3.06 percent in real wages (Calderon and Ibañez 2009). The results of Calderon and Ibañez (2009) support the findings obtained by

Kondylis (2010) showing that forced displacement has a negative impact on labor market outcomes.

Other efforts, including those of Ibañez and Moya (2010a) and Fiala (2009), have attempted to explore the relationship between displacement, household assets and consumption patterns. Ibañez and Moya (2010a) study the relationship between forced displacement, household assets and poverty in Colombia. According to these researchers, a conflict-induced shock generates considerable asset losses upon displaced households. Ibañez and Moya (2010a) find that regardless of the extent of asset loss, displaced households are left with an asset base insufficient to escape poverty. The results obtained by these researchers show that displaced households do not catch up even after they have consolidated their settlement at destination sites (Ibañez and Moya 2010a). Ibañez and Moya (2010a) highlight that individuals who had been displaced become trapped in a low level economic trajectory, and in that sense, their return to a high level economic trajectory becomes very unlikely. Obtaining completely different results in comparison to those obtained by Ibañez and Moya (2010a), Fiala (2009) also studies the association between displacement and household assets using a cross-sectional household survey collected in 2004 in Northern Uganda. Exploiting an exogenous variation in migratory movements and using a discontinuity design to control for endogenous factors, Fiala (2009) finds that being displaced is associated with an increase in the value of assets for households that originally had little or no assets and a decrease in the value of assets of all other households. In stark contrast to the results obtained by Ibañez and Moya (2010a), this researcher argues that displacement is not related to asset loss across all asset levels (Fiala 2009).

Other researchers, including Ibañez and Moya (2010b) and Ibañez and Velez (2008), have focused their work on the relationship between displacement and welfare losses. Using data from a survey administered to 2,322 displaced households in Colombia, Ibañez and Moya (2010b) examine welfare changes of displaced households and identify the impact of income generation programs that were established to reduce their vulnerability. Their analysis shows that displaced households are not able to isolate consumption from variations in transitory income, meaning that they are not able to smooth consumption. The results obtained by Ibañez and Moya (2010b) suggest that variations in transitory income turn into variations in consumption. According to these researchers, the harsh conditions in receptor sites, poor labor conditions and a severe disruption of risk-sharing mechanisms are the principal factors limiting the capacity of households to smooth consumption (Ibañez and Moya 2010b). In line with Ibañez and Moya (2010b), Ibañez and Velez (2008) use compensating variation in order to measure the welfare losses associated to forced displacement in Colombia. The compensating variation is the amount of money that displaced households

consider to be necessary for them to be indifferent between their receptor location and their hometown. Through the use of this technique, Ibañez and Velez (2008) find that welfare losses from displacement are substantial. These authors find that the costs of displacement amount to 37 percent of the net present value of aggregated rural consumption. The article also highlights that the economic affectation of displacement is higher for poor households (Ibañez and Velez 2008).

Kondylis (2008) studies the impact of displacement on the output of subsistence agricultural households in post-conflict rural Rwanda. This researcher argues that conflict induced displacement can be interpreted as an exogenous shock, and therefore the individuals that were not displaced can be used as a control group for the analysis. Kondylis (2008) finds that individuals who had been displaced were more likely to resettle in the areas of Rwanda with higher agricultural productivity. This author also finds that the returns to on-farm labor were higher for displaced individuals. Kondylis (2008) explains that these results could be driven by the higher motivation of displaced individuals to improve their economic situation compared to those individuals who were not displaced because of conflict.

A number of researchers have also contributed to the literature by carrying out descriptive analyses of the effects of displacement. For instance, Kirchhoff and Ibañez (2002) use a survey of displaced households conducted in three major receptor cities in Colombia (Bogota, Medellin and Cartagena) to carry out a descriptive analysis of the impact of displacement in this country. These two researchers find descriptive evidence that shows that for many households, displacement causes the loss of agricultural land and the associated way of life (Kirchhoff and Ibañez 2002). Kirchhoff and Ibañez (2002) also find that at the receptor locations, affected households often times have to look for jobs for which they are not trained for. Additionally, the survey shows that a substantial part of the individuals affected by displacement end up being long-term unemployed (Kirchhoff and Ibañez 2002). In line with Kirchhoff and Ibañez (2002), Carrillo (2009) uses descriptive statistics to analyse the humanitarian consequences faced by the internally displaced population in Colombia when they move to an urban environment. According to this researcher, displaced individuals are unable to become self-reliant and meet their basic needs without permanent assistance from the government and humanitarian organizations (Carrillo 2009). Carrillo (2009) highlights that a very important aspect that should be considered in the analysis of Colombian internal displacement is how to rebuild social capital and how to reconstruct the country, when the country is still affected by an armed conflict that still drives hundreds of people away from their hometowns.

Using descriptive statistics, like Carrillo's (2009) and Kirchhoff and Ibañez's (2002) contributions, Ferris and Winthrop (2010) show a general descriptive outlook of the relationship between displacement and access to education. These two researchers highlight that the lack of data on IDPs' access to education makes it very hard, if not impossible, to make comparisons between displaced children and children affected by conflict who have not been displaced (Ferris and Winthrop 2010). Ferris and Winthrop (2010) further highlight that conflicts have a profound impact on family and community life which affects children and their educational process.

It is also possible to find empirical research focusing on the effects of displacement on the health status of affected individuals. Mels et al (2010) examine the impact of war induced displacement on the mental health of Eastern Congolese adolescents. These researchers also compare the mental health status of currently internally displaced adolescents to returnees and non-displaced peers. Using univariate analysis of covariance (ANCOVA), Mels et al (2010) find that the high level of exposure to violence and daily stressors that displaced adolescents have to face causes a high level of psychological distress, when compared to returnees and non-displaced adolescents. These researchers further highlight that the difference in mental status between displaced adolescent who have returned to their hometowns and the ones who have stayed in the receptor locations illustrates that improving the living conditions of war affected adolescents could stimulate resilient outcomes, despite former displacement status (Mels et al 2010). Mels et al (2010) conclude that governments should prioritize the return of displaced children, and the reunification of separated families while enhancing their socio-economic conditions. However, this could only be done under peaceful conditions.

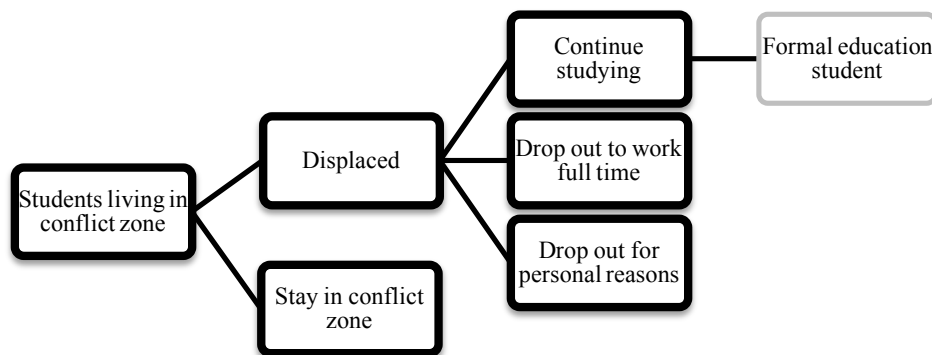
This review of the literature has attempted to capture the state of the art in the academic research of the effects of forced displacement. Nevertheless, those readers who want to complement this review can also refer to the work by Ruiz and Vargas (2013). Ruiz and Vargas (2013) make an important contribution to the literature by reviewing and analysing the available economic research on the impact of displacement. In their paper, Ruiz and Vargas (2013) highlight that the lack of research studying the effects of displacement, explained in large part by the unavailability of data to conduct proper analysis, limits the understanding of this phenomenon. As these two researchers highlight, the economics literature that studies the impacts of forced migration is still in its early stages. In fact, Ruiz and Vargas (2013) point out that the majority of studies on this subject were written in the last five years. Nevertheless, these researchers highlight that the available literature for developing countries has shown that there are serious consequences of forced migration in terms of consumption, employment, and health conditions (Ruiz and Vargas 2013).

As it was reviewed in this section, a number of researchers have contributed to the literature that studies the effects of forced displacement in countries affected by civil conflict. However, there is still a lot of ground to cover in this area of research. There are many topics that are understudied and could potentially play a very important role in policy analysis. The objective of this chapter is precisely to study one of those understudied topics: the relationship between internal displacement and educational achievement measured by the performance in high school exit examinations.

4. Conceptual Background

The families of school-age individuals living in conflict zones face an important decision: to stay in the conflict zone or to migrate to a non-conflict zone. Forced displacement is triggered by the need to find a safe place to live given the violent environment generated by the armed conflict. In fact, Ibañez and Velez (2003) highlight that violence and forced displacement are strongly linked because the terrorist strategies adopted by the illegal armed groups (e.g. massacres, life threats, selected homicides) can potentially trigger forced migration. In that sense, we can think of those families (students) that have been forced to flee from their hometown as involuntary migrants affected by exogenous factors (conflict). Nevertheless, Kirchhoff and Ibañez (2002) point out that civilians living in regions with high levels of violence could also stay in their hometown for two reasons. First, individuals living in conflict areas might have different levels of exposure to the actions of illegal armed groups. Those individuals living in conflict zones that are not exposed directly to the violent events associated to civil conflict, and therefore have a lower risk might decide to stay. Two, civilians analyse the costs behind the displacement decision (e.g. cost of leaving behind important assets). If the costs of migrating are higher than the benefits, individuals might decide to stay in a region with high levels of violence. The two conditions discussed above suggest that households that are directly affected by conflict, and therefore are at higher risk, are the ones that are more likely to opt for displacement (Kirchhoff and Ibañez 2002). Diagram 1 is a graphic representation of the decisions that individuals living in conflict zones have to face.

Diagram 1



Source: author's elaboration

Those students who have been displaced from their hometown often times face harsh conditions in their new place of residency, forcing them to decide whether to continue with their studies or go on to a different activity. More specifically, they have to decide whether to continue studying, to dropout for personal reasons or to find a job to contribute to the income of the household (*diagram 1*). In taking that decision, the individuals weigh the costs associated with a given level of education, the expected benefits from a higher level of education, the expected benefits from dropping out for other personal reasons (e.g. teenage pregnancy), and the expected income derived from finding a full time job at the new place of residency. The decision that each individual takes seeks to maximize their expected utility. It should be noted that the families that have been forced to move because of conflict often times have to abandon their land and other valuable assets. Losing those assets can lead to a reduction in the availability of resources that can be devoted to the education of those family members in school age. Additionally, the kids that belong to those families that have been forced to relocate might be forced to work to help raise the income in the new place of settlement.

Those students affected by the civil conflict that have been forced to move to non-conflict zones, and decide to continue with their formal education, are the ones that attract our attention in this chapter. As mentioned previously, the information provided by the Secretariat of Education of Bogota was a key element in the process of identifying those students who have been displaced because of the civil conflict.

5. Theoretical Framework and Identification Strategy

Consider an education production function that follows the specification proposed by Hanushek (1979) and Todd and Wolpin (2003). Following Todd and Wolpin's (2003) notation, let T_{it} be a measure of achievement for student i who lives in Bogotá in period t . Let's recall that in this chapter we are only going to concentrate in those students who participate in the formal education system of the capital city of Colombia. We are going to assume that knowledge acquisition is a production process in which current and past inputs are combined with the student's genetic endowment to produce a certain level of cognitive achievement (Todd and Wolpin 2003).

Inputs are given by the choices that are being made by parents and schools. Following Todd and Wolpin's (2003) notation, the vector of inputs supplied by the parents at a given age is represented by F_i , and the vector of school-supplied inputs is represented by S_i . Additionally, D_i denotes external factors (in this case the displacement status) affecting student i . The displacement status indicates that a student has been directly affected by the civil conflict, and therefore it is a fundamental factor in the identification of the effect of direct exposure to conflict on educational achievement. The student's endowed ability is denoted as μ_{i0} , and a measurement error in test scores as ε_i . Given those inputs, the production function is represented as:

$$T_i = T(F_i, S_i, D_i, \mu_{i0}, \varepsilon_i) \quad (1)$$

Missing data for past inputs poses an important problem at the time of estimation. In order to overcome that problem, Todd and Wolpin (2003) came up with a specification that only depends on contemporaneous measures of school and family inputs. This specification is often times adopted because of data limitations associated with limited availability of historical input measures or historical test score results. Considering that this specification only depends on contemporaneous observations, it is known as "Contemporaneous Specification". Behind the implementation of a contemporaneous specification we are assuming that ε_{it} is an additive error that includes any omitted factors, including past inputs and endowed capacity, and the possibility of measurement error. Therefore, in equation (1) the residual term includes all the omitted factors. That implies that although this specification has less data requirements, it requires strong assumptions in order to justify its application (Todd and Wolpin 2003).

An important problem that comes up in the specification of an education production function to be estimated empirically is the non-existence of a valid measure of innate abilities. In fact, as Hanushek (1979) highlights, the most consistent and obvious divergence between the empirical specification and the theoretical educational production function is the non-existence of a measure of innate or endowed abilities. However, omitting such an important variable can be problematic because it can bias the estimated regression coefficients. The size of the bias depends on how strong is the effect of the variable on achievement and also on the correlation of the omitted variable with the other variables that are included in the model. Hanushek (1979) points out that if endowed capacity is not correlated with the variables that are included in the model, the estimated coefficients of the model would not be biased. The only effect in such case would be an increase in the residual variance. Nevertheless, it is very unlikely that endowed capacity is not correlated with the other determinants of achievement. This implies that omitting the endowed capacity of the student from the empirical specification will likely cause an upward bias in the estimated impact of the other variables on achievement (Hanushek 1979). The possibility of this upward bias should not be ignored, and therefore it should be considered carefully in the interpretation of the results of the empirical exercise.

The reduced form linear relationship of the educational production function that is going to be estimated is given by:

$$T_i = \beta_0 + \beta_1 D_i + \alpha F_i + \gamma S_i + \varepsilon_i \quad (2)$$

Where D_i is a dummy variable that takes a value of 1 if the student was displaced from his/her hometown, and 0 otherwise. As mentioned before, the displacement status indicates that a student has been directly affected by the civil conflict, and therefore it is a fundamental factor in the identification of the effect of direct exposure to conflict on educational achievement. F_i is a vector of individual, parental and household characteristics that are included as control variables (gender, age, whether the student has taken the High School exit examination before, level of education of the mother and the father, father's occupation, family income, availability of household appliances, household floor material, number of people living in the household). Given the importance of the parents' education and professional background in the educational decisions of children, parental variables are included. A direct measure of household income and indirect (proxy) measures of household income (availability of household appliances) are also included as controls. These variables are included because the available income is an important determinant of the family's ability to invest in the education of the

school-aged household members. S_i is a vector of school characteristics that are also included as control variables (class time, type of school, tuition, number of teachers per student, level of education of the teachers). The inclusion of all these variables is particularly important in order to reduce potential omitted variable bias.

In addition to the controls discussed above, we also included a variable that captures the (lagged) intensity of the conflict in the municipality where the student lived before being displaced is also included. This variable was included to control for the differing levels of violence to which students could have been exposed before being forced to flee (displacement). The variable included in the specification was modeled as an interaction term with displacement status. For our purposes, the intensity of the conflict is measured as the average intensity of conflict in the last five years before the date in which the exam was taken at the municipality where the displaced student used to live before being displaced.

All the regressions are estimated using ordinary least squares (OLS), clustering standard errors at the school level. OLS regression assumes that the residuals are independent. However it is possible that the scores in the high school exit examination within each school may not be independent, and this might lead to residuals that are not independent within schools. In order to overcome that difficulty, we can cluster the observations into schools. In doing so, we can acknowledge that the observations can possibly be correlated within schools, but are independent between them. Section 9 explains in more detail the methodology employed in this empirical exercise.

6. Displacement Status

According to the normative established by the Secretariat of Education of Bogota, for a student studying in the public education system to be considered internally displaced because of conflict, he/she has to be registered in the Unique Registry of the Displaced Population (Registro Único de Población Desplazada -RUPD). This registry, which was created and regulated by the Decree 2569 of the year 2000, is managed by a government agency³⁷ that is in charge of administering the aid packages for the victims of conflict (DPS 2014). The objective of the registry is to have an up-to-date source of information about the displaced population in order to provide the assistance that this group of people needs and to keep track of the aid that has been provided (DPS 2014).

³⁷ Departamento para la Prosperidad Social – Unidad para la Atención y Reparación Integral a las Víctimas.

According to the Law 387 of 1997, any person who has been forced to migrate within the national territory, abandoning their place of residence and their usual activities, because their life, their physical integrity, their security or their personal freedom have been violated or are in danger of being violated, should be considered displaced (DPS 2014). The Colombian law also establishes that every person who has been displaced should testify before a Governmental Office within one year of the events that led to the forced migration. After this initial registration, the public officials in charge of this process evaluate the information that has been provided and within 15 days they determine if the individual/household should be recognized officially as a victim of forced displacement (DPS 2014).

For a student in Bogota to be eligible to receive the special subsidies awarded by law to the victims of forced displacement³⁸, he/she has to provide official proof of their displacement status (RUPD registry). Those students who provide proof of their displacement status are included in a special database administered by the Secretariat of Education of Bogota. As discussed in the next section, the information contained in that database is a major input for this research.

7. Data

The information used to construct the dataset that is used in this empirical exercise comes from two different sources: the Secretariat of Education of Bogota³⁹ and the Colombian Institute for the Evaluation of Education (ICFES). The Secretariat of Education of Bogota provided us detailed information about students who are attending public school in Bogota that have been displaced from their hometown because of the Colombian civil conflict for the years 2011 and 2012. The Secretariat of Education of Bogota is a public entity in charge of managing and executing policies and programs in order to guarantee the rights of individuals to access education in the capital city of Colombia. This Secretariat keeps track of all the students that are enrolled in the public education system in Bogota, paying special attention to those students that are at risk of being excluded from the educational system including the victims of the civil conflict. In addition to the information about the students that had been displaced because of the civil conflict, the Secretariat of Education of Bogota also provided us information about the teachers who work in each school and the level of education of these teachers.

³⁸ Displaced students can receive the following benefits from the Secretariat of Education of Bogota: tuition waiver, registration in the public school closest to the place of residence, and registration for all school-aged family members in the same school.

³⁹ Bogota is the capital city of Colombia.

Databases containing the nation-wide results of the Colombian high school exit examination (Saber 11), as well as information about the socioeconomic features of the students who took the exam, which were made public by the Colombian Institute for the Evaluation of Education (ICFES) for research purposes, are also employed in this chapter. Detailed information about this dataset can be found in Chapter 7 of this dissertation. It should be noted that in this chapter we are only going to use the scores obtained by the students in the mathematics and language examinations as proxies of educational achievement. The results in these two examinations are particularly important considering that they are designed to measure quantitative and verbal skills. Taking into account that we only have information about displaced students from the public education system in Bogota for the years 2011 and 2012, we are only going to employ a sub-set of the ICFES dataset which contains data from students from the public education system who took the exit examination in Bogota during the second semester in those two years. For the year 2011 the dataset contains information about 52446 public education students who took the exam in the second semester of the year in Bogota. On the other hand, for the year 2012 the dataset contains information about 52100 students who took the exam in the second semester of the year.

Using the information provided by the Secretariat of Education of Bogota we were able to identify those public school students who took the high school exit examination and are also victims of the conflict (internally displaced). However, it should be noted that the dataset that was provided by the Secretariat of Education of Bogota originally did not have information about the performance of students in the Saber11 examination. Nevertheless, it was possible to match this dataset with the ICFES data in order to have a complete dataset containing information about all the students who took the exam in Bogota and identifying those students who have been displaced. This one to one matching was done using three different matching criteria for each student: school, date of birth and gender. For an observation contained in the dataset of the Secretariat of Education to be matched with an observation in the dataset of ICFES it was necessary that these three criteria were met.

According to the Secretariat of Education of Bogota, in the year 2011 there were 44520 students in the city's public school system (all grades) registered as victims of internal displacement (*Table 1*). Of those 44520 students, 2394 students were in the last year of high school and therefore had to take the high school exit examination in order to graduate from school. For the year 2012, 36778 students were registered as victims of internal displacement and 2144 of those students were in the last year of high school. After matching the information for these two years with the dataset of ICFES, 1661 observations of students who had been internally displaced were left for the year 2011 and 1502 observations were left for

the year 2012. A T-test shows that the data does not render systematic differences between those students who were matched and those who were not. Hence, it can be assumed that the matched sample is representative of the population.

Table 1. Reported Victims of Internal Displacement in Bogota's School System			
	Total Number of Students in the School System	Students in the Last Year of High School	Students in the Last Year with ICFES Results
2011	44520	2394	1661
2012	36778	2144	1502

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá

It should be acknowledged that a possible weakness related to the dataset used in this chapter is the fact that the information on forced displacement comes from individuals' self-reports. It is possible that some of the cases reported are misreports. Furthermore, there is also a possibility of under-reporting. However, the incidence of misreports should be low because in order for individuals to be officially considered victims of the conflict by virtue of forced displacement, they have to be registered in the RUPD registry. In the case of under-reporting, it is plausible that some of the students who were not categorized as victims of displacement to be victims of displacement. In fact, some students could have migrated to Bogotá without applying for the status of victim of displacement. However, it is impossible to have an estimation of those individuals who do not report their displacement status. This weakness could be solved, at least partially, by controlling in the empirical model for the place of birth and/or the intensity of conflict at the place of birth. However, the Colombian Institute for the Evaluation of Education (ICFES) unfortunately does not report the municipality of birth of the test takers. Nevertheless, we do have information about the expulsion site (municipality) and we control for that in all the specifications considered in the empirical exercise to overcome this limitation at least partially. Yet, this limitation should be considered closely and it is particularly important to take into consideration the possibility of under-reporting in the analysis of the empirical results as these might be capturing a lower-bound of the effect.

Another possible weakness related to the dataset is that it is not possible to identify whether a student that is considered a victim of displacement has been forced to move more than once. It is possible for those students that have been forced to move more than once to have a different level of affectation because of the internal

conflict. However, it is impossible to have an estimation of those individuals who were forced to move more than once using the information that is available.

8. Descriptive Statistics

This section presents descriptive statistics of the dataset described in the previous section. These descriptive statistics are presented to provide additional motivation for the empirical exercise that is exposed in the sections that follow. *Tables 2 and 3* present the summary statistics of the scores obtained by the students in the sample in the Saber11 exam. Out of a total of 52446 observations corresponding to the year 2011, 1661 observations are from students that have been identified as internally displaced. On the other hand, for the year 2012 out of a total of 52100 observations, 1502 correspond to students that have been identified as internally displaced. This means that the students that have been displaced roughly represent 3.2 per cent of the total sample in the year 2011, and 2.9 per cent in the year 2012. As we can see in *Tables 2 and 3*, the average score in the mathematics and the language examination for those students that have been displaced is lower than the average score of those students that have not been displaced in the two years considered. The average score in mathematics for the entire sample in 2011 is 47.15, and in 2012 it is 48.44. In the case of the language examination the average score for the entire sample in 2011 is 47.62, and in 2012 it is 47.28. *Graphs 2 and 3* illustrate the distribution of the scores in the mathematics and language examinations for the two years that are under analysis.

Table 2. Scores in the Saber11 Exam 2011						
Variable	Students that have been displaced		Students that have not been displaced		All the students in the sample	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Math Score in the Saber11 Exam	46.23	10.55	47.18	10.82	47.15	10.81
Language Score in the Saber11 Exam	46.71	9.79	47.65	10.02	47.62	10.02
Observations	1661		50785		52446	

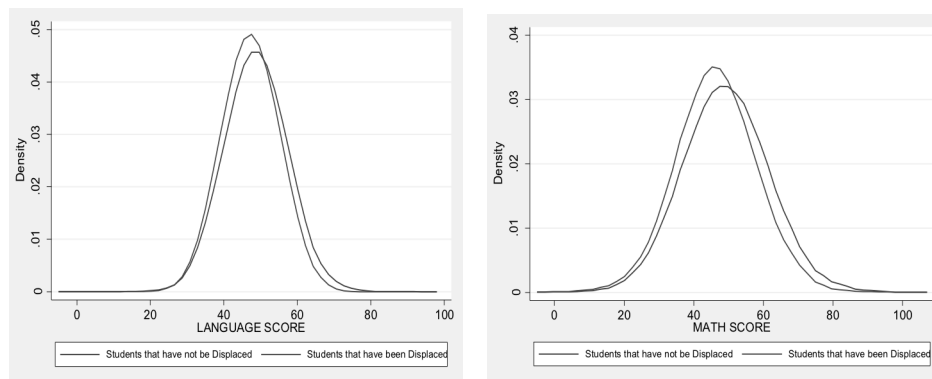
Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá

Table 3. Scores in the Saber11 Exam 2012						
Variable	Students that have been displaced		Students that have not been displaced		All the students in the sample	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Math Score in the Saber11 Exam	45.54	9.94	46.48	10.52	46.45	10.50
Language Score in the Saber11 Exam	47.17	6.12	47.49	6.53	47.48	6.52
Observations	1502		50598		52100	

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá

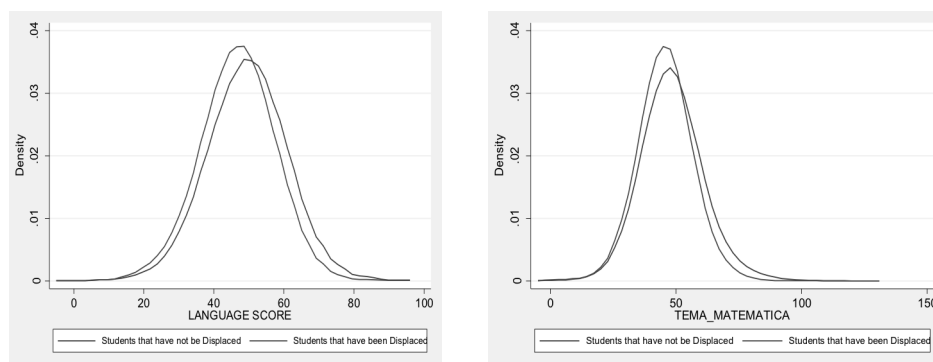
The descriptive statistics discussed above suggest that those students that have been directly affected by the Colombian civil conflict, and therefore have been forced to flee from their hometown, have a lower level of performance than those students that have not been directly affected. However, there are other determinants of educational performance (e.g. parental education, gender, age, family income, High School teachers) that should also be analysed in order to determine the strength of the relationship between displacement and educational performance. These additional determinants are going to play an important role in the econometric exercise that is going to be presented in the other sections of this chapter.

Graph 2. Distribution of Scores in the Saber11 Exam 2011



Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá

Graph 3. Distribution of Scores in the Saber11 Exam 2012



Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá

Information about the gender of the students in the sample is presented in *Table 4*. Around 55.5 per cent of the students in the public education system taking the examination in Bogotá in 2011 are female, and 54.8 per cent in 2012. Restricting the sample to those students that have been displaced, we can see that 58.1 per cent of the students in the 2011 sample that had been classified as internally displaced are female and 41.9 per cent are male. For the 2012 sample, 57.4 per cent of the students that had been classified as internally displaced are female and 42.6 per cent are male. In other words, there are more female students than male students in the sample as a whole and also if the sample is restricted to only those students that have been displaced.

Table 4. Student's Gender (as a percentage %)						
	Students that have been displaced		Students that have not been displaced		All the students in the sample	
	2011	2012	2011	2012	2011	2012
Male	41.9	42.6	44.6	45.3	44.5	45.2
Female	58.1	57.4	55.4	54.7	55.5	54.8

Source: author's calculations using the databases from ICFES.

Table 5 presents information about the educational background of the parents of the students in the sample. The educational background of the parents, which is captured in the variables *level of education of the student's father* and *level of education of the student's mother*, can go from no formal education to completion of graduate

school. For the results of the 2011 examination, on average 23.2 per cent of the mothers of the students that had been displaced had completed at least High School, while 25.6 per cent of the mothers of the students that had not been displaced had reached that educational level. For 2012, 22.5 percent of the mothers of the students that had been displaced had completed at least High School, while 26.9 percent of the mothers of the other students had reached that same level of schooling. On contrast, the number of fathers reaching that educational level is lower for both displaced and non-displaced students. For the case of the displaced students, 20.5 per cent of the fathers of those students taking the examination in 2011 had completed High School and 18.8 per cent had reached that same level for those students taking the examination in 2012. That same figure for the fathers of the other (non-displaced) students in the sample shows that 23.5 per cent of the fathers in the 2011 data had completed High School and that 23.9 per cent of the fathers in the 2012 data had also reached that level of schooling. To sum up, the dataset suggests that the average level of education of the student's mothers for both displaced and non-displaced students is higher than that of the father. It should also be highlighted that the number of parents of displaced students that did not complete any schooling or did not complete at least elementary school is almost double in comparison to the numbers of the parents of the students that had not been displaced because of the civil conflict (*Table 4*). For the econometric exercise presented in this chapter, we have decided to divide the parental education variable into four categories comprising the different levels of education discussed above. The four categories are the following: no education (parent didn't complete any studies), low level of education (elementary school), mid level of education (High school and/or technical studies) and high level of education (university and/or graduate school).

Even though a negative correlation does not necessarily imply a negative causal relationship between the variables under study, it indeed motivates a further exploration of this relationship. The correlation between displacement and the score obtained in the language and mathematics examinations is presented in *Table 6*. As expected, the correlation between these two variables is negative. The econometrical exercise presented in the next section further explores the relationship between these variables, controlling for other factors suggested in the literature.

Table 5. Parental Education								
	Father's Education				Mother's Education			
	Displaced Students		Other Students		Displaced Students		Other Students	
	2011	2012	2011	2012	2011	2012	2011	2012
	%	%	%	%	%	%	%	%
Didn't complete any studies	4.82	4	2.97	3.05	1.7	1.5	1.34	1.6
Incomplete Elementary School	22.9	24.63	17.6	17.5	21	21.0	14.9	14.7
Completed Elementary School	19.1	19.4	16.5	15.3	18.9	20.3	17.0	15.5
Incomplete High School	17.0	18.04	20.3	19.9	24.2	23.0	24.6	23.9
Completed High School	20.5	18.8	23.5	23.9	23.2	22.5	25.6	26.9
Technical Program not completed	0.84	1.13	1.76	1.78	1.2	1.1	2.34	2.4
Technical program completed	2.41	3	4.0	4.62	3.61	5.8	5.49	6.3
University degree not completed	1.74	1.13	1.71	1.8	0.96	0.9	1.91	1.94
University degree completed	3.61	3.2	4.24	4.47	2.8	1.8	3.9	4.01
Graduate school	0.48	0.47	0.8	0.74	0.73	0.4	0.85	0.84
Did not report any information	6.6	6.2	6.62	6.94	1.7	1.5	2.16	1.91

Source: author's calculations using the databases from ICFES

Table 6. Correlation between Scores and Displacement Status					
2011			2012		
Correlation	Math Exam	Language Exam	Correlation	Math Exam	Language Exam
displaced	-0.0157	-0.0164	displaced	-0.015	-0.0084

Source: author's calculations

It should be highlighted that a very high percentage of the students attending a public school in Bogotá do not pay tuition. *Table 7* presents information about the percentage of students who pay and do not pay tuition to attend public schooling in Bogotá.

Table 7. Students not paying tuition		
Variable	2011	2012
	%	%
Student does not pay tuition	96	97
Student pays tuition	4	3

Source: author's calculations using the databases from ICFES.

Descriptive statistics of all the other variables included in the econometric exercise are reported in the appendix 1 of this chapter (*Tables 8 to 12*).

9. Cluster Robust Inference

In this chapter, we consider a regression model where observations are grouped into clusters, with model errors uncorrelated across clusters but correlated within clusters. In this case we have student-level cross-sectional data with school clustering. The model errors for students in the same school may be correlated, while model errors for students in different schools are assumed to be uncorrelated. Refer to appendix 4 for a complete explanation of cluster-robust inference.

According to Cameron and Miller (2013), failure to control for within-cluster error correlation can lead to misleadingly small standard errors, and therefore misleadingly narrow confidence intervals, large t-statistics and low p-values. Clustered errors have two principal repercussions: there is a reduction in the precision of $\hat{\beta}$ and the estimator for the variance of $\hat{\beta}$ is usually biased downward from the true variance (Cameron and Miller 2013). In fact, as Cameron and Miller (2013) highlight, in most cases standard errors that control for within-cluster correlation are considerably larger than default standard errors that ignore such correlation.

Following Cameron and Miller (2013), let i denote the i th of N students in the sample, and g denote the g th of G clusters (schools). Then for individual i in cluster g the linear model with clustering is:

$$T_{ig} = x'_{ig}\beta + u_{ig} \quad (3)$$

where x_{ig} is a $K \times 1$ vector. It is also assumed that $E[u_{ig}|x_{ig}] = 0$. In this case the central assumption is that errors are uncorrelated across clusters, while errors for individuals belonging to the same cluster may be correlated. This implies that:

$$E[u_{ig}u_{jg'}|x_{ig}x_{jg'}] = 0, \text{ unless } g = g'.$$

10. Cluster Robust Inference Results

Tables 13 and 14 present the results of the main specification presented in equation (2) for the years 2011 and 2012. Section A presents the results using the score in the language exam as dependent variable, and section B presents the results using the score in the mathematics exam as dependent variable. Each of the columns in these tables present the results for the different specifications considered. These specifications include a set of controls variables to account for individual, household, parental, and school characteristics. All the regressions are estimated using ordinary least squares (OLS), clustering standard errors at the school level. *Column 1* corresponds to the results with no controls. In *Column 2*, a set of controls for individual characteristics are included as part of the specification⁴⁰. *Column 3* displays the results for the specification with household and parental controls⁴¹. In *Column 4*, the results controlling for school characteristics are reported⁴². *Column 5* displays the results obtained with individual, household, parental and school characteristics.

For the year 2011, the results show a negative and significant effect of displacement on the results in the mathematics exit examination in four out of five of the specifications that were considered in the analysis. The magnitude of the effect slightly differs depending on the control variables considered. Nevertheless it is consistently significant and negative. The complete specification shows that on average the students who were reported as displaced scored 0.52 points less than those students who were not reported as displaced in 2011. To get an idea of the size of the effect, we calculated the size of the coefficient as the proportion of a standard deviation of the test score which is being analyzed. This calculation indicates that on average an increase of one standard deviation in the variable that captures the displacement status causes a reduction of approximately 0.01 standard deviations in the score in the mathematics examination. On the other

⁴⁰ The school characteristics included as controls are: gender, age and whether the student had taken the exit examination before.

⁴¹ The household and parental characteristics included as controls are: level of education of the mother and the father, father's occupation, family income, availability of household appliances, household floor material, number of people living in the household.

⁴² The school characteristics included as controls are: class time, type of school, tuition, number of teachers per student, level of education of the teachers.

hand, for the year 2012 we only found a negative and significant effect in two of the five specifications which were considered. The specification that did not include controls and the specification with only school controls turned out to be significant. According to these specifications on average the students who were reported as displaced scored 1.16 (no controls) to 1.22 (school characteristics controls) points less than those students who were not reported as displaced. On average an increase of one standard deviation in the variable that captures the displacement status causes a reduction of approximately 0.02 standard deviations in the score in the mathematics examination in the case of the specification with school characteristics controls and a reduction of approximately 0.02 standard deviations in the score in the mathematics examination in the case of the specification with no controls.

In the case of the language examination, we also found a negative and statistically significant effect of displacement on the examination results for the year 2011 and for the year 2012. For the year 2011, the results show a negative and significant effect of displacement on the results in the language examination in four out of the five specifications that were considered in the analysis. As in the case of the mathematics examination, the magnitude of the effect differs depending on the control variables considered. Yet, the results show that the effect is consistently significant and negative under the different specifications. The complete specification shows that on average the students who were reported as displaced scored 1.14 points less than those students who were not reported as displaced in 2011. This indicates that on average an increase of one standard deviation in the variable that captures the displacement status causes a reduction of approximately 0.02 standard deviations in the score in the language examination. For the year 2012 we found a negative and significant effect in all of the specifications which were considered. The complete specification shows that on average the students who were reported as displaced scored 1.14 points less than those students who were not reported as displaced in 2012. This indicates that on average an increase of one standard deviation in the variable that captures the displacement status causes a reduction of approximately 0.03 standard deviations in the score in the language examination.

11. Instrumental Variables Approach

Following Calderon and Ibañez (2009), in this section we are going to use an instrumental variable in order to deal with the possibility of endogeneity as a source of bias. In fact, the instrumental variables approach considered in this section is our preferred method of estimation. Calderon and Ibañez (2009) propose an instrumental variable to correct for the possible bias generated by the possible

endogenous decision to locate in Bogota, the capital city of Colombia, taken by internally displaced populations. The instrument that these researchers propose is the massacres endured by the civil population in origin municipalities weighted by the distance to the destination site, in this case, Bogota. As Calderon and Ibañez (2009) have suggested, forced displacement which is triggered by armed conflicts and chronic violence does not respond to economic incentives. Instead, the massive flows of migration could be explained by the large number of individuals who are fleeing rural areas in search for safety. Although the migration decision is not related to educational decisions, the decision to locate in any given city may depend on the schooling opportunities that are being offered at each particular place. In other words, the decision to locate in Bogota might be associated with the opportunities, including educational, that the city offers. In order to correct for this possible source of endogeneity, we instrument the displacement status with a similar instrument to the one mentioned above. The instrument that we are going to use in this section is given by the ten-year average of conflict-related killings in the origin municipalities weighted by the distance to the destination site. In this case, I was able to verify that the instrument is strongly related to displacement, and it is exogenous to the schooling decisions taken by the students.

The information to construct the instrument was obtained from two different sources. Information about conflict-related killings was obtained from the Resource Centre for Conflict Analysis (CERAC). CERAC is a think tank based in Bogotá (Colombia) with three main lines of research: violent conflict, political violence, and violence associated to drug trafficking. Following these lines of research, an important area of study at CERAC is the quantification of violence directly associated to the internal armed conflict in Colombia. To quantify violent armed conflict, the research team at CERAC developed a methodology to collect, document, code and register information about this type of violence. The construction of CERAC's database started in the year 2002. On the other hand, the information that was used to construct the database with the approximate distances between the different municipalities and Bogota was obtained from the Instituto Geográfico Agustín Codazzi. This governmental institute is in charge of collecting spatial information of the different regions in Colombia.

Table 13. OLS clustering standard errors at the school level – Year 2011					
Section A. Dependent variable: language					
	(1)	(2)	(3)	(4)	(5)
Displaced	-1.88* (1.166)	-2.05* (1.172)	-2.32** (1.190)	-0.46 (1.120)	-1.142* (0.716)
Displaced * Gender		-0.520 (0.476)			-0.426 (0.482)
Displaced * Night School			-0.304 (1.108)		-0.369 (1.299)
Constant	47.70*** (0.269)	53.87** (0.420)	47.94*** (0.283)	41.36*** (0.553)	45.68*** (0.633)
Individual Characteristics Controls	No	Yes	No	No	Yes
Household Characteristics Controls	No	No	No	Yes	Yes
Parental Characteristics Controls	No	No	No	Yes	Yes
School Characteristics Controls	No	No	Yes	No	Yes
R-squared	0.003	0.026	0.03	0.07	0.09
Observations	52446	52445	50787	52445	50786
Section B. Dependent variable: mathematics					
	(1)	(2)	(3)	(4)	(5)
Displaced	-0.66** (0.328)	-0.96** (0.373)	-1.01*** (0.313)	-0.26 (0.297)	-0.523* (0.309)
Displaced * Gender		-0.532 (0.552)			-0.440 (0.560)
Displaced * Night School			-1.024 (1.733)		-1.211 (1.798)
Constant	47.50*** (0.252)	55.35*** (0.387)	47.923*** (0.192)	40.45*** (0.532)	46.56*** (0.590)
Individual Characteristics Controls	No	Yes	No	No	Yes
Household Characteristics Controls	No	No	No	Yes	Yes
Parental Characteristics Controls	No	No	No	Yes	Yes
School Characteristics Controls	No	No	Yes	No	Yes
R-squared	0.005	0.058	0.05	0.08	0.116
Observations	52446	52445	50787	52445	50786

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 14. OLS clustering standard errors at the school level – Year 2012					
Section A. Dependent variable: language					
	(1)	(2)	(3)	(4)	(5)
Displaced	-1.47*** (0.472)	- 1.12*** (0.358)	-1.62*** (0.469)	-0.93* (0.450)	-1.14** (0.465)
Displaced * Gender		-0.079 (0.350)			-0.010 (0.332)
Displaced * Night School			-1.511 (1.068)		-0.736 (0.959)
Constant	47.38*** (0.160)	51.59* ** (0.261)	47.65*** (0.159)	43.66*** (0.363)	46.32*** (0.385)
Individual Characteristics Controls	No	Yes	No	No	Yes
Household Characteristics Controls	No	No	No	Yes	Yes
Parental Characteristics Controls	No	No	No	Yes	Yes
School Characteristics Controls	No	No	Yes	No	Yes
R-squared	0.002	0.02	0.04	0.07	0.1
Observations	52016	52016	50486	52012	50482
Section B. Dependent variable: mathematics					
	(1)	(2)	(3)	(4)	(5)
Displaced	-1.16* (0.706)	-0.95 (0.784)	-1.22* (0.760)	-0.43 (0.716)	-0.434 (0.704)
Displaced * Gender		-0.210 (0.548)			-0.330 (0.512)
Displaced * Night School			-1.04 (2.093)		0.052 (2.144)
Constant	46.58*** (0.255)	50.13* ** (0.470)	46.88*** (0.259)	41.33*** (0.528)	42.95*** (0.640)
Individual Characteristics Controls	No	Yes	No	No	Yes
Household Characteristics Controls	No	No	No	Yes	Yes
Parental Characteristics Controls	No	No	No	Yes	Yes
School Characteristics Controls	No	No	Yes	No	Yes
R-squared	0.002	0.05	0.03	0.06	0.11
Observations	52016	52016	50486	52012	50482

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

12. Instrumental Variables Approach Results

Tables 15 and 16 show the results of the instrumental variable approach. Section A presents the results using the score in the language exam as dependent variable, and section B presents the results using the score in the mathematics exam as dependent variable. Each of the columns in these tables present the results for the different specifications considered. These specifications include a set of controls variables to account for individual, household, parental, and school characteristics.

For both years, the results for the mathematics examination show the same trend as the one found using OLS cluster robust inference. The results confirm that there is a negative and significant effect of displacement on the performance in the mathematics exit examination. However, the results obtained using the instrumental variables approach show a larger effect of displacement on the performance in the examination. The larger effect obtained using an instrumental variables⁴³ approach suggests that the bias in using clustered ordinary least squares is significant. In that sense, the results obtained using OLS cluster robust inference seem to be a lower bound of the effect. As it can be seen in *table 15 (Column 5- section B)*, the results obtained for 2011 using the complete specification show that the score in the mathematics examination for a student who had been displaced is about 2.29 points lower than the score obtained by a student who had not been displaced. To get an idea of the size of the effect, we calculated the size of the coefficient as the proportion of a standard deviation of the test score which is being analyzed. This calculation indicates that on average an increase of one standard deviation in the variable that captures the displacement status causes a reduction of approximately 0.04 standard deviations in the score in the mathematics examination. For the year 2012 we did not find a significant effect using the complete specification. However, we found a significant effect in the other specifications that were considered. The specification that did not include controls, the specification with only school controls, and the specification with only household and parental controls turned out to be significant. As it can be seen in *table 16 (section B)*, the results show that the students who were reported as displaced scored 3.48 (household and parental characteristics) to 3.66 (school characteristics controls) points less than those students who were not reported as displaced. Those results imply that on average an increase of one standard deviation in the variable that captures the displacement status causes a reduction of approximately 0.056 standard deviations in the score in the mathematics examination in the case of the specification with household and parental characteristics controls and a reduction of approximately 0.059 standard deviations

⁴³ An instrumental variable (IV) approach is commonly used to address endogeneity issues. Some of the problems that are usually found in using an ordinary least regression and can be tackled using IV are the following: omitted variable bias, measurement error, and simultaneity or reverse causality.

in the score in the mathematics examination in the case of the specification with school characteristics controls.

Table 15. Instrumental Variables Approach (cluster by school) – Year 2011					
Section A. Dependent variable: language					
	(1)	(2)	(3)	(4)	(5)
Displaced	-2.61* (1.554)	-2.69* (1.517)	-3.16** (1.579)	-1.67 (1.589)	-2.27** (1.084)
Constant	47.83*** (0.393)	53.99*** (0.490)	48.10*** (0.402)	41.22*** (0.628)	45.63*** (0.705)
Individual Characteristics Controls	No	Yes	No	No	Yes
Household Characteristics Controls	No	No	No	Yes	Yes
Parental Characteristics Controls	No	No	No	Yes	Yes
School Characteristics Controls	No	No	Yes	No	Yes
Root MSE	10.01	9.887	9.87	9.666	9.577
Hansen J Statistic	2.326	2.296	2.122	2.235	1.456
P-value	0.1273	0.1297	0.1452	0.1349	0.2275
Observations	52082	52082	50439	52081	50438
Section B. Dependent variable: mathematics					
	(1)	(2)	(3)	(4)	(5)
Displaced	-3.03** (1.227)	-3.05*** (1.092)	-3.42*** (1.251)	-1.91 (1.273)	-2.29** (1.159)
Constant	47.22*** (0.326)	55.04*** (0.453)	47.69*** (0.323)	40.14*** (0.609)	46.25*** (0.646)
Individual Characteristics Controls	No	Yes	No	No	Yes
Household Characteristics Controls	No	No	No	Yes	Yes
Parental Characteristics Controls	No	No	No	Yes	Yes
School Characteristics Controls	No	No	Yes	No	Yes
Root MSE	10.8	10.49	10.55	10.42	10.15
Hansen J Statistic	2.282	2.243	2.091	1.742	1.283
P-value	0.1309	0.1342	0.1482	0.1869	0.2574
Observations	52082	52081	50439	52081	50438

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 16. Instrumental Variables Approach (cluster by school) – 2012					
Section A. Dependent variable: language					
	(1)	(2)	(3)	(4)	(5)
Displaced	-1.92* (1.098)	-1.90* (1.077)	-2.07** (1.108)	-2.05* (1.114)	-2.15* (1.127)
Constant	47.36*** (0.219)	51.42*** (0.330)	47.63*** (0.222)	43.61*** (0.399)	46.23*** (0.426)
Individual Characteristics Controls	No	Yes	No	No	Yes
Household Characteristics Controls	No	No	No	Yes	Yes
Parental Characteristics Controls	No	No	No	Yes	Yes
School Characteristics Controls	No	No	Yes	No	Yes
Root MSE	6.518	6.455	6.392	6.296	6.226
Hansen J Statistic	2.425	2.700	2.706	1.632	2.191
P-value	0.1194	0.1003	0.1001	0.2015	0.1388
Observations	51706	51706	50177	51702	50173
Section B. Dependent variable: mathematics					
	(1)	(2)	(3)	(4)	(5)
Displaced	-3.58** (1.486)	-1.70 (1.502)	-3.66** (1.583)	-3.48** (1.468)	-1.85 (1.539)
Constant	46.37*** (0.311)	49.89*** (0.505)	46.66*** (0.328)	41.07*** (0.567)	42.70*** (0.677)
Individual Characteristics Controls	No	Yes	No	No	Yes
Household Characteristics Controls	No	No	No	Yes	Yes
Parental Characteristics Controls	No	No	No	Yes	Yes
School Characteristics Controls	No	No	Yes	No	Yes
Root MSE	10.5	10.23	10.36	10.22	9.954
Hansen J Statistic	1.182	1.256	1.392	0.725	0.988
P-value	0.2769	0.2624	0.2381	0.3944	0.3203
Observations	51706	51706	50177	51702	50173

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. **Note:** the coefficients obtained for the individual, household, parental and school controls are included in the appendix of this chapter. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

In the case of the language examination, the results also confirm that there is a negative and significant effect of displacement on the performance in the exam. For the year 2011, we find a statistically significant effect of displacement on the

examination results for the year 2011 in four of the five specifications considered⁴⁴. As it can be seen in *table 15 (section A – Column 5)*, the complete specification shows that on average the students who were reported as displaced scored 2.27 points less than those students who were not reported as displaced in 2011. This indicates that on average an increase of one standard deviation in the variable that captures the displacement status causes a reduction of approximately 0.04 standard deviations in the score in the language examination. In line with the findings obtained using OLS cluster robust inference, these results show a negative and significant effect of displacement on the performance in the language examination. Yet, the results obtained using the instrumental variable approach show a larger effect of displacement on the performance in the language examination than the results obtained using OLS cluster robust inference. As mentioned previously, the larger effect obtained using an instrumental variables approach suggests that the bias in using clustered ordinary least squares is significant. For the year 2012 we also found a negative and significant effect in all the specifications that were considered for the analysis. The complete specification, depicted in *table 16 (section A – Column 5)*, shows that on average the students who were reported as displaced scored 2.15 points less than those students who were not reported as displaced in 2012. This indicates that on average an increase of one standard deviation in the variable that captures the displacement status causes a reduction of approximately 0.055 standard deviations in the score in the language examination for that year.

13. Discussion of the Results

The main finding of this chapter is that being displaced has a negative and significant effect on the levels of achievement of students in mathematical and language exit examinations. The results presented are not surprising if we take into consideration the findings of a number of researchers who have studied the effects of displacement. Although to the best of my knowledge this is the first attempt to try to understand and quantify the impact of displacement on educational achievement measured by standardized exams, it is possible to find in the literature very important contributions that have studied the effects on welfare of the displacement. These contributions are going to help us establish possible factors through which displacement could affect the performance of students in exit examinations, and in doing so, helps us contextualize and get a better understanding of the results obtained in this chapter.

⁴⁴ The variable that captures the displacement status is not statistically significant in the specification that includes household and parental characteristics controls.

There are a number of channels through which internal displacement could potentially affect the performance of students in High School exit examinations. The literature has shown that there are a number of needs and vulnerabilities associated to the displacement status that affect the well-being of individuals. According to the literature, some of the factors that generate and heighten existing vulnerabilities are: loss of property (Ibañez and Moya 2010a); reduced consumption levels (Ibañez and Moya 2010b); early entrance into the labor market (Rodriguez and Sanchez 2012); fear and psychological trauma; and separation from family members, social networks and communities (Ferris and Winthrop 2010). All of these factors can be seen as channels through which the displacement status could potentially affect educational results.

The first factor that can explain the lower levels of achievement amongst displaced students is asset loss. Ibañez and Moya (2010a) have found that regardless of the extent of asset loss, displaced households are left with an asset base insufficient to escape poverty. According to these researchers, displaced households do not catch up even after they have consolidated their settlement at destination sites. Ibañez and Moya (2010a) highlight that individuals who had been displaced become trapped in a low level economic trajectory, and in that sense, their return to a high level economic trajectory becomes very unlikely. As a consequence, individuals who have been displaced are more likely to have fewer resources available to spend in education and educational materials. Additionally, as a consequence of the asset loss, it is more likely for those students who have been displaced to enter the labor market earlier (Rodriguez and Sanchez 2012). If that is the case, those students who have been displaced might have less time to study and this can reflect in lower levels of achievement.

A second factor studied in the literature that could help us understand our results is the reduced consumption levels amongst displaced families. Ibañez and Moya (2010b) show that displaced households are not able to isolate consumption from variations in transitory income, meaning that they are not able to smooth consumption. The results obtained by Ibañez and Moya (2010b) suggest that variations in transitory income turn into variations in consumption. According to these researchers, the harsh conditions in receptor sites, poor labor conditions and a severe disruption of risk-sharing mechanisms are the principal factors limiting the capacity of households to smooth consumption (Ibañez and Moya 2010b). The lower levels of consumption could potentially be affecting the educational results of those students coming from families that have been displaced. Moreover, the limited levels of consumption could have negative effects on the overall welfare levels of families affected by displacement. In fact, Ibañez and Velez (2008) find that welfare losses from displacement are substantial. These authors find that the costs of displacement amount to 37 percent of the net present value of aggregated

rural consumption. These researchers also highlight that the economic affectation of displacement is higher for poor households (Ibañez and Velez 2008). If that is the case, it is very possible that those students who have been victims of displacement dedicate less time to study in order to get a job to earn an income to help increase the family income and meet the basic needs. In fact, as Rodriguez and Sanchez (2012) have highlighted it is more likely for those students who have been displaced to enter the labor market earlier than those who have not.

A third factor is early entrance into the labor market. It is more likely for those students who have been displaced to enter the labor market earlier in order to help improve the family income if their parents have a low pay job or are unemployed (Rodriguez and Sanchez 2012). If that is the case, those students who have been displaced might have less time to study and this can reflect in lower levels of achievement.

A fourth factor is psychological trauma. According to Hoeffler and Reynal-Querol (2003), being exposed to the brutality of conflict and displacement can leave individuals psychologically scarred. These researchers highlight that the victims of conflict are prone to mental health problems even after exposure to shootings and bombings has ended (Hoeffler and Reynal-Querol 2003). The psychological trauma associated to conflict and the subsequent displacement status could have a negative effect in terms of educational achievement that can explain our results.

Those students who have been victims of displacement are indeed facing more needs than those students who have not, and this might be reflecting in lower levels of achievement. However, those needs are not limited to those associated to economic difficulties. Individuals can also be affected by social and community factors. A fifth factor is separation from family members, social networks and communities. Research done by Ferris and Winthrop (2010) has highlighted that conflicts have a profound impact on family and community life which ultimately affects children and their educational process.

All of the factors discussed above impose a heavy burden upon individuals that have been displaced, and can help us explain the results presented. According to the literature the changes that displaced individuals have to confront are often times associated to the deterioration of their quality of life (Kirschhoff and Ibañez 2002). As highlighted by Carrillo (2009), displaced individuals often times are unable to become self-reliant and meet their basic needs without permanent assistance from the government and humanitarian organizations. In that sense, our results ratify the conclusions of other authors regarding the need for governmental

programs specially designed to accommodate the particular needs of the displaced population in the short and medium term.

14. Concluding Remarks

This chapter studies the effects of forced displacement on the performance of Colombian students in the national high school exit examination. The findings of this chapter contribute to the existing literature that studies the impact of civil conflict on education, and to the literature that studies the impact of internal displacement on the welfare of civilians. To the best of my knowledge this is the first effort to try to understand and quantify the impact of forced displacement on education achievement (measured by standardized exams).

The results obtained in this chapter provide clear evidence of the negative repercussions that displacement can have on human capital accumulation. For both years, the results of the estimations using OLS cluster robust inference and instrumental variable show that there is a negative and significant effect of displacement on the performance in the mathematics and language exit examinations. However, the results obtained using the instrumental variables approach show a larger effect of displacement on the performance in the examinations. The results show that school age individuals who have been displaced from conflict zones are prone to lower levels of academic achievement, and therefore to lower levels of human capital accumulation. In that sense, our results ratify the conclusions of other authors regarding the need for governmental programs specially designed to accommodate the particular needs of the displaced population in the short and medium term.

There are a number of factors through which internal displacement could have potentially affected the performance of students in High School exit examinations. The literature has shown that there are a number of needs and vulnerabilities associated to the displacement status that affect the well-being of individuals, and in doing so, might also be affecting the performance of students in standardized tests. According to the literature, some of the factors that generate and heighten existing vulnerabilities are: loss of property (Ibañez and Moya 2010a); reduced consumption levels (Ibañez and Moya 2010b); early entrance to the labor market (Rodriguez and Sanchez 2012); fear and psychological trauma; and separation from family members, social networks and communities (Ferris and Winthrop 2010). All of these factors can be seen as channels through which the displacement status could have potentially affected educational results.

The evidence obtained in this chapter can be used to guide the formulation of educational policy targeted specifically for the displaced population. Policy makers should work in the creation of new and better incentives to keep displaced students in school with a full time dedication, and delay their entrance into the labor market. If school age individuals can dedicate their time and effort exclusively to schooling, the results in the examinations will likely be better. Most of the transmission channels through which internal displacement could have potentially affected the performance of students in High School exit examinations were related to the low levels of income amongst displaced households. Consequently, the authorities should implement programs to help displaced families earn a higher income (i.e. employment programs, conditional cash transfer programs, and other forms of subsidies) in order to reduce some of the needs and vulnerabilities associated to the displacement status that affect the well-being of students. Moreover, more governmental resources should be spent in order to meet the special needs of displaced students at schools (i.e. psychological help, school supplies, special remedial tutorials) and to supervise their progress at educational institutions.

Appendix 1. Descriptive Statistics

Table 8. Descriptive Statistics (Results in the Examination)				
Variable	2011		2012	
	Mean	S.D.	Mean	S.D.
Mathematics	47.15	10.81	46.45	10.50
Language	47.62	10.02	47.48	6.52
Number of Observations	52446		52100	

Source: author's calculations using the databases from ICFES.

Table 9. Descriptive Statistics (Individual Characteristics)				
Variable	2011		2012	
	Mean	S.D.	Mean	S.D.
Gender	0.445	0.497	0.452	0.498
Age	17.43	3.86	17.383	3.743
Student has taken the exit exam before	0.069	0.253	0.035	0.183
Displacement status	0.032	0.175	0.029	0.167
Number of Observations	52446		52100	

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá.

Table 10. Descriptive Statistics (School Characteristics)				
Variable	2011		2012	
	Mean	S.D.	Mean	S.D.
Classes are at night	0.078	0.268	0.076	0.265
Technical/vocational school	0.162	0.368	0.160	0.367
Total teachers per student	0.179	0.279	0.186	0.388
Female teachers per student	0.018	0.027	0.017	0.026
Male teachers per student	0.017	0.026	0.016	0.023
Teachers with college degree per student	0.091	0.213	0.097	0.347
Teachers with teaching diploma per student	0.008	0.016	0.001	0.002
Teachers with technical degree per student	0.002	0.008	0.0007	0.003
Teachers with high school degree per student	0.005	0.006	0.001	0.012
Number of Observations	52446		52100	

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá.

Table 11. Descriptive Statistics (Household Characteristics)				
Variable	2011		2012	
	Mean	S.D.	Mean	S.D.
Very low income (1 to 2 Colombian min.wages)	0.170	0.376	0.158	0.364
Low income (2 to 3 Colombian min.wages)	0.749	0.434	0.759	0.427
Mid income (3 to 9 Colombian min.wages)	0.075	0.263	0.076	0.266
High income (10+ Colombian min.wages)	0.006	0.078	0.007	0.081
Number of people living in the household	4.94	1.94	4.88	1.94
Household owns a computer	0.644	0.479	0.716	0.451
Household has internet service	0.510	0.499	0.600	0.490
Household owns a refrigerator	0.886	0.318	0.896	0.305
Household floor material – earth/soil (<i>reference category</i>)	0.015	0.121	0.013	0.114
Household floor material – cement	0.327	0.469	0.320	0.466
Household floor material – low quality tiles and/or wood	0.078	0.266	0.078	0.268
Household floor material – high quality tiles and/or carpet	0.581	0.493	0.590	0.492
Number of Observations	52446		52100	

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá.

Table 12. Descriptive Statistics (Parental Characteristics)				
Variable	2011		2012	
	Mean	S.D.	Mean	S.D.
Father - didn't complete any studies	0.030	0.171	0.031	0.173
Father – low level of education (up to elementary school)	0.343	0.475	0.331	0.471
Father – mid level of education (high school and/or technical studies)	0.493	0.500	0.499	0.500
Father – high level of education (university and/or graduate school)	0.067	0.250	0.069	0.254
Mother - didn't complete any studies	0.014	0.115	0.016	0.125
Mother – low level of education (up to elementary school)	0.321	0.467	0.305	0.461
Mother – mid level of education (high school and/or technical studies)	0.578	0.494	0.593	0.491
Mother – high level of education (university and/or graduate school)	0.066	0.248	0.066	0.249
Father occupation – inactive	0.539	0.498	0.550	0.497
Father occupation – unemployed	0.160	0.366	0.153	0.360
Father occupation – qualified independent	0.032	0.175	0.029	0.168
Father occupation – not qualified independent	0.051	0.219	0.051	0.221
Father occupation – qualified dependent	0.096	0.295	0.100	0.300
Father occupation – not qualified dependent	0.123	0.328	0.117	0.321
Number of Observations	52446		52100	

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá.

Appendix 2. Results OLS Clustering Standard Errors at the School Level

Table 17. OLS clustering standard errors at the school level - No controls and Individual Characteristics Controls								
	Language 2011		Mathematics 2011		Language 2012		Mathematics 2012	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Constant	47.70*** (0.269)	53.87*** (0.420)	47.50*** (0.193)	55.35*** (0.387)	47.38*** (0.160)	51.59*** (0.261)	46.58*** (0.255)	50.13*** (0.470)
Displaced	-1.88* (1.166)	-2.05* (1.172)	-0.66** (0.328)	-0.96** (0.373)	-1.47*** (0.472)	-1.12*** (0.373)	-1.16* (0.706)	-0.95 (0.784)
Expulsion Municipality	-4.07e-06 (0.000)	-5.01e-06 (0.000)			0.00001 (0.000)	0.00002 (0.000)	-9.72e-06 (0.000)	0.00001 (0.000)
Gender		1.62*** (0.454)		3.59*** (0.539)		0.262 (0.338)		3.94*** (0.533)
Age		-0.383*** (0.015)		-0.53*** (0.018)		-0.24*** (0.012)		-0.31*** (0.018)
Student has taken the exit exam before		0.20 (0.245)		0.460* (0.268)		0.079 (0.197)		0.191 (0.329)
Displaced * Gender		-0.520 (0.476)		-0.532 (0.552)		-0.079 (0.350)		-0.205 (0.548)
R-squared	0.003	0.026	0.003	0.058	0.002	0.02	0.002	0.052
Observations	52446	52445	52446	52445	52016	51827	52016	52016

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. Significance

levels: *** p<0.01, ** p<0.05, * p<0.1

Table 18. OLS clustering standard errors at the school level - School Characteristics Controls				
	Language 2011	Mathematics 2011	Language 2012	Mathematics 2012
	(1)	(1)	(1)	(1)
Constant	48.03*** (0.161)	47.92*** (0.192)	47.65*** (0.159)	46.88*** (0.259)
Displaced	-2.32* (1.199)	-1.01*** (0.313)	-1.63*** (0.469)	-1.22* (0.750)
Expulsion Municipality	-5.43e-06 (0.00002)	-0.00002** (0.00001)	-9.10e-06 (0.00001)	-0.00001 (0.00001)
Classes are at night	-5.30*** (1.048)	-9.12*** (1.743)	-2.72** (1.094)	-4.25** (2.128)
Displaced * Classes are at night	-0.30 (1.108)	-1.02 (1.734)	-1.51 (1.068)	-1.04 (2.093)
Technical/vocation al school	1.02*** (0.390)	0.99*** (0.377)	0.529* (0.287)	1.01** (0.430)
Teachers with teaching diploma per student	0.34 (6.171)	2.46 (6.833)	6.00 (14.638)	25.27 (18.903)
R-squared	0.03	0.05	0.04	0.03
Observations	50787	50787	50486	50486

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 19. OLS clustering standard errors at the school level Household and Parental Characteristics Controls				
	Language 2011	Mathematics 2011	Language 2012	Mathematics 2012
Constant	41.36*** (0.553)	40.45*** (0.532)	43.66*** (0.363)	41.33*** (0.528)
Displaced	-0.46 (1.120)	-0.26 (0.298)	-0.93** (0.450)	-0.43 (0.716)
Expulsion Municipality	-0.00001 (0.00002)	-0.00002 (0.00001)	0.00001 (9.62e-06)	-2.57e-06 (0.00001)
Low income (2 to 3 Colombian min.wages)	1.21*** (0.134)	1.31*** (0.142)	0.74*** (0.089)	1.40*** (0.141)
Mid income (3 to 9 Colombian min.wages)	2.35*** (0.266)	2.80*** (0.301)	1.65*** (0.161)	3.38*** (0.313)
High income (10+ Colombian min.wages)	1.54** (0.697)	1.44* (0.766)	1.99*** (0.616)	3.67*** (1.020)
# of People living in the household	-0.16*** (0.024)	-0.16*** (0.027)	-0.11*** (0.015)	-0.10*** (0.024)
Household owns a computer	1.21*** (0.125)	1.32*** (0.146)	0.30*** (0.028)	0.39*** (0.049)
Household has internet service	0.40*** (0.124)	0.65*** (0.137)	0.28*** (0.086)	0.24* (0.138)
Household floor material – cement	1.97*** (0.399)	1.68*** (0.344)	1.23*** (0.277)	1.51*** (0.389)
Household floor material – low quality tiles and/or wood	2.51*** (0.443)	2.15*** (0.393)	1.66*** (0.302)	1.75*** (0.434)
Household floor material – high quality tiles and/or carpet	2.89*** (0.415)	2.40*** (0.350)	1.80*** (0.278)	2.11*** (0.391)
Father – low level of education (up to elementary school)	-1.30 (0.985)	-0.60 (0.493)	-0.78 (0.526)	-0.87 (0.688)
Father – mid level of education (high school and/or technical studies)	0.14 (0.181)	0.76*** (0.190)	0.15 (0.127)	0.22 (0.178)
Father – high level of education (university and/or graduate school)	2.77*** (0.260)	3.32*** (0.303)	1.65*** (0.211)	2.74*** (0.292)
Mother – low level of education (up to elementary school)	2.06*** (0.296)	2.00*** (0.359)	0.64*** (0.183)	0.58** (0.262)
Mother – mid level of education (high school and/or technical studies)	3.57*** (0.299)	3.90*** (0.374)	1.64*** (0.183)	2.12*** (0.270)
Mother – high level of education (university)	5.37*** (0.350)	6.01*** (0.427)	3.16*** (0.230)	4.28*** (0.374)

and/or graduate school)				
Father occupation – unemployed	-1.14*** (0.133)	-1.28*** (0.156)	-0.75*** (0.096)	-0.94*** (0.145)
Father occupation – qualified independent	-2.50*** (0.232)	-3.11 (1.957)	-1.71*** (0.163)	-2.38*** (0.292)
Father occupation – not qualified independent	0.222 (0.219)	0.02 (0.240)	0.09 (0.142)	0.07 (0.253)
Father occupation – qualified dependent	0.08 (0.151)	-0.203 (0.159)	-0.01 (0.101)	-0.11 (0.160)
Father occupation – not qualified dependent	-0.41*** (0.133)	-0.73*** (0.150)	-0.34*** (0.093)	-0.52*** (0.148)
R-squared	0.08	0.07	0.07	0.05
Observations	52245	52445	52012	52012

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 20. OLS clustering standard errors at the school level Complete Specification				
	Language 2011	Mathematics 2011	Language 2012	Mathematics 2012
	(1)	(1)	(1)	(1)
Constant	45.68*** (0.633)	46.56*** (0.590)	46.32*** (0.385)	42.95*** (0.640)
Displaced	-1.14* (0.716)	-0.52* (0.309)	-1.14** (0.465)	-0.43 (0.704)
Expulsion Municipality	-3.78e-06 (0.00002)	-0.00002 (0.00001)	0.00001 (9.69e-06)	-6.26e-06 (0.00001)
Displaced*Gender	-0.43 (0.482)	-0.44 (0.560)	-0.01 (0.332)	-0.33 (0.512)
Displaced*Night School	-0.37 (1.299)	-1.21 (1.798)	-0.74 (0.959)	-0.05 (2.144)
Gender	1.29*** (0.467)	3.28*** (0.551)	0.06 (0.325)	3.65*** (0.505)
Age	-0.18*** (0.016)	-0.26*** (0.019)	-0.07*** (0.012)	-0.12*** (0.019)
Student has taken the exit exam before	-0.61*** (0.207)	-0.39* (0.222)	-0.16 (0.173)	-0.17 (0.312)
Low income (2 to 3 Colombian min.wages)	1.13*** (0.131)	1.12*** (0.133)	0.70*** (0.088)	1.28*** (0.135)
Mid income (3 to 9 Colombian min.wages)	2.29*** (0.237)	2.50*** (0.236)	1.53*** (0.131)	3.01*** (0.231)
High income (10+ Colombian min.wages)	1.78*** (0.660)	1.27* (0.704)	1.63*** (0.556)	3.31*** (0.769)
# of People living in the household	-0.18*** (0.024)	-0.19*** (0.025)	-0.13 (0.015)	-0.10*** (0.024)
Household owns a computer	0.91*** (0.124)	0.85*** (0.138)	0.23*** (0.027)	0.27*** (0.049)
Household has internet service	0.37*** (0.121)	0.58*** (0.131)	0.24*** (0.085)	0.16 (0.133)
Household floor material – cement	1.87*** (0.374)	1.10*** (0.363)	1.10*** (0.270)	1.24*** (0.408)
Household floor material – low quality tiles and/or wood	2.42*** (0.408)	1.60*** (0.397)	1.48*** (0.305)	1.58*** (0.470)
Household floor material – high quality tiles and/or carpet	2.83*** (0.385)	1.92*** (0.366)	1.63*** (0.271)	1.98*** (0.410)
Father – low level of education (up to elementary school)	-1.31 (0.184)	-0.52 (0.391)	-0.76 (0.425)	-0.74 (0.483)
Father – mid level of education (high school and/or technical studies)	-0.12 (0.180)	0.42** (0.190)	0.02 (0.127)	0.04 (0.174)

Father – high level of education (university and/or graduate school)	2.48*** (0.252)	2.90*** (0.281)	1.61*** (0.184)	2.53*** (0.260)
Mother – low level of education (up to elementary school)	1.49*** (0.293)	1.11*** (0.290)	0.23 (0.176)	0.32 (0.268)
Mother – mid level of education (high school and/or technical studies)	2.68*** (0.290)	2.45*** (0.289)	1.00*** (0.173)	1.53*** (0.271)
Mother – high level of education (university and/or graduate school)	4.40*** (0.337)	4.39*** (0.341)	2.45*** (0.212)	3.66*** (0.323)
Father occupation – unemployed	-1.04*** (0.135)	-1.07*** (0.147)	-0.65*** (0.090)	-0.73*** (0.132)
Father occupation – qualified independent	-2.51 (1.436)	-3.28 (1.957)	-1.70 (1.158)	-2.42 (1.280)
Father occupation – not qualified independent	0.16 (0.219)	-0.13 (0.234)	-0.03 (0.139)	-0.20 (0.231)
Father occupation – qualified dependent	0.05 (0.153)	-0.28 (0.253)	-0.01 (0.100)	-0.16 (0.153)
Father occupation – not qualified dependent	-0.41*** (0.130)	-0.76*** (0.141)	-0.33*** (0.091)	-0.53*** (0.144)
Classes are at night	-2.04*** (0.281)	-5.35*** (1.797)	-1.92** (0.981)	-2.85 (2.215)
Technical/vocational school	0.66** (0.292)	0.63** (0.287)	0.29** (0.209)	0.75** (0.340)
Total teachers per student	-0.24 (0.215)	-0.15 (0.190)	-0.26** (0.103)	-0.36 (0.240)
R-squared	0.08	0.115	0.1	0.11
Observations	50786	50786	50482	50482

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. **Significance**

levels: *** p<0.01, ** p<0.05, * p<0.1

Appendix 3. Results Instrumental Variables Approach Clustering at the School Level

Table 21. Instrumental Variables Approach clustering at the school level No controls and Individual Characteristics Controls								
	Language 2011		Mathematics 2011		Language 2012		Mathematics 2012	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Displaced	-2.06** (0.957)	-2.07** (0.939)	-3.03*** (1.227)	-3.05*** (1.092)	-1.92*** (1.098)	-1.90* (1.077)	-3.58** (1.486)	-1.70*** (1.502)
Expulsion Municipality	-0.00002 (0.000)	-0.00001 (0.000)	2.66e-06 (0.000)	4.47e-06 (0.000)	0.00001 (0.000)	0.00001 (0.000)	0.00001 (0.000)	0.00001 (0.000)
Gender		1.10*** (0.153)		3.08*** (0.156)		4.16*** (0.197)		4.16*** (0.200)
Age		-0.38*** (0.015)		-0.53*** (0.018)		-0.24*** (0.012)		-0.31*** (0.017)
Student has taken the exit exam before		0.24 (0.244)		0.45* (0.266)		0.04 (0.196)		0.17 (0.327)
Constant	47.70*** (0.149)	53.88*** (0.335)	47.22*** (0.326)	55.04*** (0.453)	47.36*** (0.219)	51.42*** (0.330)	46.37*** (0.311)	49.89*** (0.505)
First Stage F-Test	4600	4300	4600	4300	217.73	217.71	217.73	217.71
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Hansen's J Statistic	2.633	2.609	2.282	2.243	2.425	2.700	1.182	1.526
P-value	0.1047	0.1062	0.1309	0.1342	0.1194	0.1003	0.2769	0.2624
Observations	52083	52082	52082	52081	51706	51706	51706	51706

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. Significance

levels: *** p<0.01, ** p<0.05, * p<0.1

Table 22. Instrumental Variables Approach clustering at the school level School Characteristics Controls				
	Language 2011	Mathematics 2011	Language 2012	Mathematics 2012
	(1)	(1)	(1)	(1)
Displaced	-2.32** (0.954)	-3.42*** (1.251)	-2.07* (1.108)	-3.66** (1.583)
Expulsion Municipality	6.69e-09 (1.54e-08)	5.47e-06 (0.00002)	0.00001 (0.00002)	8.23e-06 (0.00003)
Classes are at night	-5.65*** (0.249)	-8.15*** (0.334)	-4.21*** (0.178)	-5.30*** (0.277)
Technical/vocational school	1.01*** (0.387)	0.96** (0.375)	0.53* (0.288)	1.00** (0.429)
Teachers with teaching diploma per student	0.23 (6.137)	2.42 (6.798)	6.68 (14.451)	26.30 (18.578)
Constant	48.05*** (0.160)	47.69*** (0.323)	47.63*** (0.222)	46.66*** (0.328)
First Stage F-Test Prob > F	3700 0.000	3400 0.000	189.59 0.000	189.59 0.000
Hansen's J Statistic P-value	2.055 0.1517	2.091 0.1482	2.766 0.1000	1.392 0.2381
Observations	50382	50439	50177	50177

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 23. Instrumental Variables Approach clustering at the school level - Household and Parental Characteristics Controls				
	Language 2011	Mathematics 2011	Language 2012	Mathematics 2012
	(1)	(1)	(1)	(1)
Displaced	-1.07 (0.959)	-1.91 (1.273)	-2.05* (1.114)	-3.48** (1.468)
Expulsion Municipality	-0.00001 (0.00004)	2.37e-06 (0.00002)	0.00003 (0.00002)	0.00003 (0.00002)
Low income (2 to 3 Colombian min.wages)	1.21*** (0.134)	1.29*** (0.143)	0.74*** (0.088)	1.39*** (0.141)
Mid income (3 to 9 Colombian min.wages)	2.34*** (0.266)	2.78*** (0.300)	1.65*** (0.161)	3.40*** (0.310)
High income (10+ Colombian min.wages)	1.53** (0.697)	1.44* (0.762)	2.05*** (0.612)	3.76*** (1.014)
Mother – low level of education (up to elementary school)	2.10*** (0.297)	2.06*** (0.360)	0.66*** (0.184)	0.55** (0.263)
Mother – mid level of education (high school and/or technical studies)	3.62*** (0.299)	3.95*** (0.373)	1.66*** (0.184)	2.09*** (0.271)
Mother – high level of education (university and/or graduate school)	5.39*** (0.350)	6.07*** (0.426)	3.17*** (0.231)	4.26*** (0.374)
Father – low level of education (up to elementary school)	-1.33*** (0.185)	-0.60*** (0.193)	-0.81*** (0.126)	-0.87*** (0.186)
Father – mid level of education (high school and/or technical studies)	0.10 (0.180)	0.74*** (0.190)	0.11 (0.125)	0.21 (0.177)
Father – high level of education (university and/or graduate school)	2.74*** (0.261)	3.31*** (0.302)	1.63*** (0.212)	2.73*** (0.294)
# of People living in the household	-0.15*** (0.024)	-0.16*** (0.027)	-0.11*** (0.015)	-0.10*** (0.024)
Household owns a computer	1.22*** (0.126)	1.32*** (0.146)	0.30*** (0.028)	0.39*** (0.049)
Household has internet service	0.38*** (0.124)	0.65*** (0.137)	0.28*** (0.086)	0.23* (0.138)
Household floor material – cement	2.09*** (0.396)	1.73*** (0.344)	1.20*** (0.276)	1.50*** (0.381)
Household floor material – low quality tiles and/or wood	2.66*** (0.438)	2.22*** (0.392)	1.63*** (0.302)	1.76*** (0.431)
Household floor material – high quality tiles and/or carpet	3.02*** (0.408)	2.44*** (0.350)	1.76*** (0.277)	2.11*** (0.384)
Father occupation –	-1.16***	-1.29***	-0.77***	-0.94***

unemployed	(0.134)	(0.157)	(0.095)	(0.146)
Father occupation – qualified independent	-2.46*** (0.233)	-3.12*** (0.272)	-1.75*** (0.161)	-2.37*** (0.292)
Father occupation – not qualified independent	0.22 (0.219)	-0.011 (0.240)	0.08 (0.139)	0.07 (0.253)
Father occupation – qualified dependent	0.08 (0.151)	-1.91 (0.159)	-0.03 (0.100)	-0.12 (0.158)
Father occupation – not qualified dependent	-0.40*** (0.133)	-0.72*** (0.150)	-0.36*** (0.093)	-0.53*** (0.147)
Constant	41.11*** (0.507)	40.14*** (0.609)	43.61*** (0.399)	41.07*** (0.567)
First Stage F-Test Prob > F	17000 0.000	17000 0.000	217.77 0.000	217.77 0.000
Hansen's J Statistic P-value	2.107 0.1467	1.742 0.1869	1.632 0.2015	0.725 0.3944
Observations	52082	52081	51702	51702

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 24. Instrumental Variables Approach clustering at the school level Complete Specification				
	Language 2011	Mathematics 2011	Language 2012	Mathematics 2012
	(1)	(1)	(1)	(1)
Displaced	-1.34 (0.954)	-2.29** (1.159)	-2.15* (1.127)	-1.85* (1.004)
Expulsion Municipality	-2.04e-06 (0.00004)	3.73e-06 (0.00003)	0.00002 (0.00002)	0.00002 (0.00003)
Gender	0.87*** (0.112)	2.86*** (0.115)	0.05 (0.080)	3.98*** (0.139)
Age	-0.18*** (0.016)	-0.26*** (0.020)	-0.07*** (0.012)	-0.12*** (0.019)
Student has taken the exit exam before	-0.60*** (0.207)	-0.41* (0.221)	-0.20 (0.171)	-0.19 (0.310)
Low income (2 to 3 Colombian min.wages)	1.12*** (0.130)	1.10*** (0.134)	0.69*** (0.087)	1.28*** (0.135)
Mid income (3 to 9 Colombian min.wages)	2.26*** (0.237)	2.48*** (0.235)	1.53*** (0.130)	3.03*** (0.229)
High income (10+ Colombian min.wages)	1.75*** (0.658)	1.31* (0.698)	1.70*** (0.553)	3.40*** (0.762)
# of People living in the household	-0.17*** (0.024)	-0.19*** (0.025)	-0.13*** (0.015)	-0.10*** (0.024)
Household owns a computer	0.92*** (0.123)	0.84*** (0.138)	0.23*** (0.027)	0.27*** (0.048)
Household has internet service	0.35*** (0.120)	0.60*** (0.131)	0.24*** (0.085)	0.16 (0.134)
Household floor material – cement	1.84*** (0.375)	1.13*** (0.363)	1.06*** (0.269)	1.20*** (0.400)
Household floor material – low quality tiles and/or wood	2.40*** (0.410)	1.65*** (0.396)	1.46*** (0.305)	1.56*** (0.466)
Household floor material – high quality tiles and/or carpet	2.81*** (0.385)	1.95*** (0.365)	1.60*** (0.270)	1.95*** (0.403)
Father – low level of education (up to elementary school)	-1.33*** (0.184)	-0.53*** (0.191)	-0.79*** (0.125)	-0.73*** (0.181)
Father – mid level of education (high school and/or technical studies)	-0.15 (0.179)	0.40** (0.189)	-0.02 (0.126)	0.03 (0.174)
Father – high level of	2.46***	2.89***	1.59***	2.52***

education (university and/or graduate school)	(0.253)	(0.278)	(0.184)	(0.262)
Mother – low level of education (up to elementary school)	1.51*** (0.292)	1.14*** (0.290)	0.24 (0.177)	0.28 (0.270)
Mother – mid level of education (high school and/or technical studies)	2.72*** (0.289)	2.48*** (0.289)	1.03*** (0.174)	1.50*** (0.272)
Mother – high level of education (university and/or graduate school)	4.42*** (0.336)	4.43*** (0.339)	2.46*** (0.213)	3.62*** (0.324)
Father occupation – unemployed	-1.04*** (0.136)	-1.09*** (0.147)	-0.67*** (0.089)	-0.74*** (0.133)
Father occupation – qualified independent	-2.49*** (0.237)	-3.28*** (0.255)	-1.72*** (0.157)	-2.40*** (0.281)
Father occupation – not qualified independent	0.16 (0.219)	-0.17 (0.234)	-0.04 (0.135)	-0.19 (0.231)
Father occupation – qualified dependent	0.08 (0.153)	-0.25 (0.151)	-0.02 (0.100)	-0.17 (0.152)
Father occupation – not qualified dependent	-0.39*** (0.129)	-0.74*** (0.139)	-0.35*** (0.090)	-0.54*** (0.143)
Classes are at night	-2.44*** (0.241)	-4.20*** (0.316)	-2.68*** (0.166)	-2.82*** (0.290)
Technical/vocational school	0.65** (0.289)	0.62** (0.285)	0.29 (0.210)	0.75** (0.341)
Total teachers per student	-0.25 (0.217)	-0.18 (0.190)	-0.25** (0.102)	-0.36 (0.242)
Constant	45.63*** (0.601)	46.25*** (0.646)	46.23*** (0.426)	42.70*** (0.677)
First Stage F-Test Prob > F	1500 0.000	1500 0.000	189.23 0.000	189.23 0.000
Hansen's J Statistic P-value	1.413 0.2345	1.286 0.2567	2.191 0.1388	0.988 0.3203
Observations	50381	50438	50173	50173

Source: author's calculations using the databases from ICFES and the Secretariat of Education of Bogotá. Significance

levels: *** p<0.01, ** p<0.05, * p<0.1

Appendix 4. Cluster Robust Inference

Following Cameron and Miller (2013), consider an ordinary least squares (OLS) model with no intercept, and with a single regressor that is nonstochastic⁴⁵. Let $y_i = \beta x_i + u_i$, $i = 1, \dots, N$, where x_i is non-stochastic and $E[u_i] = 0$. The OLS estimator $\hat{\beta} = \sum_i x_i y_i / \sum_i x_i^2$ can be re-written as $\hat{\beta} - \beta = \sum_i x_i u_i / \sum_i x_i^2$. Therefore,

$$V[\hat{\beta}] = E[(\hat{\beta} - \beta)^2] = V[\sum_i x_i u_i] / (\sum_i x_i^2)^2 \quad (1)$$

If errors are uncorrelated over i , then $V[\sum_i x_i u_i] = \sum_i V[x_i u_i] = \sum_i x_i^2 V[u_i]$.

For the simple case of homoscedastic errors, we have that $V[u_i] = \sigma^2$, and therefore, equation (1) can be re-written as $V[\hat{\beta}] = \sigma^2 / \sum_i x_i^2$. For the case of heteroscedastic errors, we have that equation (1) can be re-written as $V_{HET}[\hat{\beta}] = (\sum_i x_i^2 E[u_i^2]) / (\sum_i x_i^2)^2$. This is so because $V[u_i] = E[u_i^2]$, since $E[u_i] = 0$.

Although implementation of the above seems to require consistent estimates of each of the N error variances $E[u_i^2]$, White (1980), cited by Cameron and Miller (2013), noted that it is only necessary to have an estimate of the scalar $\sum_i x_i^2 E[u_i^2]$. Therefore, it is possible to use $\sum_i x_i^2 \hat{u}_i^2$, where $\hat{u}_i = y_i - \hat{\beta} x_i$ is the OLS residual if $N \rightarrow \infty$. The estimated variance is therefore given by:

$$\hat{V}_{HET}[\hat{\beta}] = (\sum_i x_i^2 \hat{u}_i^2) / (\sum_i x_i^2)^2 \quad (2)$$

The standard error for $\hat{\beta}$ is known as a heteroscedastic-robust standard error.

⁴⁵ The results presented above can be extended to a multiple regression set-up with stochastic regressors.

WORKS CITED

- (CERAC), Centro de Recursos para el Análisis del Conflicto (2013). *Conflict data*.
- (DPS), Departamento para la Prosperidad Social (2014). *Unidad de Atención a Víctimas – Registro Único de Población Desplazada (RUPD)*. Available in: <http://www.dps.gov.co/contenido/contenido.aspx?catID=295&conID=1953>. Date Accessed: July 23, 2014.
- (ICFES), Instituto Colombiano para la Evaluación de la Educación (2013). FTP Databases from Instituto Colombiano para la Evaluación de la Educación (ICFES).
- Akresh, R. (2008). Armed conflict and schooling: Evidence from the 1994 Rwandan genocide (Vol. 3516). *World Bank Publications*.
- Albuja, S., Anwar, A., Birkeland, N., Caterina, M., Charron, G., Dolores, R., Grayson, C. (2013). Global Overview 2012: People internally displaced by conflict and violence. *Internally Displacement Monitoring Centre (iDMC) and Norwegian Refugee Council (NRC), April*.
- Bundervoet, T. (2012). War, Health, and Educational Attainment: A Panel of Children during Burundi's Civil War: Households in Conflict Network.
- Calderón, V., and Ibáñez, A. M. (2009). *Labor market effects of migration-related supply shocks: Evidence from internally displaced populations in Colombia*: CEDE, Centro de Estudios sobre Desarrollo Económico, Facultad de Economía, Universidad de los Andes.
- Cameron, A. C., and Miller, D. L. (2013). A Practitioner's Guide to Cluster-Robust Inference. *Forthcoming in Journal of Human Resources*.
- Carrillo, A. C. (2009). Internal displacement in Colombia: humanitarian, economic and social consequences in urban settings and current challenges. *International Review of the Red Cross*, 91(875), 527-546.
- Dabalén, A., and Paul, S. (2012). Estimating the causal effects of conflict on education in Cote d'Ivoire. *World Bank Policy Research Working Paper*(6077).
- Secretaría de Educación del Distrito (2012). Bases de Datos Colegios. Bogotá - Colombia.

- Ferris, E., and Winthrop, R. (2010). Education and Displacement: Assessing Conditions for Refugees and Internally Displaced Persons affected by Conflict. *Background paper for the EFA Global Monitoring Report 2011. The hidden crisis: Armed conflict and education.*
- Fiala, N. (2009). The consequences of forced displacement in northern Uganda: Households in Conflict Network.
- Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources*, 14(3), 351-388.
- Hoeffler, A., and Reynal-Querol, M. (2003). Measuring the costs of conflict. Washington, DC: World Bank.
- Hudson, R. (2010). *Colombia: A Country Study*. Washington D.C.: Library of Congress.
- Ibáñez, A. M., and Moya, A. (2010a). Vulnerability of victims of civil conflicts: empirical evidence for the displaced population in Colombia. *World Development*, 38(4), 647-663.
- Ibáñez, A. M., and Moya, A. (2010b). Do conflicts create poverty traps? Asset losses and recovery for displaced households in Colombia *The Economics of Crime: Lessons for and from Latin America* (pp. 137-172): University of Chicago Press.
- Ibáñez, A. M., and Velez, C. E. (2008). Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia. *World Development*, 36(4), 659-676. doi: 10.1016/j.worlddev.2007.04.013
- Kirchhoff, S., and Ibáñez, A. M. (2001). Displacement due to violence in Columbia: Determinants and consequences at the household level: ZEF Discussion Papers on Development Policy.
- Kondylis, F. (2008). Agricultural outputs and conflict displacement: Evidence from a policy intervention in Rwanda. *Economic Development and Cultural Change*, 57(1), 31-66.
- Kondylis, F. (2010). Conflict displacement and labor market outcomes in post-war Bosnia and Herzegovina. *Journal of Development Economics*, 93(2), 235-248.

- Lai, B. a. C. T. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, 44(3), 277-292. doi: 10.1177/0022343307076631
- LeGrand, C. C. (2001). The Colombian crisis in historical perspective. *Canadian Journal of Latin American and Caribbean Studies*, 28(55-56), 165-209.
- Mels, C., Derluyn, I., Broekaert, E., and Rosseel, Y. (2010). The psychological impact of forced displacement and related risk factors on Eastern Congolese adolescents affected by war. *Journal of child psychology and psychiatry*, 51(10), 1096-1104.
- Merrouche, O. (2006). The human capital cost of landmine contamination in Cambodia: Households in Conflict Network.
- Moya, A. (2012). *Violence, emotional distress and induced changes in risk aversion among the displaced population in Colombia*. Paper presented at the Pacific Development Economics Conference, University of California, Davis.
- Rodríguez, C., and Sánchez, F. (2010). Books and guns: the quality of schools in conflict zones. *Documento CEDE*(38).
- Rodriguez, C., and Sanchez, F. (2012). Armed conflict exposure, human capital investments, and child labor: Evidence from Colombia. *Defence and peace economics*, 23(2), 161-184.
- Ruiz, I., and Vargas-Silva, C. (2013). The economics of forced migration. *The Journal of Development Studies*, 49(6), 772-784.
- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, 95(2), 186-200. doi: 10.1016/j.jdeveco.2010.05.002
- Todd, P. E., and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485), F3-F33.

HIGHER EDUCATION: ASSESSING THE EFFECT OF THE CIVIL CONFLICT AMONGST THOSE INDIVIDUALS WHO ATTEND UNIVERSITY

Getting a better understanding of the possible effects of conflict on human capital is particularly important for the effective implementation of public policy. People who live in conflict zones are not only affected by the evident and most often documented negative repercussions of warfare (i.e. death, physical injuries, destruction of private property). Civilians could also be affected negatively because of the effects of civil war on income, health and education. According to Quinn et al. (2007), the most vulnerable people (poverty, geographical location, lack of access to public services) living in countries with a civil conflict are more likely to be affected by warfare. In that sense, civil conflict can help perpetuate inequality and poverty. A better understanding of the negative repercussions of non-international armed conflicts can help in the formulation of public policy that specifically targets those individuals who are more affected.

A number of efforts have been made recently to study the relationship between years of education, school dropouts and civil conflict⁴⁶. However, these efforts have not examined the relationship between civil conflict and the results of students in standardized examinations. This chapter seeks to contribute to the literature that studies civil conflict and education, and also open a new research agenda to study the relationship between civil conflict and performance in standardized examinations. More specifically, the objective of this chapter is to estimate the effect of civil conflict on student achievement gains in standardized examinations. We are going to study the possibility of differential achievement gains amongst students coming from conflict zones and non conflict zones by quantitatively analyzing the differences in high school exit examination scores and university exit examination scores. To the best of my knowledge, this is the first attempt to study the relationship between civil conflict and academic achievement measured by cognitive examinations at both high school and university levels. More specifically, by considering students who have been exposed to a conflict environment and students who have not been exposed to a conflict environment, we want to find out whether Colombian students affected by conflict have

**** The contents of this chapter correspond to the fourth publishable paper of this PhD dissertation.

⁴⁶ Amongst those contributions we can highlight the work of Swee (2008), Akresh and de Walque (2008), Merrouche (2006), Lai and Thyne (2007), Shemyakina (2011), Debalen and Paul (2012), Bundervoet (2012), and Bellows and Miguel (2006).

differential gains or losses in performance in comparison to those students who have not been affected, using the results from the Colombian high school exit examination (Saber11) and the Colombian college exit examination (SaberPro). To tackle this research question we are going to use the theoretical framework employed in educational value added models. Educational value added models are models that use student examination results at different points of time in order to measure their learning gain. Following this theoretical backbone, we are going to apply a difference in differences estimation strategy (equivalent to a value added specification with individual fixed effects) in order to quantify the student's learning gains using information at two points in time.

As Cunha and Miller (2014) have highlighted, in recent years we have seen a mounting pressure on colleges and universities to quantify the value that these institutions are adding to their students. Such pressure, related to the increasing interest in outcomes-based measures, has generated a growing interest in value added measures of academic achievement that take into consideration the fact that students have different academic backgrounds (Cunha and Miller 2014). The objective of a value-added model is to unravel the effect of institutions from other factors (e.g. individual characteristics, parental characteristics) that also contribute to student achievement. In order to do so, a peculiarity of value added models is that they control for a student's incoming achievement level. As Domingue (2012) highlights, the inclusion of an initial measure of achievement is important to ensure that an institution is not held accountable for preexisting differences among students.

The construction of a value added model is particularly important in studying higher education outcomes in a country like Colombia where the background of educational institutions and the students attending are very heterogeneous. In using a value added model, we are also trying to take into account differences in the backgrounds of students that attend the different universities offering higher education programs. Those differences are quantifiable in the Colombian case using the results of the High School exit examination (Saber11) and the individual, parental and socioeconomical characteristics reported by the students who take the examinations.

Higher education exit examinations have not been traditionally administered around the world. It is probable that the lack of standardized data has limited educational research at this level of instruction. The availability of a rich database of high school and university standardized exit examinations results and a set of student characteristics, maintained by the Colombian Institute for Educational Evaluation (ICFES), opens new possibilities for the construction of a value added model. Specifically, the information that is available to researchers offers a unique

opportunity for the construction of a value added model as the same students take the High School exit examination (Saber11) and the University exit examination (SaberPro). As mentioned previously, a value-added model plays a key role in disentangling the effect of institutions from other factors (e.g. individual characteristics, parental characteristics) that also contribute to student achievement. In that sense, the construction of a value added model allows for a better analysis of the relationship between conflict and education.

This chapter contributes to the existing literature that studies the impact of civil conflict on education, including Rodríguez and Sánchez (2010), Akresh and de Walque (2008), Lai and Thyne (2007), Shemyakina (2011), Dabalen and Paul (2012), Bundervoet (2012) and Merrouche (2006)⁴⁷. This chapter also contributes to the literature that use standardized test score to estimate value added models in higher education. Amongst the research that has contributed to this literature we can find the work of Klein et al. (2005), Liu (2011a and 2011b), Saavedra and Saavedra (2011), Cunha and Miller (2014), Winters et al. (2012), Plecki et al. (2012), Kim and Lalancette (2013), and Balcazar and Ñopo (2014). In terms of methodology, this chapter also contributes to the literature that uses program evaluation methodologies to study the relationship between conflict and education. The contributions to this literature include the work of Akbulut (2009), Akresh and deWalque (2008), Galdo (2010), Kecmanovic (2012), Shemyakina (2007) and Verwimp and Van Bavel (2011).

The rest of this chapter is organized as follows: section 2 presents a brief description of the civil conflict in Colombia; section 3 presents an overview of Higher Education in Colombia; section 4 reviews the existing literature; section 5 presents a conceptual background; section 6 provides a historical overview of the exit examinations in Colombia; section 7 presents the dataset and the variables that are going to be employed in this empirical exercise; section 8 describes the individuals under analysis; section 9 introduces the concept of selection bias; section 10 presents the methodology/identification strategy used in this chapter; section 11 reports and analyzes the main results obtained using a difference in differences estimation strategy; section 12 presents the discussion of the results and section 13 presents concluding remarks.

2. The Civil Conflict in Colombia

The Colombian conflict is considered to be a long duration conflict with variable intensity. In fact, the origins of the Colombian civil conflict can be traced back to

⁴⁷ This literature was reviewed in detail in Chapter 4.

the time of independence. The political enmity between the two political parties that emerged after independence, the conservative party and the liberal party, has marked Colombian history ever since. In April 1948, the assassination of the Liberal leader and presidential candidate Jorge Eliecer Gaitán led to a major violent outburst now known as the times of the Violence (*La Violencia*). During *La Violencia*, which went from 1946 to 1965, political tensions between the leaders of the Liberal and Conservative parties led to clashes in the countryside between peasants supporting the two sides (Legrand 2001). Armed groups of peasants supporting the Conservative party and supporting the Liberal party fought each other affecting principally the rural population and the inhabitants of small cities. In an effort to ease the confrontation between the supporters of the two parties, the Conservatives and the Liberals came together in 1958 to sign the National Front (Frente Nacional) pact establishing that they would alternate the presidency of the country (Legrand 2001). Although elections were still held everyone knew who would win: a Liberal candidate, then a Conservative candidate, and so on. This agreement, which lasted until 1974, helped to greatly reduce the violence between the supporters of the two parties in the early sixties. However, the exclusion of dissident political forces from the National Front contributed to the emergence of guerrilla groups in the mid-1960s, who were also looking for participatory power. The Cuban revolution and their successful guerrilla tactics inspired these new guerrilla movements. In 1964, the Fuerzas Armadas Revolucionarias de Colombia (FARC), which later on became the largest guerrilla group in Colombia, was founded, and in 1965, the Ejército de Liberación Nacional (ELN) and the Ejército de Liberación Popular (EPL) began operations.

Although guerrilla war started in Colombia in the 1960s with the emergence of the left wing guerrilla groups, Ibañez and Moya (2010) highlight that this war intensified during the late seventies and early eighties, when these illegal groups got involved in the illegal production of marijuana and cocaine. In fact, these two authors point out that illicit drug trade has provided massive financial resources to rebel groups that have helped sustain the conflict ever since (Ibañez and Moya 2010).

During the first half of the nineties the guerrillas expanded their territorial presence throughout the country, but the levels of violence did not increase (Restrepo et al. 2006). It was only in the year 1996 that the dynamic of conflict started to change significantly as an important increment in the intensity of conflict, and also in the frequency of combats and attacks became evident. By the year 1998, the FARC had gained military advantage over the governmental forces. In fact, Restrepo et al. (2006) called this period the *escalation of violence period* given the increase in the intensity of the conflict and the frequency of the combats and the attacks. According to these authors, the *escalation period* took place between the

years 1996 and 2002, year in which the armed conflict reached the highest intensity. The recrudescence of violence in Colombia in the period between the years 1996 and 2000 could be explained by a number of factors (Restrepo et al. 2006). Some of these factors include: the unification of the majority of paramilitary groups into one organization (Autodefensas Unidas de Colombia), the implementation of a terror strategy against civilians, and the modernization of the Colombian military forces through the implementation of new technologies in order to have a better response against the attacks of illegal groups. During *the escalation period*, the illegal groups implemented a strategy that focused in attacking civilians as an instrument to terrorize the society. In fact, the highest number of civilian fatalities during the conflict (global maximum) was reported in the year 2001.

It is important to note that in the year 1998, the Colombian government established a demilitarized zone (zona de despeje) in the southern part of Colombia in order to start peace talks with the FARC. The demilitarized zone was maintained until the year 2002 when the peace talks came to an end. As mentioned above, during those years there was an increase in the intensity of the conflict and in the number of municipalities registering conflict events. Nevertheless, the governmental forces always had a greater military power than the FARC during those years (Fundación Ideas para la Paz 2014).

In the period that goes from the years 2003 to 2008, the Colombian government introduced several changes in order to deal with the escalation of violence and the problems associated with it. In fact, Restrepo et al. (2006) have called this period *the reorganization period*. Amongst the changes that were introduced it is important to highlight the professionalization and expansion of the Colombian military forces, which led to important changes in the strategic scenario of the anti-guerrilla war (Restrepo et al. 2006). During this time period not only did the government introduce important reforms, but also the illegal military groups were forced to go through a re-organization period in order to react to the increasing pressure from the governmental military forces (Restrepo et al. 2006). In effect, the organizational and operative changes introduced by the governmental military forces led to a substantial difference in the warfare capacity between the governmental forces and the insurgent forces (Restrepo et al. 2006). The Democratic Security Policy (Política de Seguridad Democrática) established by ex-President Uribe, as well as the implementation of the Patriot Plan (Plan Patriota) in the year 2004, the Liberty Plan I and II in the years 2003 and 2005, amongst other operations, were vital in the re-organization process that led to the strengthening of the warfare capacity of the governmental forces.

Between the years 2003 and 2005, the number of conflict events remained relatively

unchanged, but from mid-2005 to mid-2006 the number of events decreased significantly. After that year, conflict events started to increase and in 2008 the highest number of conflict events was registered in the series (Restrepo et al. 2006). Even though, the number of attacks registered during this period is the highest in the series, the intensity of the conflict decreased. This decrease could be explained by two factors. First, the number of civilians dying because of the conflict decreased starting in the year 2001 due to a reduction in the number of victims caused by paramilitary attacks. In fact, the demobilization of a growing number of paramilitaries from the year 2003 to the year 2007 contributed to the expansion of governmental control over areas of the Colombian territory, and therefore helped in the improvement of overall security conditions (DeShazo et al. 2007). It should be noted that the paramilitary groups were responsible for many of the civilian casualties in the Colombian conflict during the nineties and beginning of the 21st century (Restrepo et al. 2006). Second, starting in the year 2003 there was a decline in the number of casualties of combatants. Such reduction can be explained by the guerrilla's operational re-organization, given their weaker position in comparison to the governmental forces, which forced them to move to remote areas in the highlands and the jungle (Restrepo et al. 2006).

As Restrepo et al. (2006) highlight, even though there was a reduction in the victimization levels, the number of victims was still very high, reaching levels above those registered in the mid-nineties. There are two aspects that could explain the persistence of high levels of violence associated to the civil conflict during this time period. First, the neo-paramilitary forces, which emerged after the demobilization of the paramilitary forces, kept the paramilitary strategies that mainly targeted civilians (Restrepo et al. 2006). Second, the guerrilla war against the FARC was still causing many civilian casualties (Restrepo et al. 2006).

The Democratic Security Policy came to an end in August 2010. The number of military actions carried out by the FARC increased at this time following the establishment of the FARC's Reborn Plan (Plan Renacer). Nevertheless, the military power of the governmental forces was still higher than the military power of the FARC (Fundación Ideas para la Paz 2014). The Reborn Plan implemented by the FARC involved the use of explosive devices, sniper attacks, and attacks against economic infrastructure (Fundación Ideas para la Paz 2014). It should be noted that this strategy did not involve a significant military effort or a significant mobilization of personnel. This could potentially be an indicator of the weakened military capacity of the FARC (Fundación Ideas para la Paz 2014).

It should be highlighted that the Colombian government has made important efforts to protect the rights of those individuals who have been affected directly or indirectly by the conflict. Law 975 of 2005 promotes a number of instruments to

guarantee the rights of the victims to justice, truth, and reparation. Additionally, Law 1448 of 2011, which is known as the Law of Victims and Land Restitution, has established judicial, administrative, social and economic actions to assist the needs of the victims of the armed conflict.

Even though in the last ten years we have seen a significant improvement in the levels of security and protection of civil rights, the Colombian government still has an important challenge: the construction of peace in the entire territory. As the Colombian National Development Plan highlights, the persistence of the armed conflict in Colombia over many decades has been an important obstacle for the development of the country, not only in terms of economic development but also in terms of social and institutional development (DNP 2014). Hoping to meet the peace challenge, the Colombian government started peace talks with the FARC in Cuba in November 2012.

3. Higher Education in Colombia

Higher education became a priority for policy makers in Colombia back in the 1930s as the country's development needs, and more specifically the fast passed urbanization process, demanded higher levels of education. Nevertheless, as Melo et al. (2014) highlight, these efforts did not lead to significant changes in the education coverage rate. According to these researchers, the coverage rate went from 3.9% in 1970 to 8.9% in 1980, and in 1990 it went up to 13.4% (Melo et al. 2014).

However, in recent years the gross coverage rate has increased at a faster rate, going from 24% in the year 2000 to 42.4% in 2012 (Melo et al. 2014). In fact, the enrollment rate increased almost four times from the year 1990 to the year 2012, going from 487448 students to 1841282 students (Melo et al. 2014). It should also be noted that the total enrollment in undergraduate programs has gone from 582672 students in the year 2000 to 913538 in the year 2012 (Melo et al. 2014). Amongst public institutions, the enrollment went from 234210 students in the year 2000 to 480155 students in the year 2012 (Melo et al. 2014).

From an international perspective, the higher education coverage rate is still very low even though enrollment has increased significantly in the last two decades. As Melo et al. (2014) have highlighted, Colombia is well behind in terms of coverage rate of higher education in comparison to both developed and developing countries. Specifically, the country lags behind other Latin American countries like Argentina, Chile, Cuba, Puerto Rico and Uruguay, where the coverage rate exceeds 60% (Melo et al. 2014).

Higher Education in Colombia is defined and regulated by Colombian law. Law 30 of 1992 and Law 115 of 1994 mandate the general standards to regulate higher education. According to these Acts, the norms to regulate this service should be grounded on the right that every individual should have to receive an education. Law 30 of 1992 establishes that higher education is a public service that can be provided by the State and/or the private sector. According to Law 30 of 1992, there are different types of institutions that can offer a higher education program, including: technical professional institutes, technological institutes, and universities. This law also establishes that the quality of the service that is being provided should be monitored by a national system of quality assurance. The State is in charge of guaranteeing the high quality of education through the inspection and supervision of the educational practices at each institution. In order to do so, the National government established a System of Quality Assurance, which is run by the National Ministry of Education (MEN) and the Higher Education National Council (CESU). Operationally, there are two commissions in charge of evaluating the quality of higher education: the National Commission for Quality Assurance in Higher Education (CONACES⁴⁸) and the National Accreditation Commission (CNA⁴⁹). CONACES, establishes whether or not a program complies with the minimum quality standards, and reports these findings to the Ministry of Education. A positive report from the CONACES is necessary in order to obtain the Qualified Registry. The Qualified Registry is required by law for all higher education programs in order to legally operate. In order to improve the quality of higher education, the Colombian authorities also established a voluntary accreditation system, which is managed by the CNA. The National System of Accreditation is in charge of publically recognizing the high quality of those higher education institutions and higher education programs that have gone through a process of auto-evaluation, and demonstrated outstanding practices.

As mentioned previously, Law 30 of 1992 establishes that there are four different types of institutions that can offer a higher education program in Colombia. The first category includes universities that offer undergraduate and graduate programs (“Specialization”, Masters, and Doctoral programs), and also contribute in scientific research. The second category encompasses “university institutions” that offer undergraduate and “specialization” programs. The “specialization” programs are career-related programs that provide a level of qualification above an undergraduate degree but below a master’s degree). The third category

⁴⁸ The letters stand for Comisión Nacional de Aseguramiento de la Calidad de la Educación Superior, which translates to English, National Commission for Quality Assurance in Higher Education.

⁴⁹ The letters stand for Comisión Nacional de Acreditación, which translates to English, National Accreditation Commission.

includes technological institutions that offer programs up to the technologist level. In some cases students who are enrolled in these institutions can continue their studies in order to obtain a professional degree. The fourth category encompasses professional technical institutions that offer training for a particular job or career. This last category includes those institutions that offer short courses to prepare individuals to perform a very specific job.

According to Melo et al. (2014), in the year 2012 Colombia had 288 higher education institutions. Of those 288 institutions, 81 were universities (first category), 120 were “university institutions” (second category), 50 were technological institutions (third category) and 37 were professional technical institutions (fourth category). Of the total, 61 institutions were public, 208 were private, and 19 were part of a special regime. It should also be noted that by the year 2012 only 8.1% of all academic programs had received a high quality accreditation conferred by the CAN (Melo et al. 2014). It is important to highlight that the quality of higher education institutions is very heterogeneous in Colombia.

The OECD (2012) has pointed out that some students in Colombia have a wider choice of tertiary institutions than others. Higher education institutions are not evenly distributed across Colombian Departments and Municipalities. *Table 1* shows the gross enrollment rates by Colombian Department from 2002 to 2010.

The basic requirement for entry to higher education in Colombia is a High School graduation certificate. Nevertheless, every institution is free to decide its own admission standards. According to the OECD (2012), 78% per cent of the tertiary institutions use the results of the national High School exit examination (Saber11) as their only admission standard. The remaining percentage is mainly made up of public universities⁵⁰ that have their own admissions examination. Nevertheless, Sanchez et al. (2002) have found evidence showing that admissions decisions at public universities are mostly determined by the results in the Saber11 examination. These researchers have found that there is a high correlation between the results in the Saber11 examination and being admitted at a public university (Sanchez et al. 2002).

⁵⁰ An example of this is Universidad Nacional de Colombia. This includes the main campus in Bogotá and all the other campuses in different cities in Colombia. There are other public universities that also have their own admissions examinations, including: Universidad del Magdalena, Universidad del Atlántico, Universidad de la Guajira, Universidad de Cartagena, Universidad de Antioquia, Colegio Mayor de Cundinamarca (Sanchez et al. 2002)

Table 1 - Gross Tertiary Enrollment by Department (%)									
Department	2002	2003	2004	2005	2006	2007	2008	2009	2010
Amazonas	1.5	4	5.1	4.4	6.4	6.5	6.5	12.4	13.3
Antioquia	26.6	28	29.6	31.3	33.3	33.1	35.1	39.6	40.9
Arauca	1.6	1.7	3	3.2	4.5	8.6	12.5	14	12.7
Atlántico	34	32.2	32.2	34.9	35.2	36	36.5	33.4	37.9
Bogota	55.4	55.5	59.9	61.3	66.8	63	68.3	71.7	73.7
Bolívar	13.2	17.9	18.3	18.5	18.3	22.2	24.9	21.8	28
Boyacá	21	22.5	23.1	26.3	25.7	33.7	36.5	37.4	39.7
Caldas	22.4	23.2	25	26.5	26.2	29.3	28.3	33.7	35
Caquetá	7.6	7.5	8.9	12.2	14.8	20.3	22.5	26.1	19.1
Casanare	2.6	4.5	5	8.2	9.9	18.4	26	26.1	23.8
Cauca	12.8	13.5	15.1	15.8	16.4	20.1	22.1	23.2	26.6
Cesar	10.9	11.7	12	14	15.5	19.2	21	25	21.6
Chocó	19.1	17	18.4	19.3	22	19.3	19.5	22.1	25.8
Córdoba	11.1	12.1	12.5	12.7	15.2	17.6	17.4	10.9	17
Cundinamarca	11.5	13.4	13.6	13.8	14.8	15.9	18.8	21.4	21.1
Guainía	N/A	0	3.3	4.2	9.7	17	19.4	14	11.5
Guaviare	N/A	0	1.7	3.1	7.3	11.6	13	14.2	12.8
Huila	11.5	13.7	14.4	16.2	17	21.1	23.3	26	25.7
La Guajira	13	13.2	12.8	14.3	15.3	14.6	17.7	20.8	17.5
Magdalena	6.7	7.9	9.4	11.5	13	21.5	23.1	24.6	20.5
Meta	13.2	14.2	14.1	17.9	20	24.9	26.5	25.3	24.4
Nariño	10.6	11	10.6	11.9	12.2	16.6	17.5	18.9	18.3
Norte de Santander	21.9	26.9	25.9	29	26.2	36.6	39.8	42.2	42.8
Putumayo	2.8	3.3	4.2	4.1	5.1	6.1	9.1	6.8	11.5
Quindío	22.7	25	25.3	24.6	29.6	40.6	47.8	49.4	50.4
Risaralda	17.6	21	24.2	26.6	28.7	35.3	39.4	37.1	42.2
San Andrés	18.1	7.1	9.4	7.2	12.2	18.7	19.2	17.3	25.7
Santander	31.2	32.2	34.4	36.1	36.1	39.7	44.8	38.2	48
Sucre	9.2	10.6	9.1	10.7	11.4	14.8	17.3	17.2	17
Tolima	18.1	25.8	27.6	27.9	27.9	24.2	26.5	26.5	25.6
Valle del Cauca	23.8	22.9	23.2	24.3	24.7	26.5	27.8	29.7	31.7
Vaupés	N/A	0	0.7	2.7	4.1	12	7.8	9.6	4.2
Vichada	N/A	0	0.5	2	2.7	7.6	8.3	10.9	9.9
Total (National)	24.5	25.6	27	28.4	30	31.7	34.1	35.3	37.1

Source: OECD (2012)

4. Literature Review

This chapter contributes to the existing literature that studies the impact of civil conflict on education, the literature that studies the use of value added models to understand the determinants of educational achievement, and the literature that uses program evaluation methodologies to study the relationship between conflict and education. The existing literature that is significant in the study of the relationship between conflict and education was reviewed in detail in Chapter 4 of this dissertation. Amongst the contributions that were reviewed in that chapter we can find the work of Rodríguez and Sánchez (2010), Akresh and de Walque (2008), Lai and Thyne (2007), Shemyakina (2011), Dabalen and Paul (2012), Bundervoet (2012) and Merrouche (2006). It is also imperative to review the literature that studies educational achievement through the use of a value-added approach in order to have a better understanding of the state of the art in this subject. In fact, value added models of student achievement have received a lot of attention in recent years because of the growing interest in test-based accountability. A number of researchers have contributed to this literature by analysing the different determinants of student achievement through the use of a value added model. Amongst the academic research that has contributed to this literature we can find the work of Klein et al. (2005), Liu (2011a and 2001b), Plecki et al. (2012), Winters et al. (2012), Cunha and Miller (2014), and Kim and Lalancette (2013). We can also find some contributions that have applied the value added model to study the Colombian case including the work of Saavedra and Saavedra (2011), Balcazar and Ñopo (2014), Bogoya and Bogoya (2013), Rodríguez (2014) and Isáziga-David et al. (2014). However, what is notable is that to the best of my knowledge no effort has been made to try to quantify the impact of the Colombian civil conflict on education achievement using a value added model for higher education. In terms of methodology, this chapter also contributes to the literature that uses program evaluation methodologies to study the relationship between conflict and education. The contributions to this literature include the work of Akbulut (2014), Akresh and deWalque (2008), Galdo (2010), Kecmanovic (2012), Shemyakina (2011), Dabalen and Paul (2012), Merrouche (2006), Swee (2008) and Verwimp and Van Bavel (2011).

It should be recognized that the literature review presented in this chapter concentrates on the economic literature that is currently available. Nonetheless, there are other disciplines that have also done scientific research on the relationship between civil conflict and education, including Sociology and Political Science. However, the economic literature has a stronger quantitative tradition. For our purposes, and considering that the analysis done in this chapter concentrates on the application of quantitative methods, we have decided to center our attention on the economic literature that is currently available.

In what follows, I will review the available literature that uses value added models to study student achievement and the available literature that uses program evaluation methodologies to study the relationship between conflict and education in order to contextualize the topic under study and justify the relevance of this chapter.

4.1 Value Added Models in the Literature

International Literature

A number of researchers, including Klein et al. (2005), Liu (2011a and 2001b), Plecki et al. (2012), Winters et al. (2012), Cunha and Miller (2014), and Kim and Lalancette (2013), have contributed to the literature that studies educational value added models. Educational value added models are models that use student examination results at different points of time in order to measure their learning gain. These models have been used in the literature in order to have a better understanding of the role of teachers in the students' learning process and to determine which factors are associated with the academic gain of students. This literature review first examines how value added models have been used in the literature to quantify the importance of teachers in student achievement, and then examines how the literature has studied the added value of higher education other factors affecting achievement gains.

Using data from the Washington State Office of the Superintendent of Public Instruction (OSPI) in the United States, Plecki et al. (2012) study the effects of teacher education programs in value-added scores. Plecki et al. (2012) estimate value-added measures in reading and mathematics for fifth-grade teachers, and use that information to examine whether differences in value-added measures exist by teacher experience and by the institutions where these teachers obtained their initial teacher training. To do so, the researchers apply a univariate nonclassified two-level hierarchical linear model. According to Plecki et al. (2012), this model reduces the multivariate data to a univariate outcome. The researchers find evidence suggesting that experience indeed matters to student achievement particularly in the early years of a teacher's career, and there is variation across teacher education programs when measuring the effectiveness of their graduates using value-added measures (Plecki et al. 2012).

In line with Plecki et al. (2012), Winters et al. (2012) measure the impact of observed teacher characteristics on student math and reading proficiency using a dataset that links individual teachers and students throughout Florida elementary

schools from 2000–2001 to 2003–2004 school years. These researchers account for nonrandom attrition of teachers from the classroom in a sample selection framework. Winters et al. (2012) use a modified version of the sample selection model developed by Heckman (1979) to account for teacher attrition when evaluating the relationship between teacher characteristics and the impact a teacher makes on student achievement gains. The results obtained by these researchers show that sample selection is present in the estimation of student gains, but failure to account for it does not substantially bias the estimation.

The literature has also studied the application of added value models in higher education. The objective of these models is to get a better understanding of the factors that are affecting achievement gains. Cunha and Miller (2014) develop a general methodology for measuring the value added of institutions of higher education using commonly available administrative data. The research approach proposed by Cunha and Miller (2014) combines information from different administrative sources in the state of Texas to develop value-added estimates for 30 traditional four year public colleges in that State. The researchers follow the universe of Texas college applicants from the time of application through public college and into the labor market. In specifications that do not control for selection, Cunha and Miller (2014) find large and significant differences across colleges in terms of persistence, graduation, and earnings. However, the differences decrease considerably when the researcher controls for selection (Cunha and Miller 2014).

Following the same line of analysis of Cunha and Miller (2014), Klein et al. (2005) administered several open-ended tests to 1365 students from 14 diverse colleges in the United States to investigate whether these measures were sensitive to changes in student ability over time (difference from first year students to last year students). To do so, the researchers used a regression model. The results showed that student performance on the tests was related to grades in university. Additionally, after controlling on the high school exit exam results (SAT scores) the mean scores increased consistently from first year students to last year students. These findings suggest that the results in the examinations are sensitive to student learning over time. In that sense, the results obtained by Klein et al. (2005) can be taken as evidence of value added in higher education.

In line with Cunha and Miller (2014) and Klein et al. (2005), Liu (2011b) studies how students progress in college by analyzing the differences in the performance of students in first semester and students in the last year of studies after controlling for admission scores. In order to measure the learning outcomes of general college education, Liu (2011b) uses information from 6196 students enrolled in 23 institutions who took the Measure of Academic Proficiency and Progress™ (MAPP™) test. The MAPP measures college-level skills in critical

thinking, reading, writing and mathematics. An ordinary least squares (OLS) regression model was estimated between mean SAT scores and mean MAPP scores in order to measure the student learning growth. It should be noted that Liu (2011b) carried out her analysis at the institutional level. The results obtained by this researcher show that the MAPP is able to differentiate performance between first semester students and students in the last year of studies after controlling for their admission scores. As a consequence, Liu (2011b) concludes that the score difference can be used as one indicator of student learning at an institution.

Other researchers have also contributed to the higher education added value literature by developing new methodological approaches. Liu (2011a) proposes an alternative way to quantify the added value of higher education using multilevel models. Liu (2011a) uses a two-level hierarchical linear model to calculate value-added scores. This researcher uses written communication and critical thinking examinations as the core educational outcomes under study. The first level of analysis of the hierarchical model is at the student level and the second level of analysis is at the academic institution level. The results obtained using the multilevel model are compared with the results obtained from an ordinary least squares estimation. According to the results obtained by Liu (2011a) the institutional value-added was significantly different for some of the institutions when using a multilevel model in comparison to those obtained when using ordinary least squares.

For a comprehensive review of the literature on value-added measurement in higher education refer to Kim and Lalancette (2013). These two researchers review the existing literature on value-added measurement approaches, methodologies, and challenges within the higher education contexts.

Colombian Literature

Some studies, including the work of Saavedra and Saavedra (2011), Balcazar and Ñopo (2014), Rodriguez (2014), Bogoya and Bogoya (2013), and Isáziga-David et al. (2014), have contributed to the literature that studies educational value added models in the Colombian educational context. In contrast to the international literature, there are no studies specifically studying the role of teachers in the students' learning process. All of the studies reviewed study the factors associated with the academic gain of students in higher education.

Using data from a 2009 collegiate assessment pilot study in Colombia, Saavedra and Saavedra (2011) investigate how much value college enrollment adds to students' critical thinking, problem solving and communication skills, and the role

college inputs play in developing these competencies. According to these researchers, the estimation of student gains during college (value added analysis) with cross sectional data is subject to two empirical challenges. The first challenge is associated to potential changes in cohort quality over time, and the second challenge is related to non-random attrition from college between freshman and senior year (Saavedra and Saavedra 2011). In order to account for potential bias due to differences in observable characteristics between first and last-year students, the researchers control for student characteristics including their college entry scores, income and parental education, and area of study and college fixed effects. Saavedra and Saavedra (2011) also use propensity scores to re-weight observations in order to impose identical covariate distributions for first and last-year students. According to their results, in comparison with similar first year students, students in their final year of college score about half of a standard deviation higher. These researchers also find that this effect is particularly significant amongst students from private universities. On the other hand, Saavedra and Saavedra (2011) find that variables such as the number of professors with a PhD, the percentage of fulltime faculty, and the expenditures per student are not correlated with higher gains in critical thinking, problem solving, interpersonal and communication skills.

A number of studies have used the same information that is being used in this chapter, ICFES's Saber11 and the SaberPro examination results, in order to understand the added value of higher education in the Colombian context. Balcazar and Ñopo (2014) use a panel data set that combines two standardized tests for Colombian students: the Saber11, taken at the end of senior year in high school, and the SaberPro, taken when students are near graduation from college, to test the extent to which education majors relatively improve or deteriorate their skills in comparison to students in other programs. In that sense these two researchers contribute to the literature that studies the added value of higher education. Balcazar and Ñopo (2014) analyze three different skills: quantitative reasoning, Spanish (native language) and English (foreign language). These researchers measure test score-rankings using the z-scores of the three skills evaluated in both tests and compare changes in the relative position of students in the distributions of skills (Balcazar and Ñopo 2014). In order to account for potential bias due to differences in observable characteristics between education and non-education majors, Balcazar and Ñopo (2014) use two approaches: a two-step heckit model and a non-parametric matching model. The objective of the researchers is to compare education majors and non-education majors based on the same observable characteristics. According to their results, the quantitative reasoning skills of students in education majors deteriorate in comparison to those enrolled in other majors (Balcazar and Ñopo 2014). However, the results obtained for native language and foreign language skills do not show evidence of robust

changes. Rodriguez (2014) also studies the value-added of engineer programs in Colombian universities using information from the Institute for the Evaluation of Education Quality (ICFES). To do so, the researcher uses a multilevel model that includes variables such as gender, sociocultural and socioeconomic level, previous academic performance (score in the Saber11 examination) and tuition of the university. The sample employed in this study includes 20 universities and 3000 engineering students. The results obtained by Rodriguez (2014) show that the previous test score (Saber11 score) is a significant variable to determine the performance in higher education.

Other researchers have also contributed to the Colombian higher education added value literature by applying new methodological approaches. Bogoya and Bogoya (2013) propose a mathematical model to study higher education effectiveness in Colombia from an academic added-value perspective. To solve the mathematical model and show its application, academic value added effectiveness was calculated for a sample of 10,782 Business Administration students who finished their higher education in Colombia in the year 2009. In this model, according to the researchers, two variables captured the learning level of students enrolled in a given program (the first at the beginning of the higher education program and the second at the end), a third variable captured the students' socioeconomic strata and a fourth captured academic value added effectiveness (Bogoya and Bogoya 2013). According to the results obtained by Bogoya and Bogoya (2013) controlling for context variables, an academic program is more effective and more significant for student learning than another when the first program has higher academic value added effectiveness than the second one. Isáziga-David et al. (2014) also propose the application of a new methodology to obtain an estimate of the added value of higher education. These researchers propose a methodological approach based on the percentage of maximum effectiveness on a test to estimate the added value of higher education in Colombia. According to Isáziga-David et al. (2014), the percentage of maximum effectiveness on a test is a theoretical construct that attempts to capture the maximum attainable score that a student can get in a test in a given year. The aim of these researchers is to find the difference between the students' relative performance before and after a higher education program, taking as a reference the Saber11 and SaberPro examinations. In order to calculate this difference Isáziga-David et al. (2014) use different paired sample tests, including the Student's t-test, the Wilcoxon test, and the sign test. The results obtained by these researchers show that there is a significant difference between the percentage of maximum effectiveness before and after a higher education program. According to Isáziga-David et al. (2014), these results suggest that higher education institutions through their different academic programs can generate value for students, and in that sense they can generate considerable social profit.

4.2 Program Evaluation Methodologies to Study the Relationship between Conflict and Education in the Literature

A number of researchers have contributed to the literature by applying program evaluation methodologies to study the relationship between civil conflict and educational outcomes. Amongst the academic research that has contributed to this literature we can find the work of Galdo (2010), Verwimp and Van Bavel (2011), Shemyakina (2011), Dabalen and Paul (2012), Merrouche (2006), Kecmanovic (2012), Swee (2008), Akresh and de Walque (2008) and Akbulut-Yuksel (2014).

Some authors, including Galdo (2010), Kecmanovic (2012) and Akbulut-Yuksel (2014) have studied the relationship between civil conflict and labor market outcomes using program evaluation methodologies. Galdo (2010) exploits the district-level variation in the timing and intensity of Peruvian civil war violence to study how early life exposure to civil war affects labor-market outcomes later in life. This researcher uses a differences-in-differences approach to test whether children born in districts affected by armed violence have more adverse labor-market earnings later in life. To do so, Galdo (2010) uses two sources of information: the Peruvian household survey to capture the socio-demographic and labor market earnings of civilians and the Peruvian Truth and Reconciliation Commission (TRC) database to obtain information on civil war measures. This researcher finds that the most sensitive period to early-life exposure to civil war violence is the first 36 months of life. Galdo (2010) also finds that health is the most important channel in connecting early life exposure to civil war and adult earnings rather than schooling. Also exploring the relationship between civil conflict and labor outcomes, Kecmanovic (2012) studies the effect that the war in Croatia from 1991 to 1995 had on the educational, employment and income possibilities of the 1971 male birth cohort. The 1971 birth cohort was chosen because it is very likely that an important number of men from this cohort were drafted into the armed forces. The author uses the occurrence of war and the subsequent effects as a natural experiment, and applies the differences in differences technique to find the impact of conflict on education. Kecmanovic (2012) finds lower levels of education amongst the cohort of males affected by war in Croatia compared to women and older cohorts. The author explains this finding by pointing out the importance of males in the conflict as soldiers. In line with Galdo's (2010) and Kecmanovic's (2012) work, Akbulut-Yuksel (2014) uses a differences in differences approach to provide causal evidence on the long-term consequences of large-scale physical destruction on the educational attainment, health status and labour market outcomes of German children after World War II. Akbulut-Yuksel (2014) shows that German children that were school-aged during WWII had 0.4 fewer years of education in adulthood. The author highlights that an important mechanism to explain this impact was the destruction of schools and the absence of teachers.

Other researchers, including Verwimp and Van Bavel (2011), Shemyakina (2011), Swee (2008), and Akresh and de Walque (2008) have studied the relationship between civil conflict, enrollment and the probability of completion of primary schooling using program evaluation methodologies. Verwimp and Van Bavel (2011) study the effect of exposure to violent conflict on the completion of primary schooling. The identification strategy used in this paper is based on a differences-in-differences approach. Verwimp and Van Bavel (2011) exploit the spatial and temporal variation of violent conflict in Burundi to find the effect of exposure on child schooling. These researchers compare children who were exposed to several years of conflict in their area of residence with children of the same age residing in areas who were not much affected. To do so, Verwimp and Van Bavel (2011) use a nationwide household survey that collected detailed education, migration, gender and wealth data. These researchers find that the probability of completing primary school for a child exposed to violence went down by 40 to 50% compared to a child who was not exposed to violence (Verwimp and Van Bavel 2011). Verwimp and Van Bavel (2011) conclude that the completion of primary schooling in Burundi for boys and girls is severely effected by the massacres (1993-1994) and the civil war (1995-2005). Shemyakina (2011) also studies the relationship between civil conflict and the probability of completion of primary schooling. This researcher studies the specific case of Tajikistan using a difference in differences technique to determine whether the exposure to conflict affected the probability of completion of mandatory schooling. The results show that exposure to the conflict had a significant negative effect on the school enrolment of girls, and little, or no, effect on enrolment of boys (Shemyakina 2011).

In line with Shemyakina (2011) and Verwimp and Van Bavel (2011), Swee (2008) exploits the variation in the Bosnian War (1992–1995) intensity and birth cohorts of children to identify the effects of the civil war on schooling attainment. The empirical strategy exploits the variation in birth cohorts and in war intensity represented by the number of war casualties per capita across Bosnian municipalities to adopt a difference in differences approach that takes into account unobserved pre-war differences across municipalities. The author finds that individuals in the cohorts affected by war are less likely to complete secondary schooling, if they resided in municipalities that experienced higher levels of war intensity, but there are no significant effects of war on the completion of primary schooling. Using these results, Swee (2008) argues that the Bosnian war influenced school attainment through youth soldiering and no other direct mechanisms like exodus of teachers and damaged or destroyed infrastructure. The author discusses different mechanisms through which civil war might affect schooling attainment given the availability of data on individuals' physical and mental health, war damage to housing units, and out-migration during the war. In another study,

Akresh and de Walque (2008) examine the impact of the Rwandan genocide on children's human capital investment. The authors use two cross-sectional household surveys collected before and after the Rwandan genocide in order to examine if and how genocide affect children's school enrolment and the probability of completing a particular grade (Akresh and de Walque 2008). To do so, the authors compare children in the same age group who were and were not exposed to genocide. Akresh and de Walque (2008) point out that civil war, and more specifically genocide in this particular case, are among the most destructive of social phenomena especially for children of school-going age. The authors find that children exposed to the genocide experienced a drop in educational achievement of almost one-half year of completed schooling, and are 15 percentage points less likely to complete third or fourth grade (Akresh and de Walque 2008). In that sense the results obtained by Akresh and de Walque (2008) contradict the findings of Swee (2008). The results obtained by Swee (2008) for the Bosnian case show no significant effect on the completion of primary schooling whereas Akresh's and de Walque's (2008) results for the Rwandan genocide show significant negative effects of civil conflict on completing schooling at the primary level.

Dabalen and Paul (2012) and Merrouche (2006) have also contributed to the literature by studying the relationship between civil war and years of education. Dabalen and Paul (2012) use two different identification strategies to understand this relationship. In the first strategy, the year and department of birth is used to determine an individual's exposure to war and a difference in differences approach to quantify differences in the years of education for kids that live in war affected regions. The second strategy uses victimization indicators to implement the propensity score matching technique, in an effort to try to minimize selection bias and confounding in the causal effect. The results robustly show that war has a detrimental impact on education. In line with Dabalen and Paul (2012), Merrouche (2006) uses instrumental variable regressions and a difference in differences approach to identify the effects of landmine contamination on years of schooling using data from Cambodia. This researcher found that exposure to landmines in Cambodia resulted in an average loss of 0.4 years of education. The author claims that there are three possible factors that could explain this finding: the downgrading of educated people during the regime, direct effects of landmines on the returns to education, and the destruction of physical capital and technological delay.

The existing literature that employs program evaluation techniques to study the relationship between civil conflict and educational outcomes is still limited. One of the contributions of this chapter is precisely related to the application of these techniques. To the best of my knowledge, this is the first attempt to employ a

matching and a differences-in-differences technique to try to understand and quantify the impact of civil conflict on educational added value (gains in educational achievement) measured by the differences in standardized examination results which were collected at two points in time. We are contributing in the search for new and better ways to understand the relationship between civil conflict and educational achievement.

5. Conceptual Background

This section introduces a simple model of achievement in order to rationalize the analytical relationship between educational achievement and civil conflict. The model, proposed by Todd and Wolpin (2003), sets forth the idea that as students progress through the formal education curricula, their level of achievement in cognitive exams is a function of the history of family inputs, the school inputs and innate ability. However, for our purposes, in this case the model has been extended to incorporate the role of conflict as a determinant of educational achievement. The intensity of conflict and/or terrorist attacks can be seen as an environmental factor that causes stress amongst school aged individuals, and therefore should be included as a determinant of educational achievement.

Following Todd and Wolpin's (2003) notation, let T_{ija} be a measure of achievement for student i who lives in municipality j at age a . In this model we assume that knowledge acquisition is a production process in which current and past inputs are combined with the student's genetic endowment to produce a certain level of cognitive achievement (Todd and Wolpin 2003).

In this conceptual framework we assume that inputs are given by the choices that are being made by parents, schools and the actors of the conflict (causing environmental stress). Following Todd and Wolpin's (2003) notation, the vector of inputs supplied by the parents at a given age is represented by F_{ija} , the vector of school-supplied inputs is represented by S_{ija} and the vectors of their respective input histories up to age a is given by $F_{ij}(a)$ and $S_{ij}(a)$. Additionally, C_{ija} denotes environmental factors (in this case the intensity of civil conflict events) affecting student i who lives in municipality j at age a . The student's endowed ability is denoted as μ_{ij0} , and a measurement error in test scores as ε_{ija} . Given those inputs, the production function is represented as:

$$T_{ija} = T_a(F_{ij}(a), S_{ij}(a), C_{ija}, \mu_{ij0}, \varepsilon_{ija}) \quad (1)$$

Following Todd and Wolpin's (2003) notation, the a subscript on $T_a(.)$ is used in order to denote that the impact of inputs and of the genetic endowment depends on the age of the student.

Todd and Wolpin (2003) propose a specification to deal with the problem of missing data for past inputs known as the "Value-Added Approach" for estimating achievement production functions. This "Value-Added Approach" assumes that a previous test score is a sufficient statistic for the missing historical inputs (Todd and Wolpin 2007). In other words, it is necessary to assume that the lagged achievement measure is a sufficient statistic for unobserved input histories as well as the unobserved endowment of mental capacity (Todd and Wolpin 2003).

Following Todd and Wolpin's (2003) notation, let X denote the vector of family and school inputs and $X(a)$ denote the input histories up to age a . C_{ija} denotes the intensity of civil conflict events for student i who lives in municipality j at age a . Under this specification we are assuming that equation (3) can be written as a function only of a baseline test score and contemporaneous inputs (Todd and Wolpin 2003). Without loss of generality, let's assume that the lagged test is taken by the student at time $a - 1$. Given this notation, the value-added model would be specified by:

$$T_{ija} = T_a(X_{ija}, C_{ija}, T_{a-1}[X_{ij}(a-1), \mu_{ij0}], \eta_{ija}) \quad (2)$$

If we assume that the arguments in equation (1) are additively separable and the parameters are non-age varying, we have the following equation that can be estimated empirically:

$$T_{ija} = X_{ija}\alpha + C_{ija}\varphi + \gamma T_{ij,a-1} + \eta_{ija} \quad (3)$$

6. ICFES and the SaberPro Examination

The Colombian Institute for the Evaluation of Education (ICFES), which was established in the year 1968, is responsible for the evaluation of the outcomes at all levels of instruction, including higher education. According to Law 1324 of 2009, ICFES is a public institution with autonomous funding, governed by a board of directors appointed by the President of Colombia and the National Ministry of

Education. ICFES administers the National Basic Education Examinations (Saber5 for 5th graders and Saber9 for 9th graders) and the National State Examinations (Saber11: High School exit examination and SaberPro: Higher Education).

The Saber11 examination is the Colombian High School examination. This examination, which is usually taken by students in their last year of High School, provides information to determine admission to Higher Education. The Saber11 examination, which was originally known as the ICFES exam, was first administered in the year 1969. This examination was initially designed to assess specific areas of knowledge, which are necessary for admission to higher education academic programs. However, in the year 2000 several changes were introduced. The design of the examination changed in order to assess competencies instead of evaluating knowledge in a specific subject area. It should be highlighted that an OECD (2012) report points out that the psychometric quality of the Saber11 examination significantly improved after the major revisions that were introduced to the content and psychometric indicators of the exam.

Colombia has made an important effort to develop an examination of learning outcomes for higher education. ICFES has administered standardized examinations to evaluate the outcomes of students at the end of their undergraduate programs since the year 2004. According to the decree 1781 of 2003, the objectives of the higher education exit examination are to:

- a) Make sure that students have the sufficient skills/competences when they graduate from university.
- b) Build indicators of the value of an undergraduate education.
- c) Provide information that will enable comparisons amongst academic programs, institutions and learning methodologies, and follow its progress over time.
- d) Provide information for the construction of quality indicators for academic programs and higher education institutions. This information is meant to help in the design of policies, and aid in the decision taking process in educational matters.

The SaberPro examination, which was originally known as ECAES⁵¹, assesses individual competencies of undergraduate students during their final year of schooling. The OECD (2012) highlights that given that this examination is calibrated with the Saber11 High School exit examination, it offers a great opportunity to determine the effect of programs and learning systems on the

⁵¹ Originally Exámenes de Calidad de la Educación Superior (ECAES), which translates to Higher Education Quality Examinations.

educational outcomes of students. The calibration is particularly useful for the construction of higher education quality indicators, including value-added estimates and performance trends (OECD 2012). The examination was initially voluntary, but with the introduction of Law 1324 of 2009 it became mandatory for all students in order to graduate from an undergraduate program starting in the year 2009. Law 1324 also establishes that the test structure should be the same for a period of at least 12 years.

In the year 2009, ICFES administered 55 examinations of specific competencies for academic programs. Each of these examinations also had a generic competencies component that tested reading comprehension and English. The students of those programs of study that did not have a specific examination had to take a generic skills test that evaluated problem solving skills, critical thinking, writing and interpersonal communication (OECD 2012). In the first semester of 2010 ICFES administered 33 degree specific examinations, and in the second semester 31 examinations. Just like in the year 2009, the students of all other degree programs had to take the generic skills test. It should be noted that the results of the examinations of specific competencies could only be used to make comparisons amongst the same program of study. It was not possible to make comparisons between students from different programs and evaluate the overall performance of all students enrolled in higher education using these results.

In the second semester of 2011 important changes were introduced. The degree specific examinations were eliminated. ICFES decided to maintain the existing generic skills test, and complement it with new competencies examinations. The new competencies that were included in the examination are: critical reading, quantitative reasoning and writing skills. All the students who are enrolled in higher education academic programs have to take the generic competencies examination during their final year of studies. The application of the generic competencies examination eliminates the problem of comparability between students from different academic programs and different institutions (Melo et al. 2014).

The OECD (2012) has highlighted some characteristics of the new SaberPro examination that makes ICFES's evaluation program a leader in terms of outcomes evaluation for tertiary education:

- The SaberPro assesses all of Colombia's tertiary education population.
- The SaberPro examinations are developed and maintained following standard psychometric procedures that maintain the measurement scale and the comparability of results across time and groups.
- The SaberPro examination evaluates common competencies between

- diverse programs.
- The SaberPro examination can measure value added in tertiary education, using the new Saber11 examinations as the entry measure, because the new SABER 11 examinations measure many of the same competencies.
- The SaberPro examination concentrates on the assessment of more basic skills, not expected to fluctuate, and which are the outcome of an aggregation of educational content and processes over the course of the whole tertiary education program.

It is very important to highlight that the results of the generic competencies examinations allow for comparisons amongst the different areas of knowledge. Furthermore, amongst those generic competencies, the Saber11 and the SaberPro examinations have three components in common that allow researchers to compare the students' learning gain, i.e. added value of higher education. The three common components are Spanish (language), quantitative reasoning (mathematics) and English. The language (Spanish) component in the Saber11 examination can be compared to the reading (Spanish) component in the SaberPro. The mathematics component in the Saber11 examination can be compared to the quantitative reasoning component in the SaberPro examination. The English component in both examinations evaluates aptitude in English. For the purposes of this chapter, we are only going to use the results of the mathematics and language components.

7. Data

In this chapter we use information from two national standardized examinations: the Colombian High School exit examination (Saber11) and the Colombian Higher Education (SaberPro) exit examination. This information was obtained from the Colombian Institute for the Evaluation of Education (*Instituto Colombiano para la Evaluación de la Educación – ICFES*). We also used information from two additional sources. We obtained information about Colombian municipalities from the Colombian National Planning Department (*DNP – Departamento Nacional de Planeación*). The planning agency provided us annual information about the social and financial status of Colombian municipalities. The conflict data that we use in this chapter was obtained from the Resource Centre for Conflict Analysis (*Centro de Recursos para el Análisis del Conflicto – CERAC*).

7.1 Main Sources of Information

Colombian Exit Examinations – ICFES Saber11 and ICFES SaberPro

Databases containing the nation-wide results of the SaberPro and Saber11 examinations, as well as information about socioeconomic variables of the students who took the exam, have been made public by ICFES⁵² for research purposes. Those databases are the main source of information for this empirical chapter. The database containing the SaberPro examination results that is employed in this chapter corresponds to the second semester of the year 2011. The databases containing the Saber11 examination results that were used in this empirical exercise cover the years 2002 to 2007.

The sample that is used in this empirical exercise has nation-wide observations from students for which ICFES has information about both the results for the SaberPro and the Saber11 examination. Although ICFES does not have a dataset where the results from both exams are matched for each student, the Institute provides a codification strategy to do this. It is important to highlight that although Colombian law (Law 30 of 1992) recognizes four different types of institutions that can offer a higher education program in Colombia⁵³, we restricted our sample to only include those universities that offer undergraduate and graduate programs (“Specialization”, Masters, and Doctoral programs), and also contribute in scientific research. We have excluded from our sample the information that corresponds to students enrolled in “university institutions”, “technological institutions” and “professional technical institutions”. After matching and dropping all the incomplete observations (where there was no information for both SaberPro and Saber 11 results), 56306 observations are left. This is 68% of the complete sample of university students taking the examination in the second semester of 2011. In the year 2011, those students for which the matching was possible obtained a score 1.21% higher on average in comparison to the total sample in the mathematics examination, and 0.89% in the language examination. Therefore, the sample of interest is slightly biased towards higher performing students but such bias is similar for students who come and those students who do not come from zones with a high intensity of conflict.

The dataset comprises the results from 56306 students coming from 214 universities. It should also be noted that less than half of the total number of students included in the dataset (39.22%) graduated from a public university. *Table 2* shows the sample distribution in detail.

⁵² The governmental agency in charge of these exams.

⁵³ For more information refer to section 3 of this chapter.

It is very important to acknowledge that given the nature of the dataset under analysis, self-selection bias and sample selection bias may be a concern. In order to tackle these issues (address the possibility of endogeneity bias), we are going to use the differences in differences methodology, which is widely used in the program evaluation literature. *Section 9* discusses in detail the selection bias issue that we address in this chapter.

Table 2 – Number of Observations					
Examination	Year	Students	Schools / Universities	Graduating from a Public Institution	
				#	%
Saber11	2002 – 1	483	101	219	45.3%
	2002 – 2	3615	810	2230	61.7%
	2003 – 1	677	135	323	47.7%
	2003 – 2	5221	1140	3138	60.1%
	2004 – 1	905	201	377	41.7%
	2004 – 2	7732	1503	4488	58%
	2005 – 1	1444	291	627	43.4%
	2005 – 2	11515	1985	6079	52.8%
	2006 – 1	2283	383	846	37.1%
	2006 – 2	16184	2296	8230	50.9%
	2007 – 1	2366	344	615	26%
	2007 – 2	3882	768	1608	41.4%
SaberPro	2011 – 2	56306	214	22084	39.22%

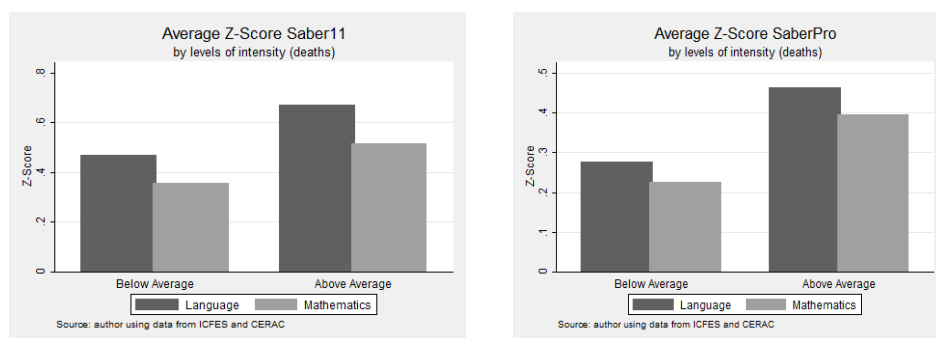
Source: author's calculations

The Saber11 and the SaberPro examinations have two testing components in common: the language and the quantitative reasoning examinations. As Balcazar and Ñopo (2014) have highlighted, the language component in the Saber11 examination is comparable to the reading component in the SaberPro examination. Both of these components evaluate competencies in the native language (Spanish) reading skills. Balcazar and Ñopo (2014) also highlight that the mathematics component in the Saber11 examination is also comparable to the quantitative reasoning component in the SaberPro examination. Nevertheless, the level of difficulty of both the language examination and the mathematics examination is higher in the SaberPro exam. This is consistent with the fact that students should reach a higher level of aptitude in these subjects during their undergraduate

career. Taken that into consideration, and following Balcazar and Ñopo (2014), we are not going to compare the scores obtained by students in the two examinations directly. Instead, we are going to use the position obtained by the student on the distribution of scores as z-scores, and with that study the change that students experience from the Saber11 to the SaberPro examination (Balcazar and Ñopo 2014).

Figure 1 shows the average z-scores for students coming from zones with a high intensity of conflict and coming from zones with a low intensity of conflict. Those students who graduated from schools located in municipalities with above average levels of intensity of conflict were categorized as students coming from high intensity of conflict zones. The average z-scores for these two groups in the quantitative reasoning and the language examinations are reported for the Saber11 and the SaberPro examinations. Figure 1 shows that students from zones with a high intensity of conflict have higher levels of performance in comparison to students from zones with a low intensity of conflict in the two subjects under study.

Figure 1 - Average z-scores for students coming from zones with a high intensity of conflict and coming from zones with a low intensity of conflict

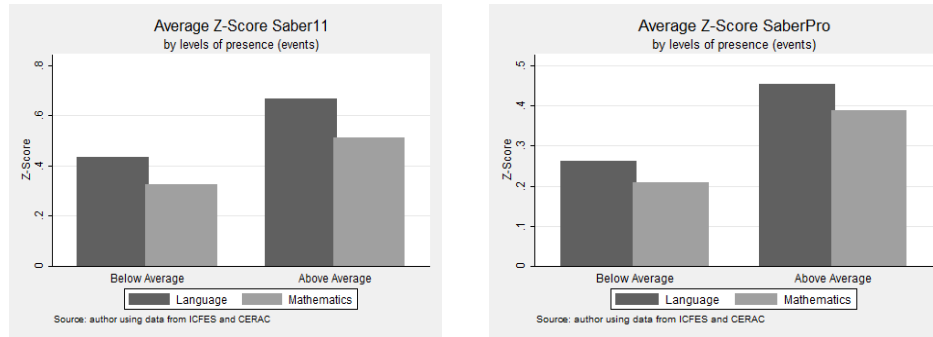


Source: author using data from ICFES and CERAC

Figure 2 shows the average z-scores for students coming from zones with a high presence of conflict and coming from zones with a low presence of conflict. The average z-scores for these two groups in the quantitative reasoning and the language examinations are reported for the Saber11 and the SaberPro examinations. We divided the students into two categories: **(1)** those students who graduated from schools located in municipalities with above average levels of presence of conflict were categorized and **(2)** those students who graduated from schools located in municipalities with below average levels of presence of conflict. Figure 2 shows that students from zones with a high presence of conflict have

higher levels of performance in comparison to students coming from zones with a low presence of conflict in the two subjects under study.

Figure 2 - Average z-scores for students coming from zones with a high presence of conflict and coming from zones with a low presence of conflict



Source: author using data from ICFES and CERAC

It is possible that those students who graduated from schools located in municipalities with above average levels of conflict are having a better performance than those students who graduated from schools located in municipalities with below average levels of conflict because an auto-selection mechanism could potentially be in action. It could be the case that those school-aged and university-aged individuals that are mostly affected by conflict are the ones that do not participate in the formal education system. Those individuals that do not have the resources (monetary, psychological, family support) to overcome or adapt to the rough conditions that are prevalent in conflict areas are probably more likely to drop out from the formal educational system. It is probable that those individuals coming from municipalities with above average levels of conflict who continue with their formal education have certain resources that allow them to stay in the system. There is auto-selection because those individuals with resources to overcome or adapt to the harsh conditions are most likely to continue with their studies. This auto-selection mechanism could be linked to previous research efforts that have studied the relationship between conflict, lower levels of enrollment and school dropouts. However that literature has only studied this relationship at the High School and Elementary School level. Rodriguez and Sanchez (2012) have found that armed conflict reduces the average years of schooling in 8.78% for a pooled sample of Colombian children, and it reduces the average years of schooling in 17.03% for children between sixteen and seventeen years old. Barrera and Ibañez (2004) have also found that after controlling for individual, household and other context variables, violence influences negatively

school enrollment. Furthermore, Sánchez and Diaz (2005) show that municipalities which were affected by the activities of illegally armed groups, had slower growth rates of primary and secondary school enrolment compared to those in municipalities that were not affected by illegal activities using a matching mechanism. Wharton and Oyelere (2012) also show that children who live in a municipality with high levels of conflict have a gap in enrolment and accumulation.

CERAC's Colombian Conflict Dataset

Information about conflict events was obtained from the Resource Centre for Conflict Analysis (CERAC). CERAC is a think tank based in Bogotá (Colombia) with three main lines of research: violent conflict, political violence, and violence associated to drug trafficking. Following these lines of research, an important area of study at CERAC is the quantification of violence directly associated to the internal armed conflict in Colombia. To quantify violent armed conflict, the research team at CERAC developed a methodology to collect, document, code and register information about this type of violence. The construction of CERAC's database started in the year 2002. This database contains information about the violent actions, the event in which these actions occur, the victims of the event, the attackers and the group to which they belong. CERAC's methodology is explained in detail in *Appendix B of Chapter 7*.

Using information from CERAC, we have included variables that capture the intensity of the conflict and the presence of conflict as the central components of this empirical analysis. The intensity of the conflict is quantified by the number of civilian deaths in each municipality. The presence of conflict is quantified by the number of conflict events in each municipality.

National Planning Department (DNP) Municipality Indicators

The Colombian National Planning Department (*DNP – Departamento Nacional de Planeación*) makes available annual information about the social and financial status of Colombian municipalities. This information is reported in the Municipal Development Indicator (*INDEMUN*). The Sustainable Territorial Development Program (*DDTS – Dirección de Desarrollo Territorial Sostenible*), which is run by the *DNP*, is in charge of calculating the Municipal Development Indicator. In the calculation of this index a number of variables are taken into consideration: the percentage of population living in the municipal urban areas, the percentage of households with clean water service, the percentage of households with sewage

system, the percentage of households with energy service, the percentage of the population without unmet basic needs in the municipal urban areas, the percentage of the population without unmet basic needs in the municipal rural areas, the percentage of the population that can read, the percentage of the population who attend school, tax income per capita, and municipal public investment per capital. A value close to 100 in this indicator is awarded to municipalities with high levels of development and a value close to 0 is awarded to municipalities with very low levels of development.

7.2 Variables

The Dependent Variables

Our analysis focuses on two outcome variables: the Z-score obtained in the mathematics examinations and the Z-score obtained in the language examinations. There is an important reason to use Z-scores in our analysis instead of the original score reported by ICFES. If we were to analyze the reported Saber11 results of a student who obtained a score of 70 in mathematics in the year 2002 and of a student who obtained a score of 80 in mathematics in the year 2003, we would be tempted to say that the student who took the examination in 2003 did better. However, the examination taken in these two years is different and therefore it is not appropriate to compare the scores of these two students. Nevertheless, if we know the mean and the standard deviations of the 2002 and 2003 distributions we could transform the scores to Z-scores in order to make the comparison between these two students possible.

A Z-score is a standardization that transforms the original distribution to one in which the mean becomes zero and the standard deviation becomes 1. In that sense, the Z-score quantifies the original score obtained in the exit examination in terms of the number of standard deviations that that score is from the mean of the distribution. Once we have done the Z-score transformation, it is possible to use these new transformed scores to compare scores from different distributions.

A negative Z-score indicates that the original score was below the mean, and a positive Z-score indicates that the original score was above the mean. The value of the Z-score can be interpreted as the number of standard deviations the score is from the mean in that direction (negative or positive).

Table 3 – Average Language and Math Z-Scores in the High School Exit Examination (Saber11) and the University Exit Examination (SaberPro)				
Year 2011	Saber11		SaberPro	
	Mean	Std. Dev.	Mean	Std. Dev.
Math Z Score	0.467	1.114	0.343	1.058
Language Z Score	0.611	0.984	0.406	0.952
Observations (students)	56306		56306	

Source: author's calculations

Table 3 presents the statistics of the Z-scores obtained by the students in the mathematics and language (high school and university) exit examinations. The number of observations reported corresponds to the number of students in the sample considered for each examination.

The Independent Variables

According to the literature, family and individual background variables play a crucial role on educational performance (Häkkinen et al. 2003, Haveman and Wolfe 1995, Woßmann 2003). The inclusion of parental education variables (level of education of the father and level of education of the mother) is particularly important considering that studies, including those of Woßmann (2003) and Häkkinen et al. (2003), have found that students whose parents have a higher level of education have a better performance than those students whose parents have a low level of education. This could be the case because of genetic transmission of skills from parents to kids, and because highly educated parents tend to provide a good cultural environment that is adequate for learning.

Table 4 presents information about the educational background of the parents of the students in the sample. The variables *level of education of the student's father* and *level of education of the student's mother* divide the different levels of education into eight different categories, which go from no formal studies to completion of graduate school. The information is available for the two periods under analysis. In *period 0* (end of high school studies), on average around 30.40% of the mothers had completed High School, while 26.46% of the fathers had completed that same level of schooling. As for university studies, around 13.23% of the mothers had completed a university degree, whereas around 17.26% of the fathers had completed a university degree. In *period 1* (end of university studies), on average around 24.03% of the mothers had completed High School, while 21.20% of the fathers had completed that same level of schooling. For that same period, around 16.42% of the mothers had completed a university degree, whereas around 18.48% of the fathers had completed a university degree.

Table 4 - Parental Education - High School (<i>period 0</i>) and University (<i>period 1</i>)				
	Mother Education		Father Education	
	High School (%)	University (%)	High School (%)	University (%)
Parent didn't complete any studies (level 1)	1.48	0.67	2.03	2.28
Elementary school was not completed (level 2)	NA	8.14	NA	10.70
Elementary school was completed (level 3)	26.74	8.27	27.70	8.81
High school was not completed (level 4)	11.95	14.01	9.51	11.92
High school was completed (level 5)	30.40	24.03	26.46	21.10
Technical program was not completed (level 6)	NA	3.61	NA	3.28
Technical program was completed (level 7)	8.63	12.01	9.51	9.16
University degree was not completed (level 8)	NA	3.64	NA	4.32
University degree was completed (level 9)	13.23	16.42	17.26	18.48
Graduate school (level 8)	6.99	8.82	6.73	9.94
No information	NA	0.37	NA	0.01

Source: author's calculations

For the econometric exercise presented in this chapter, we have decided to divide the parental education variable into four categories comprising the different levels of education discussed above. The four categories are the following: no education (parent didn't complete any studies), low level of education (elementary school), mid-level of education (High school and/or technical studies) and high level of education (university and/or graduate school). *Table 5* shows the descriptive statistics for that categorization.

Two additional variables were also considered to account for individual characteristics: a dummy variable to capture the gender of the student taking the examinations and a variable capturing the age of the students at both periods. *Table 6* contains information about the gender distribution of the sample. The percentage of female students in the sample is higher than the percentage of male students; 39.98% of the students taking the examinations were male and 60.02% were female. *Table 7* presents the average language and math Z-Scores in the high

school exit examination (Saber11) and the university exit examination (SaberPro) by gender.

Table 5 - Parental Education Categorization - High School (<i>period 0</i>) and University (<i>period 1</i>)				
	Mother Education		Father Education	
	High School (%)	University (%)	High School (%)	University (%)
No education - parent didn't complete any studies (level 1)	1.48	0.67	2.03	2.28
Low level of education – Partial or Complete Elementary School Education (level 2)	27.32	16.41	28.20	19.51
Mid-level of education – Partial or Complete High school and/or technical studies (level 3)	50.98	53.66	45.78	45.46
High level of education – Partial or Complete University and/or graduate school (level 4)	20.21	28.88	23.99	32.74

Source: author's calculations

Table 6 – Gender		
	Number of Students	%
Male	22512	39.98
Female	33794	60.02

Source: author's calculations

Table 8 contains information about the age of the students who took the high school exit examination (Saber11) and the university exit examination (SaberPro). The average age of the students taking the high school exit examination is 16.64 years with a standard deviation of 2.00, whereas the average age of the students taking the university exit examination is 23.37 with a standard deviation of 2.44.

Table 7 – Average Language and Math Z-Scores in the High School Exit Examination (Saber11) and the University Exit Examination (SaberPro) by Gender								
Year 2011	Saber11				SaberPro			
	Male		Female		Male		Female	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Math Z Score	0.706	1.192	0.307	1.028	0.649	1.133	0.140	0.951
Language Z Score	0.685	1.005	0.562	0.966	0.465	0.964	0.368	0.942
Observations (students)	22512		33794		22512		33794	

Source: author's calculations

Table 8 – Age of the Students Who Took the High School Exit Examination (Saber11) and the University Exit Examination (SaberPro)				
Year 2011	Saber11		SaberPro	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	16.64	2.00	23.37	2.44
Observations (students)	56306		56306	

Source: author's calculations

There could be significant differences in terms of the tuition that students have to pay at each of the universities and schools considered in our analysis. Such differences could help us capture dissimilarities in the amount of resources available at each university and at each school. For that reason, it is important to control for the cost of the tuition. To do so, we have included an explanatory variable (categorical) to capture the cost of attending school at *period 0* and the cost of attending university at *period 1*. For the econometric exercise presented in this chapter, we have decided to divide the cost of tuition variable into four categories comprising the different levels of tuition paid by the students. The four categories are the following: free tuition, low tuition (students pay less than \$400 dollars per semester in the case of university (*period 1*) and students pay less than \$30 dollars per month in the case of high school (*period 0*)), mid tuition (students pay between \$401 and \$2000 dollars per semester in the case of university and students pay between \$36 and \$100 dollars per month in the case of high school) and high tuition (students pay more than \$2000 dollars per semester in the case of university and students pay more than \$100 per month in the case of high school). *Table 9* presents information about the tuition that was paid by the students who are included in the sample at the high school and university level.

We have also included a dummy variable to control for differences between public and private educational institutions. *Table 10* presents the information regarding the public or private categorization of the educational institution. It should be noted that 51.11% of the students included in the dataset graduated from a public high school, whereas only 39.22% of the students included in the dataset graduated from a public university.

Table 9 – High School and University Tuition				
	High School		University	
	#	%	#	%
Free	26842	47.98	769	1.38
Low Tuition	8288	14.81	17425	31.17
Mid Tuition	18038	32.24	27455	49.11
High Tuition	2781	4.97	10251	18.34

Source: author's calculations

International literature has shown that the economic, political and social consequences of civil war are significant (Justino 2010). However, the connection between conflict and education is an under-analysed area of study. This is partly due to the difficulties involved in the systematic collection of evidence on why, where and how attacks occur. As mentioned by Justino (2010), high quality monitoring data and systematic reporting of events are very limited. Additionally, in many places affected by conflict censorship of information is also common. The availability of a rich database about conflict events in Colombia, maintained by the Resource Centre for Conflict Analysis (CERAC), opens new possibilities for the analysis of the relationship between conflict and education. Using information from CERAC, we have included variables that capture the intensity of the conflict and the presence of conflict as the central components of this empirical analysis. The intensity of the conflict is quantified by the number of civilian deaths in each municipality. However, in some of the specifications considered we also included an additional measure of intensity: a dummy variable that takes a value of 1 if the municipality recorded an intensity of conflict above the national average. The presence of conflict is quantified by the number of conflict events in each municipality. In this case we also included an additional measure of presence of conflict: a dummy variable that takes a value of 1 if the municipality recorded conflict events above the national average.

Table 10 – Public / Private Educational Institution			
Examination	Year	Students	%
Saber11	Private	27526	48.89
	Public	28780	51.11
SaberPro	Private	34194	60.78
	Public	22058	39.22

Source: author's calculations

8. The Individuals Under Analysis

The sample that is under analysis in this chapter is restricted to those students who have completed at least 75% of their academic program at a recognized university (i.e. only students enrolled in their fourth year of undergraduate studies or above are considered in the analysis). These students belong to a small group of individuals who are covered by the Colombian higher education system and have not dropped out from the higher education system. In that sense, it is important to highlight that the coverage rate of higher education programs in Colombia is still low and the drop out rate is high.

Table 11 Higher Education Coverage Rate in Latin America			
Country	2009	2010	2011
Latin American and Caribbean Average	37%	41%	42%
Argentina	71%	75%	N.A.
Brazil	36%	N.A.	N.A.
Chile	59%	66%	71%
Colombia	35.3%	37.1%	40.8%
Cuba	115%	95%	80%
El Salvador	23%	23%	25%
Mexico	27%	28%	29%
Panama	45%	46%	N.A.
Paraguay	37%	35%	N.A.
Puerto Rico	81%	86%	86%
Uruguay	63%	63%	N.A.
Venezuela	78%	N.A.	N.A.

Source: Ministerio de Educación de Colombia with information from UNESCO

In fact, by the year 2011 the higher education coverage rate only reached 40.8%. Nevertheless, according to the information reported by the Ministry of Education the coverage rate has improved significantly from the year 2005 to the year 2011, going from 28.4% to 40.8%. This rate is low in comparison to other countries in the region, and is below the Latin American and Caribbean average (*Table 11*).

Additionally, as we can see in *Table 12*, the dropout rate from higher education programs is high. The dropout rate reported for the year 2011 was 11.8%. Even though the dropout rate has decreased in comparison to the number reported for the year 2005 (13.4%), the rate observed in 2011 is still high (Ministerio de Educación 2014).

9. Selection Bias

In order to have a better understanding of how different factors determine the impact of higher education on students, it is important to take into consideration the possibility of selection bias. Selection bias can be caused by two different problems: sample selection and self-selection.

Table 12 Dropout Rates in Colombian Higher Education	
Year	Colombia
2005	13.4%
2006	12.9%
2007	10.7%
2008	12.1%
2009	12.4%
2010	12.9%
2011	11.8%
2012	11.1%
2013	10.4%

Source: Ministerio de Educación de Colombia with information from UNESCO

Sample selection bias can arise when individuals are selected into a sample in a non-random manner, often times based on data availability. For instance, if we were to use a sample consisting only of students who came from a conflict zone and enrolled in higher education, our study of the effects of conflict on higher education may be imprecise because the results are prone to bias due to sample selection. The sample under analysis would be a non-random restricted sample

because the dependent variable (the score of students coming from conflict zones in the higher education exit examination) would only be observed for those students coming from conflict zones who choose to enroll in higher education. That is indeed an important source of bias because we would be inferring the characteristics of an entire population (individuals affected by the civil conflict) from a very restricted part of the population (students who attend university). It follows that the observed results for students from conflict zones in the non-random restricted sample may be higher or lower than the results observed for the full sample, misestimating the effect of conflict on higher education.

Self-selection bias can arise when predictors of an outcome are associated with other unobserved or observed variables. If we do not take into account the possibility that certain students auto-select to enroll in a higher education program, the results might also be affected by self-selection bias. It is possible that the independent variable of interest (if the student came from a conflict zone) depends on a number of factors that are related to the levels of achievement in the Higher Education exit examination. If we ignore these factors the coefficient estimates obtained in the empirical exercise might be imprecise and our model is very likely to be mis-specified.

In order to address the problem of sample-selection and endogeneity bias, we are going to use the differences in differences methodology, which is widely used in the program evaluation literature. In this case, a difference in differences approach is applied in order to estimate the treatment effects of coming from a conflict zone on the levels of achievement in the Colombian higher education exit examination.

10. Methodology – Identification Strategy

In the dataset that was constructed for this empirical exercise we can observe higher education students coming from conflict zones and higher education students coming from non-conflict zones. However, if we were to make a simple comparison of these two groups the results would likely be biased due to the selection bias issue that was discussed in the previous section. In order to address potential bias we use a difference in differences estimation strategy. The following sub-sections explain in detail this methodology.

10.1 Differences in Differences⁵⁴

Observing the before-and-after change in scores in a standardized examination for students coming from conflict zones after participating in a higher education program will not give us the causal impact of conflict on academic achievement because many other factors are also likely to influence achievement over time. The difference in the before-and-after scores for students coming from high conflict zones, i.e. the first difference, controls for factors that are constant over time in that group, since we are comparing the same group to itself (Gertler et al. 2011). However, according to Gertler et al. (2011), after accounting for that first difference we would still be left with the outside time-varying factors. According to these researchers, one way to capture the time-varying factors is to measure the before-and-after change in outcomes for students who did not come from conflict zones but were also exposed to a higher education program, i.e. the second difference (Gertler et al. 2011). If we clean the first difference of time-varying factors that affect the outcome of interest by subtracting the second difference, then we can eliminate the main source of bias that could have affected our analysis in the simple before and after comparison (Gertler et al. 2011). “Cleaning” by using these differences is known in the literature as the differences-in-differences estimation approach. The differences-in-differences methodology compares the changes in outcomes over time between a group that is enrolled in a program, i.e. the treatment group, and a group that is not, i.e. the comparison group (Gertler et al. 2011).

Our differences in differences estimation strategy compares the change in the scores in the standardized examination before and after participating in a higher education program for individuals coming from conflict zones to changes in scores amongst those individuals who did not come from conflict zones. It should be noted that in a difference in differences estimation approach the treatment and comparison groups do not need to have the same pre-intervention conditions (Gertler et al. 2011). However, in order for the differences in differences methodology to be valid the control group must accurately represent the change in outcomes that would have been experienced by the treatment group in the absence of treatment (Gertler et al. 2011).

It should be highlighted that even though the differences-in-differences methodology can be used to deal with differences between the treatment and the comparison group that are constant over time, it will not take care of differences between the treatment and comparison groups that change over time (Gertler et al. 2011). Therefore, in order to obtain a valid estimate of the counterfactual from a

⁵⁴ Equivalent to a value added specification with individual fixed effects.

differences-in-differences estimation approach we need to assume that there is no time varying differences between the treatment and comparison groups. This implies that without treatment the scores should increase or decrease at the same rate in both groups. In other words, the outcomes for both groups should show the same trends in the absence of treatment (Gertler et al. 2011).

The availability of high school exit examination results (baseline data) and college exit examination results (post-intervention data) makes it possible to estimate the impact of the intervention by assuming that unobserved heterogeneity is time invariant and uncorrelated with the treatment over time (Khandker et al. 2010). The differences-in-differences estimate can be calculated using a regression framework. Following Khandker et al. (2010), the estimating equation could be specified as follows:

$$Y_{it} = \alpha + \beta T_{it}t + \rho T_{i1} + \gamma t + \varepsilon_{it} \quad (4)$$

In equation (4) β is the coefficient of interest that captures the average differences-in-differences effect of the intervention. β is the coefficient on the interaction between the treatment variable (T_{it}) and the time dummy ($t = 0, 1$).

Equation (4) can be re-written in expectations form:

$$E(Y_1^T - Y_0^T | T_1 = 1) = (\alpha + DD + \rho + \gamma) - (\alpha + \rho) \quad (5)$$

$$E(Y_1^c - Y_0^c | T_1 = 0) = (\alpha + \gamma) - \alpha \quad (6)$$

If we subtract equation (6) from equation (5) we can obtain the average differences-in-differences effect of the intervention. The differences-in-differences estimate is unbiased only if the potential source of selection bias is **additive** and **time invariant** (Khandker et al. 2010). For the differences-in-differences estimator to be unbiased, the following conditions must hold:

1. The model in equation (4) should be correctly specified.
2. The error term should be uncorrelated with the other variables in the equation:

$$Cov(\varepsilon_{it}, T_{it}) = 0$$

$$Cov(\varepsilon_{it}, t) = 0$$

$$Cov(\varepsilon_{it}, T_{it}) = 0$$

The last assumption, which is also known as the **parallel-trend** assumption, is the most important assumption in order to correctly apply a differences-in-differences approach (Khandker et al. 2010). This assumption implies that unobserved characteristics affecting program participation do not vary over time with treatment status.

However, there is not a direct way to prove that the differences between the treatment and comparison groups would have moved at the same rate in the absence of the program (Gertler et al. 2011). This is the case because we cannot observe what would have happened to the treatment group in the absence of the treatment (Gertler et al. 2011). In other words, we cannot observe the counterfactual. Therefore, in order to apply the differences-in-differences methodology we need to assume that in the absence of the program the outcome in the treatment group would have moved at the same rate as the outcome in the comparison group. The validity of this assumption can be analysed but it cannot be proved. According to Gertler et al. (2011), a good validity check is to compare changes in outcomes for the treatment and comparison groups before implementing an intervention. According to these researchers, if the outcomes moved at the same rate before the intervention then we can have some confidence that the outcomes would have continued to move at the same rate after the intervention (Gertler et al. 2011).

10.2 Evaluating the Impact

In this chapter we evaluate the impact of a treatment (coming from a conflict zone) on an outcome Y (score in the mathematics and language exit examinations) over a group of individuals. We assume that there are two groups: the treatment group (those students coming from conflict zones) and the control group (those students who do not come from a conflict zone). Using information from CERAC, we identify the student's exposure to conflict during high school using variables that capture the intensity ($INTENSITY_j$) of the conflict and the presence ($PRESENCE_j$) of conflict at the municipal level. The intensity of the conflict is quantified by the number of civilian deaths in each municipality. The presence of conflict is quantified by the number of conflict events in each municipality. However, in some of the specifications considered we also included a dummy variable that

takes a value of 1 if the municipality recorded an intensity of conflict above the national average as an additional measure of intensity, and a dummy variable that takes a value of 1 if the municipality recorded conflict events above the national average as an additional measure of presence of conflict. In that sense, we are going to estimate a discrete (binary) and a continuous version of the differences-in-differences model.

Using a differences-in-differences estimation strategy we compare the examination results of students who came from conflict zones to the examination results of students who came from non-conflict zones. The differences-in-differences estimating equations would be specified as follows:

$$Y_{ijt} = \alpha + \beta_1 period + \beta_2 INTENSITY_j + \beta_3 (period \times INTENSITY_j) + \beta_4 X_{ijt} + \delta_j + \rho_i + \varepsilon_{it} \quad (7)$$

$$Y_{ijt} = \alpha + \beta_1 period + \beta_2 PRESENCE_j + \beta_3 (period \times PRESENCE_j) + \beta_4 X_{ijt} + \delta_j + \rho_i + \varepsilon_{it} \quad (8)$$

ε_{ijt} is a random unobserved error term which contains all determinants of Y_{ijt} that are not included in the model. Y_{ijt} is a continuous variable that captures the score in the exit examination and $period$ is a time dummy that takes a value of 1 to indicate the second examination (SaberPro) period. β_3 is the differences-in-differences estimate. X_{ijt} is a vector of covariates that vary over time within units. δ_j is a municipality of origin specific fixed effect and ρ_i is a student fixed effect. We include the municipality of origin and student fixed effects in order to control for unobserved correlation of observations within municipalities and individuals. The inclusion of municipality of origin fixed effects eliminates all observed and unobserved municipal characteristics that are constant across students from the same municipality of origin, removing any bias that is generated by municipality characteristics.

To estimate **equations (7) and (8)**, we used the results of students who took the Saber11 examination (High School exit examination) between the year 2002 and the year 2007, and the results in the SaberPro examination (University exit examination) for the year 2011 for this same group of students. The students included in this sample graduated from High School between the years 2002 and 2007, and decided to continue with their studies in order to complete a higher education program. Therefore, these students also took the university exit examination (SaberPro) in the year 2011.

We also examined other specifications to capture the heterogeneous effect of the treatment. In particular, the specific impact of conflict by gender is particularly interesting. To calculate the gender differential effect, we estimated additional specifications. The results of these additional specifications are also reported in the results section.

10.3 Validity of the Differences in Differences Estimation Strategy

A very important assumption that needs to hold in order to implement a differences-in-differences methodology is that both groups that are being examined (students coming from conflict zones and students coming from non-conflict zones) should have similar underlying human development trends. In order to interpret β as the effect of the intensity of the conflict, it is necessary to assume that if the student would not have come from a conflict zone, the difference in the standardized examination scores between the group of students who came from a conflict zone and the group of students who did not come from a conflict zone would have been the same across municipalities of varying intensity (presence) of conflict. In other words, the differences-in-differences estimator relies on the assumption that in the absence of conflict the trends in the levels of academic achievement (exit examinations) would have been similar for both groups. However, if conflict affected municipalities have systematic differences in terms of social and economic development it is possible that this assumption might not hold. It is reasonable to think that those municipalities that have been more affected by the conflict might have a lower level of social and economic development that could explain, at least partially, the differences in the levels of achievement in the exit examinations. If that happens, our results would not be capturing the effect of conflict on educational achievement because the effect would be confounded with the poor socio-economic conditions that are prevalent in conflict zones.

In order to control for possible differences in the levels of social and economic development in municipalities with high levels of conflict and municipalities with low levels of conflict, we use an indicator constructed by the Colombian National Planning Department that captures the level of municipal development. As we explained in *Section 7.1* in the calculation of this index a number of variables are taken into consideration, including the percentage of population living in the municipal urban areas, the percentage of households with clean water service, the percentage of households with sewage system, the percentage of households with energy service, the percentage of the population without unmet basic needs in the municipal urban areas, the percentage of the population without unmet basic needs in the municipal rural areas, the percentage of the population that can read,

the percentage of the population who attend school, tax income per capita, and municipal public investment per capital. A value close to 100 in this indicator is awarded to municipalities with high levels of development and a value close to 0 is awarded to municipalities with very low levels of development.

The municipality development indicator is included in the specifications in two different ways. First, we include the development indicator of the municipality where the student took the examination corresponding to the year in which the examination was administered. Second, we include the four-year average development indicator of the municipality where the student took the examination corresponding to the four years prior to the year in which the examination was administered (average lagged measure of municipal development). The lagged measure is included to account for possible differences in the levels of social and economic development in municipalities with high and low levels of conflict during the years in which the students did their studies.

10.4 Differences in Differences Estimation Strategy vs. Latent Growth Curve Model

It should be noted that latent growth curve models could also be used to understand the change and the trajectory of change of outcome variables in education. A difference in differences approach could be used to evaluate whether differences in performance in a standardized examination between those students who come from conflict zones and those students who do not come from conflict zones changes over time (within group variation), whereas a latent growth curve model could be used to study the effects of time-varying and time-invariant explanatory variables on the trajectory of performance in a standardized examination⁵⁵ (Hess 1999) (Hox 2005). Latent growth curve models require at least three time intervals and are not designed to deal adequately with selection bias. As Hess (1999) points out, “one of the most obvious disadvantages in using latent growth model methods in assessing program impact is the fact that it requires measures of an outcomes at multiple time points.” Duncan and Duncan (2009) further highlight that growth curve models should include information at three or more points in time in order to evaluate the validity of the growth trajectory over time and in order to obtain reliable estimates of the parameters. According to these researchers, precision increases as models include information for more points in time (Duncan and Duncan 2009). On the other hand, a difference in differences⁵⁶ estimation technique requires information for only two time intervals, and it could help eliminate bias due to unobserved heterogeneity between individuals.

⁵⁵ Latent growth curve models quantify the variation across groups (intercept) and the rate of that change (slope).

⁵⁶ This is true for any fixed-effects estimator.

It should be acknowledged that latent growth models could possibly have an advantage over a difference in difference model in the analysis of longitudinal data (three periods or more) (Stoel 2004) (Hox 2005). However, in the empirical exercise presented in this chapter the unavailability of data for more than two periods restricts the use of this type of models. A latent growth curve model would have been more appropriate if we had information for three or more periods because using data from more than two points in time in a latent growth curve model can help to maximize the reliability of the results of the trajectory of individual change (Hess 1999).

11. Results

This section discusses the results that we obtained using difference in differences regressions to examine the impact of exposure to conflict on achievement gains in Colombian standardized exit examinations. By considering students who have been exposed to a conflict environment and students who have not been exposed to a conflict environment, we want to find out whether the students affected by conflict have gains or losses in performance in comparison to those students who have not been affected, using the results from the Colombian high school exit examination and the Colombian college exit examination. To tackle this research question we use the theoretical framework employed in educational value added models. As explained in a previous section, educational value added models are models that use student examination results at different points of time in order to measure their learning gain. Following this theoretical backbone, we use a difference in differences estimation strategy in order to quantify the student's learning gains using information from two points in time. Our difference in differences strategy uses the spatial and temporal variation of conflict in Colombia to estimate the effect of exposure on the results in standardized exit examinations. The difference in differences approach is useful in this case to tackle the possible concerns associated to selection bias.

In this empirical exercise we used two different measures of conflict: the intensity (*INTENSITY_j*) of the conflict and the presence (*PRESENCE_j*) of conflict at the municipal level. The intensity of the conflict is quantified by the number of civilian deaths in each municipality. The presence of conflict is quantified by the number of conflict events in each municipality. However, in some of the specifications considered we also included a dummy variable that takes a value of 1 if the municipality recorded an intensity of conflict above the national average as an additional measure of intensity, and a dummy variable that takes a value of 1 if the

municipality recorded conflict events above the national average as an additional measure of presence of conflict. In that sense, we estimated a discrete (binary) and a continuous version of the differences-in-differences model (fixed-effects model).

For the analysis of the results, it is important to consider the fact that we restricted the sample under analysis. It is necessary to recall that after matching and dropping all the incomplete observations (where there was no information for both SaberPro and Saber 11 results) we were left with a sample that is slightly biased towards higher performing students. However, such bias is similar for students who come and those students who do not come from zones with a high intensity of conflict.

The results are reported in two sub-sections. In the first subsection, we report the results that we obtained using our first measure of conflict, i.e. intensity of conflict (deaths). In the second subsection, we report the results that we obtained using presence of conflict (events) at the municipal level as our measure of conflict.

11.1 Intensity of Conflict

Mathematics

This sub-section presents the results from estimating our difference-in-differences model for the mathematics exit examination using intensity of conflict as our variable of interest. As mentioned above, the intensity of the conflict is quantified using two different measures: **(i)** a dummy variable that takes a value of 1 if the municipality recorded an intensity of conflict (deaths) above the national average, and **(ii)** a positive integer variable that captures the number of civilian deaths in each municipality.

Table 13 reports the results of the difference in differences regressions for the specifications that use our first measure of intensity of conflict, above average deaths, as variable of interest. This table includes the results for specifications that use a contemporaneous measure of intensity (intensity in the year when the student took the examination) and specifications that use a 5-year average measure of intensity. In this table, we sequentially introduce various controls for individual/parental and school characteristics. *Columns 1 and 2* report the results of the specifications that did not include any controls, *Columns 3 and 4* report the results of the specifications that included controls for individual and parental characteristics, *Columns 5 and 6* show the results of the specifications that controls for school characteristics, *Columns 7 and 8* report the results of the specifications that included individual/parental and school controls, and *Columns 9 and 10* report the results of the specifications that include all controls and an interaction term to

account for heterogeneous gender effects. Observed individual/parental characteristics include the age of the student, the mother's level of education, and the father's level of education. Observed school characteristics include a dummy indicating whether the school is public and the level of tuition that students have to pay in order to attend.

All the results reported in *Table 13* include student and municipal fixed effects. As discussed in previous sections, the difference-in-differences estimation strategy allows to control for unobserved student-level and municipal-level factors, which may have an impact on the students' exit examination scores. Additionally, all the specifications control for the size of the municipality using the logarithm of the population.

Table 13. Differences in Differences – Intensity of Conflict Discrete Version (Binary Treatment Variable) - Mathematics Examination					
	Specification				
	(1)	(2)	(3)	(4)	(5)
DiD Estimate - Period × Above Average Deaths (Dummy)	0.028 (0.114)		0.134** (0.054)		0.037 (0.114)
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)		0.033 (0.120)		0.170** (0.073)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0372	0.0371	0.1320	0.1315	0.0388
	Specification				
	(6)	(7)	(8)	(9)	(10)
DiD Estimate - Period × Above Average Deaths (Dummy)		0.145*** (0.055)		0.073 (0.060)	
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)	0.047 (0.120)		0.176** (0.075)		0.103 (0.078)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0387	0.1373	0.1367	0.1395	0.1395

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

The difference-in-differences estimate is the coefficient on the interaction between the dummy indicating the period and the dummy that indicates whether a student comes from a municipality that has reported deaths associated to the civil conflict in a quantity above the national average. This coefficient captures the gains in achievement in the exit examinations for students coming from conflict zones. The coefficient on the interaction is positive and significant in four of the ten specifications considered in *Table 13*. Of those four specifications, two correspond to specifications that use the contemporaneous measure of intensity and two correspond to specifications that use the 5-year average measure of intensity. The results suggest that students coming from conflict zones have significantly greater gains in achievement than those students coming from non conflict zones. It is important to highlight that the identification strategy is only valid if changes overtime in the scores of the standardized examinations would be similar across municipalities in the absence of conflict.

The complete output, including the coefficients for all the controls, is included in *Appendix 2 (Table 25)*. In *Columns 3 and 4* we present the results for the specifications that only control for student and parental characteristics. In these specifications, the coefficient capturing the difference-in-differences estimate is positive and significant. Students from municipalities that have an above national average intensity of conflict have an improvement (gain) of 0.134 standard deviations in the mathematics examination Z-Score. Furthermore, students from municipalities that have a 5-year above national average intensity of conflict have an improvement (gain) of 0.170 standard deviations in the mathematics examination Z-Score. The results obtained for these specifications also show that the categorical variables that capture the levels of education received by the parents are positively and significantly related to gains in achievement. (*Table 25, Appendix 2*). In the case of parental education, the reference category is no parental education.

Columns 5 and 6 present the results for the specifications that only control for school characteristics. In these specifications, the coefficient capturing the difference-in-differences estimate is not significant, but the sign of the coefficient remains positive. As we can see in *Table 25 in Appendix 2*, the tuition paid in the educational institutions is positively and significantly related to gains in achievement in the standardized examinations. The reference category for the tuition categorical variables is no tuition.

In *Columns 7 and 8* we present the results for the specifications that control for student/parental and school characteristics. As we can see in this table, in these two specifications the coefficient capturing the difference-in-differences estimate is positive and significant. Results show that students from municipalities that have

an above national average intensity of conflict have an improvement (gain) of 0.145 standard deviations in the mathematics examination Z-Score. Similarly, students from municipalities that have a 5-year above national average intensity of conflict have an improvement (gain) of 0.176 standard deviations in the mathematics examination Z-Score. These results are analogous to the results obtained in the specifications that included student and parental characteristics as controls.

Columns 9 and 10 explore the gender heterogeneity of the conflict impact. In these columns we present the results of the specifications in which we interact the difference in difference variable with a male indicator. The coefficient of this interacted term is positive and significant. These results suggest that the gains in achievement are significantly stronger for male students than for female students in the sample. Male students from municipalities that have an above national average intensity of conflict have an additional improvement (gain) of 0.150 standard deviations in the mathematics examination Z-Score. Similarly, male students from municipalities that have a 5-year above national average intensity of conflict have an additional improvement (gain) of 0.154 standard deviations in the mathematics examination Z-Score.

Table 14 presents the results of the difference in differences regressions for the specifications that use our second measure of intensity of conflict, the number of conflict-related civilian deaths at the municipality, as variable of interest. The table reports the results for specifications that use a contemporaneous measure of intensity (intensity in the year when the student took the examination) and specifications that use a 5-year average measure of intensity. In *Table 14*, we sequentially introduce various controls for individual/parental and school characteristics. All the results reported in this table include student and municipal fixed effects. Also, all the specifications control for the size of the municipality using the logarithm of the population.

The difference-in-differences estimate is the coefficient on the interaction between the dummy indicating the period and the variable that indicates the number of conflict-related deaths in the municipality where the student took the exit examination. As we can see in *Table 14*, the coefficient on the interaction is positive and significant in all the specifications considered that use a contemporaneous measure of intensity. However, only 3 out of 5 of the specifications that use a 5-year average measure of intensity turned out to be significant.

Table 14. Differences in Differences – Intensity of Conflict Continuous Version (Continuous Treatment Variable) - Mathematics Examination					
	Specification				
	(1)	(2)	(3)	(4)	(5)
DiD Estimate - Period × Deaths	0.0029*** (0.0004)		0.003*** (0.0006)		0.003*** (0.0002)
DiD Estimate - Period × 5 Year Average Deaths		0.003*** (0.0003)		0.006 (0.111)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0391	0.0387	0.1372	0.1369	0.0410
	Specification				
	(6)	(7)	(8)	(9)	(10)
DiD Estimate - Period × Deaths		0.0035*** (0.001)		0.0024*** (0.001)	
DiD Estimate - Period × 5 Year Average Deaths	0.004*** (0.0003)		0.006 (0.111)		0.004*** (0.001)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0411	0.1433	0.1421	0.1455	0.1444

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

In *Columns 3 and 4*, the results for the specifications that only control for student and parental characteristics are reported. Results show that students coming from conflict affected municipalities have an improvement (gain) of 0.003 standard deviations in the mathematics examination Z-Score per additional conflict-related death. However, for the specification that considers the 5-year average of conflict-related deaths as a measure of intensity the coefficient of the interaction was not significant. It is also important to highlight that the categorical variables that capture the levels of education received by the parents are positively and significantly related to gains in achievement⁵⁷.

⁵⁷ The complete results from the econometric exercise, including the coefficients for all the controls, are included in Appendix 2 (Table 29).

Columns 5 and 6 present the results for the specifications that only control for school characteristics. In both of these specifications, the coefficient capturing the difference-in-differences estimate is significant, and the sign of the coefficient is positive. The results reported in *Table 14* show that students coming from conflict affected municipalities have an improvement of 0.003 standard deviations in the mathematics examination Z-Score per additional conflict-related death. Similarly, students from conflict affected municipalities have an improvement of 0.004 standard deviations in the mathematics examination Z-Score per additional conflict-related death in the analysis of the 5-year average of conflict-related deaths. It should also be highlighted that the tuition paid in the educational institutions is positively and significantly related to gains in achievement in the standardized examinations⁵⁸.

Columns 7 and 8 present the results for the specifications that control for student/parental and school characteristics. Results show that students coming from conflict affected municipalities have an improvement (gain) of 0.004 standard deviations in the mathematics examination Z-Score per additional conflict-related death. However, for the specification that considers the 5-year average of conflict-related deaths as a measure of intensity the coefficient of the interaction was not significant.

Columns 9 and 10 explore the gender heterogeneity of the conflict impact. These columns present the results of the specifications that include an interaction term between the difference in difference variable and a male indicator. In these specifications the coefficient capturing the difference-in-differences estimate is positive and significant. Results show that students coming from conflict affected municipalities have an improvement of 0.0024 standard deviations in the mathematics examination Z-Score per additional conflict-related death. Likewise, students from conflict affected municipalities have an improvement of 0.004 standard deviations in the mathematics examination Z-Score per additional conflict-related death in the analysis of the 5-year average of conflict-related deaths. The results obtained in this empirical exercise suggest that the gains in achievement are stronger for male students than for female students in the sample. Male students from municipalities affected by conflict have an additional improvement (gain) of 0.0023 standard deviations in the mathematics examination Z-Score per additional conflict-related death. Similarly, male students from conflict affected municipalities have an additional improvement (gain) of 0.005 standard deviations in the mathematics examination Z-Score per additional conflict-related death in the analysis of the 5-year average of conflict-related deaths.

⁵⁸ *Table 29 - Appendix 2*

Language

This sub-section presents the results from estimating our difference-in-differences model for the language exit examination using intensity of conflict as our variable of interest. The presence of conflict is quantified using two different measures: **(i)** a dummy variable that takes a value of 1 if the municipality recorded conflict events above the national average, and **(ii)** a positive integer variable that captures the number of conflict events in each municipality.

Table 15. Differences in Differences – Intensity of Conflict Discrete Version (Binary Treatment Variable) - Language Examination					
	Specification				
	(1)	(2)	(3)	(4)	(5)
DiD Estimate - Period × Above Average Deaths (Dummy)	0.009 (0.012)		0.046 (0.040)		0.016 (0.012)
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)		0.029 (0.117)		0.046 (0.055)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0669	0.0671	0.1621	0.1621	0.0684
	Specification				
	(6)	(7)	(8)	(9)	(10)
DiD Estimate - Period × Above Average Deaths (Dummy)		0.052 (0.042)		0.007 (0.045)	
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)	0.022 (0.118)		0.054 (0.057)		0.107* (0.059)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0685	0.1643	0.1641	0.1658	0.1666

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 15 presents the results of the difference in differences estimations for the specifications that use our first measure of intensity of conflict, above average deaths, as variable of interest. This table includes the results for specifications that

use a contemporaneous measure of intensity (intensity in the year when the student took the examination) and specifications that use a 5-year average measure of intensity. In *Table 15*, we sequentially introduce various controls for individual/parental and school characteristics. All the results reported in this table include student and municipal fixed effects. Also, all the specifications control for the size of the municipality using the logarithm of the population.

The coefficient capturing the difference-in-differences estimate was not significant for any of the specifications considered. However, the coefficient on the interaction term between the difference in difference variable and a male indicator in *Columns 9 and 10* turned out to be positive and significant. According to these results, male students from municipalities that have an above national average intensity of conflict have an additional improvement (gain) of 0.095 standard deviations in the language examination Z-Score. Similarly, male students from municipalities that have a 5-year above national average presence of conflict have an additional improvement (gain) of 0.112 standard deviations in the language examination Z-Score.

Table 16 reports the results of the difference in differences estimations for the specifications that use our second measure of intensity of conflict, number of conflict-related deaths, as variable of interest. This table includes the results for specifications that use a contemporaneous measure of intensity (intensity in the year when the student took the examination) and specifications that use a 5-year average measure of intensity. All the results reported in this table include student and municipal fixed effects. Also, all the specifications control for the size of the municipality using the logarithm of the population.

The coefficient capturing the difference-in-differences estimate was not significant for any of the specifications that used a contemporaneous measure of intensity of conflict. However, in 3 out of 5 of the specifications that use a 5-year average measure of intensity the coefficient capturing the difference-in-differences estimate turned out to be significant.

In *Columns 3 and 4*, the results for the specifications that only control for student and parental characteristics are reported. The results for the specification that considers the 5-year average of conflict-related deaths as a measure of intensity show that students coming from conflict affected municipalities have an improvement (gain) of 0.002 standard deviations in the language examination Z-Score per additional conflict-related death. *Columns 5 and 6* present the results for the specifications that only control for school characteristics. The results reported in *Table 16* show that students from conflict affected municipalities have an

improvement of 0.001 standard deviations in the language examination Z-Score per additional conflict-related death in the analysis of the 5-year average of conflict-related deaths. *Columns 7 and 8* display the results for the specifications that control for student/parental and school characteristics. The results for the specification that considers the 5-year average of conflict-related deaths as a measure of intensity show that students coming from conflict affected municipalities have an improvement (gain) of 0.002 standard deviations in the language examination Z-Score per additional conflict-related death.

Table 16. Differences in Differences – Intensity of Conflict Continuous Version (Continuous Treatment Variable) - Language Examination					
	Specification				
	(1)	(2)	(3)	(4)	(5)
DiD Estimate - Period × Deaths	0.0001 (0.0002)		0.0003 (0.0004)		0.0001 (0.0002)
DiD Estimate - Period × 5 Year Average Deaths		0.0002 (0.0003)		0.002** (0.001)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0669	0.0672	0.1619	0.1631	0.0684
	Specification				
	(6)	(7)	(8)	(9)	(10)
DiD Estimate - Period × Above Average Deaths (Dummy)		0.0003 (0.0004)		0.0004 (0.001)	
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)	0.001* (0.0003)		0.002** (0.001)		0.001 (0.001)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0687	0.1641	0.1653	0.1656	0.1669

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Columns 9 and 10 report the results of the specifications that examine the gender heterogeneity of the conflict impact. These specifications include an interaction term between the difference in difference variable and a male indicator to account

for gender heterogeneity. Our results for these specifications show that the coefficient capturing the difference-in-differences estimate is not significant. However, the coefficient on the interaction term between the difference in difference variable and a male indicator is positive and significant. According to these results, male students from municipalities affected by conflict have an additional improvement (gain) of 0.0015 standard deviations in the language examination Z-Score per additional conflict-related death. Similarly, male students from conflict affected municipalities have an additional improvement (gain) of 0.003 standard deviations in the language examination Z-Score per additional conflict-related death in the analysis of the 5-year average of conflict-related deaths.

11.2 Presence of Conflict

Mathematics

This sub-section presents the results from estimating our difference-in-differences model for the mathematics exit examination using presence of conflict as our variable of interest. As mentioned above, the presence of the conflict is quantified using two different measures: **(i)** a dummy variable that takes a value of 1 if the municipality recorded conflict events above the national average, and **(ii)** a positive integers variable that captures the number of conflict events in each municipality.

Table 17 reports the results of the difference in differences regressions for the specifications that use our first measure of presence of conflict, above average conflict events, as variable of interest. This table includes the results for specifications that use a contemporaneous measure of presence (presence of conflict in the year when the student took the examination) and specifications that use a 5-year average measure of presence of conflict. In this table, we sequentially introduce various controls for individual/parental and school characteristics. *Columns 1 and 2* report the results of the specifications that did not include any controls, *Columns 3 and 4* report the results of the specifications that included controls for individual and parental characteristics, *Columns 5 and 6* show the results of the specifications that controls for school characteristics, *Columns 7 and 8* report the results of the specifications that included individual/parental and school controls, and *Columns 9 and 10* report the results of the specifications that include all controls and an interaction term to account for heterogeneous gender effects. Observed individual/parental characteristics include the age of the student, the mother's level of education, and the father's level of education. Observed school

characteristics include a dummy indicating whether the school is public and the level of tuition that students have to pay in order to attend.

Table 17. Differences in Differences – Presence of Conflict Binary Version (Binary Treatment Variable) - Mathematics Examination					
	Specification				
	(1)	(2)	(3)	(4)	(5)
DiD Estimate - Events × Deaths	0.026 (0.018)		0.087 (0.095)		0.034* (0.018)
DiD Estimate - Period × 5 Year Average Events		0.029 (0.020)		0.045 (0.082)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0371	0.0371	0.1309	0.1308	0.0387
	Specification				
	(6)	(7)	(8)	(9)	(10)
DiD Estimate - Period × Above Average Deaths (Dummy)		0.089 (0.096)		0.010 (0.098)	
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)	0.040* (0.020)		0.055 (0.084)		0.019 (0.087)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0387	0.1360	0.1359	0.1393	0.1387

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

All the results reported in *Table 17* include student and municipal fixed effects. Additionally, all the specifications control for the size of the municipality using the logarithm of the population. The difference-in-differences estimate is the coefficient on the interaction between the dummy indicating the period and the dummy that indicates whether a student comes from a municipality that has reported conflict events above the national average. This coefficient captures the gains in achievement in the exit examinations for students coming from above average conflict zones. In this case we categorize students as coming from conflict zones if they took the exit examination at a municipality that reported more

conflict events than the national average. The coefficient on the interaction is positive and significant in only two of the specifications considered in *Table 18*. The only specifications in which the difference-in-differences estimate turned out to be significant are the specifications that incorporated school controls.

The complete results from the econometric exercise, including the coefficients for all the controls, are included in *Appendix 2 (Table 33)*. *Columns 5 and 6* present the results for the specifications that only control for school characteristics. In these specifications, the coefficient capturing the difference-in-differences estimate is significant and positive. Results show that students from municipalities that have an above national average presence of conflict have an improvement (gain) of 0.034 standard deviations in the mathematics examination Z-Score. Similarly, students from municipalities that have a 5-year above national average presence of conflict have an improvement (gain) of 0.04 standard deviations in the mathematics examination Z-Score. It is also important to highlight that as we can see in *Table 32 in Appendix 2*, the tuition paid in the educational institutions is positively and significantly related to gains in achievement in the standardized examinations. In this specification, the public education dummy is also positive and significant.

Columns 9 and 10 show the results of the specifications that include an interaction to account for gender heterogeneity of the conflict impact. More specifically, in these columns we present the results of the specifications in which we interact the difference in difference variable with a male indicator. The coefficient of this interacted term is positive and significant. These results suggest that the gains in achievement are significantly stronger for male students than for female students in the sample. Male students from municipalities that have an above national average presence of conflict have an additional improvement (gain) of 0.164 standard deviations in the mathematics examination Z-Score. Similarly, male students from municipalities that have a 5-year above national average presence of conflict have an additional improvement (gain) of 0.152 standard deviations in the mathematics examination Z-Score.

Table 18 contains the results of the difference in differences regressions for the specifications that use our second measure of presence of conflict, number of conflict events, as variable of interest. This table includes the results for specifications that use a contemporaneous measure of presence of conflict (number of conflict events in the year when the student took the examination) and specifications that use a 5-year average measure of presence of conflict. All the results reported include student and municipal fixed effects. Also, all the specifications control for the size of the municipality using the logarithm of the population. In this case, the difference-in-differences estimate is the coefficient on

the interaction between the dummy indicating the period and the variable that indicates the number of conflict events in the municipality where the student took the exit examination. The coefficient on the interaction is positive and significant in all the specifications considered in *Table 18*. The results suggest that students coming from conflict zones have significantly greater gains in achievement than those students coming from non conflict zones.

Table 18. Differences in Differences – Presence of Conflict Continuous Version (Continuous Treatment Variable) - Mathematics Examination					
	Specification				
	(1)	(2)	(3)	(4)	(5)
DiD Estimate - Events × Deaths	0.001*** (0.0002)		0.004*** (0.001)		0.0014*** (0.0001)
DiD Estimate - Period × 5 Year Average Events		0.0013** * (0.0001)		0.005*** (0.001)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0377	0.0379	0.1355	0.1372	0.0398
	Specification				
	(6)	(7)	(8)	(9)	(10)
DiD Estimate - Period × Above Average Deaths (Dummy)		0.004*** (0.0008)		0.003*** (0.001)	
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)	0.0019** * (0.0002)		0.005*** (0.0009)		0.003*** (0.001)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0402	0.1405	0.1424	0.1418	0.1444

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

The complete output for these specifications, including the coefficients for all the controls, is reported in *Appendix 2 (Table 34)*. In *Columns 3 and 4* the results for the specifications that only control for student and parental characteristics are reported. In these two specifications, the coefficient capturing the difference-in-

differences estimate is positive and significant. Results show that students coming from conflict affected municipalities have an improvement (gain) of 0.004 standard deviations in the mathematics examination Z-Score per additional conflict event. Similarly, students from conflict affected municipalities have an improvement (gain) of 0.005 standard deviations in the mathematics examination Z-Score per additional conflict event in the analysis of the 5-year average of events. As we can see in *Table 34 (Appendix 2)*, it is also important to note that the categorical variables that capture the levels of education received by the parents are positively and significantly related to gains in achievement.

Columns 5 and 6 report the results for the specifications that only control for school characteristics. In these specifications, the coefficient capturing the difference-in-differences estimate is significant, and the sign of the coefficient is positive. The results reported in *Table 19* show that students coming from conflict affected municipalities have an improvement of 0.0014 standard deviations in the math examination Z-Score per additional conflict event. Similarly, students from conflict affected municipalities have an improvement of 0.002 standard deviations in the math examination Z-Score per additional conflict event in the analysis of the 5-year average of events. It should also be highlighted that the tuition paid in the educational institutions is positively and significantly related to gains in achievement in the standardized examinations (*Table 25 - Appendix 2*).

Columns 7 and 8 present the results for the specifications that control for student/parental and school characteristics. In these two specifications the coefficient capturing the difference-in-differences estimate is also positive and significant. Our results show that students coming from conflict affected municipalities have an improvement (gain) of 0.004 standard deviations in the mathematics examination Z-Score per additional conflict event. Similarly, students from conflict affected municipalities have an improvement (gain) of 0.005 standard deviations in the mathematics examination Z-Score per additional conflict event in the analysis of the 5-year average of events. Additionally, as we can see in *Table 34 (Appendix 2)*, the coefficients for the variables that capture parental education remain positive and significant.

Columns 9 and 10 explore the gender heterogeneity of the conflict impact. These columns present the results of the specifications that include an interaction term between the difference in difference variable and a male indicator. In these specifications the coefficient capturing the difference-in-differences estimate is positive and significant. Results show that students coming from conflict affected municipalities have an improvement of 0.003 standard deviations in the mathematics examination Z-Score per additional conflict event. Likewise, students from conflict affected municipalities have an improvement of 0.003 standard

deviations in the mathematics examination Z-Score per additional conflict event in the analysis of the 5-year average of events. The results obtained in this empirical exercise suggest that the gains in achievement are stronger for male students than for female students in the sample. Male students from municipalities affected by conflict have an additional improvement (gain) of 0.002 standard deviations in the mathematics examination Z-Score per additional conflict event. Similarly, male students from conflict affected municipalities have an additional improvement (gain) of 0.003 standard deviations in the mathematics examination Z-Score per additional conflict event in the analysis of the 5-year average of events.

Table 19. Differences in Differences – Presence of Conflict Discrete Version (Binary Treatment Variable) - Language Examination					
	Specification				
	(1)	(2)	(3)	(4)	(5)
DiD Estimate - Period × Above Average Events (Dummy)	0.037 (0.116)		0.038 (0.071)		0.031 (0.116)
DiD Estimate - Period × Above 5 Year Average Events (Dummy)		0.034 (0.117)		0.086 (0.062)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0670	0.0670	0.1619	0.1622	0.0684
	Specification				
	(6)	(7)	(8)	(9)	(10)
DiD Estimate - Period × Above Average Deaths (Dummy)		0.022 (0.072)		0.073 (0.324)	
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)	0.024 (0.018)		0.068 (0.063)		0.118 (0.173)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0684	0.1640	0.1642	0.1663	0.1664

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Language

This sub-section presents the difference-in-differences results for the language exit examination using presence of conflict as our variable of interest. The presence of the conflict is quantified using two different measures: (i) a dummy variable that takes a value of 1 if the municipality recorded conflict events above the national average, and (ii) a positive integers variable that captures the number of conflict events in each municipality. All the results reported include student and municipal fixed effects. Additionally, all the specifications control for the size of the municipality using the logarithm of the population. *Table 19* presents the results of the difference in differences regressions for the specifications that use our first measure of presence of conflict, above average conflict events, as variable of interest.

The coefficient capturing the difference-in-differences estimate was not significant for any of the specifications considered. However, the coefficient on the interaction term between the difference in differences variable and a male indicator in *Columns 9 and 10* turned out to be positive and significant. According to these results, male students from municipalities that have an above national average presence of conflict have an additional improvement (gain) of 0.105 standard deviations in the language examination Z-Score. Similarly, male students from municipalities that have a 5-year above national average presence of conflict have an additional improvement (gain) of 0.104 standard deviations in the language examination Z-Score.

Table 20 reports the results for the specifications that use our second measure of presence of conflict, number of conflict events, as variable of interest. This table includes the results for specifications that use a contemporaneous measure of presence of conflict (number of conflict events in the year when the student took the examination) and specifications that use a 5-year average measure of presence of conflict. All the results reported include student and municipal fixed effects. Also, all the specifications control for the size of the municipality using the logarithm of the population.

The difference-in-differences estimate is the coefficient on the interaction between the dummy indicating the period and the variable that indicates the number of conflict events in the municipality where the student took the exit examination. The coefficient on the interaction is positive and significant in five of the ten specifications considered in *Table 20*. Of those five specifications, two correspond to specifications that use the contemporaneous measure of presence of conflict and three correspond to specifications that use the 5-year average measure of presence of conflict.

Table 20. Differences in Differences – Presence of Conflict Continuous Version (Continuous Treatment Variable) - Language Examination					
	Specification				
	(1)	(2)	(3)	(4)	(5)
DiD Estimate - Period × Events	0.00002 (0.0001)		0.002*** (0.001)		0.0002 (0.0002)
DiD Estimate - Period × 5 Year Average Events		0.0002 (0.0002)		0.002** (0.001)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0669	0.0672	0.1638	0.1631	0.0684
	Specification				
	(6)	(7)	(8)	(9)	(10)
DiD Estimate - Period × Above Average Deaths (Dummy)		0.002*** (0.001)		0.001 (0.001)	
DiD Estimate - Period × Above 5 Year Average Deaths (Dummy)	0.001*** (0.00018)		0.002** (0.001)		0.001 (0.001)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0687	0.1660	0.1653	0.1680	0.1672

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Columns 3 and 4 report the results for the specifications that only control for student and parental characteristics. In these specifications, the coefficient capturing the difference-in-differences estimate is positive and significant. Students who come from conflict affected municipalities have an improvement (gain) of 0.002 standard deviations in the language examination Z-Score per additional conflict event. Similarly, students from conflict affected municipalities have an improvement (gain) of 0.002 standard deviations in the language examination Z-Score per additional conflict event in the analysis of the 5-year average of events.

Columns 5 and 6 present the results for the specifications that only control for school characteristics. In these specifications, the coefficient capturing the difference-in-differences estimate is not significant for the specification that considers the contemporaneous measure of presence of conflict. However, the coefficient for the difference-in-differences estimate is significant and positive for the specification that considers the 5-year average measure of presence of conflict. The results presented in *Table 20* show that students from conflict affected municipalities have an improvement of 0.001 standard deviations in the language examination Z-Score per additional conflict event in the analysis of the 5-year average of events.

In *Columns 7 and 8* we present the results for the specifications that control for student/parental and school characteristics. As we can see in this table, in these two specifications the coefficient capturing the difference-in-differences estimate is positive and significant. The results show that students coming from conflict affected municipalities have an improvement (gain) of 0.002 standard deviations in the language examination Z-Score per additional conflict event. Likewise, students from conflict affected municipalities have an improvement (gain) of 0.002 standard deviations in the language examination Z-Score per additional conflict event in the analysis of the 5-year average of events.

Columns 9 and 10 explore the gender heterogeneity of the conflict impact. These columns present the results of the specifications that include an interaction term between the difference in difference variable and a male indicator. In these specifications the coefficient capturing the difference-in-differences estimate is not significant. However, the coefficients of the interaction between the difference in difference variable and the male indicator are positive and significant. According to our results, male students from municipalities affected by conflict have an additional improvement (gain) of 0.002 standard deviations in the language examination Z-Score per additional conflict event. Similarly, male students from conflict affected municipalities have an additional improvement (gain) of 0.002 standard deviations in the language examination Z-Score per additional conflict event in the analysis of the 5-year average of events.

12. Discussion of the Results

The main finding of this chapter is that students who have been exposed to conflict have more significant improvements in their academic performance in comparison to those students who have not been affected. This finding is robust to the different measures of conflict (intensity and presence) that were used in the empirical exercise of this chapter. There are a number of factors that could be taken into

consideration to explain the results obtained in the econometric exercise. It is particularly important to establish the factors that could explain why those students who came from conflict zones experienced more significant achievement gains in the standardized examinations, and in doing so, help us get a better understanding of the results obtained in this chapter. There are three factors that could play an important role in the rationalization of the results obtained in the econometric exercise: **(i)** resilience, **(ii)** future expectations and **(iii)** self-selection.

The first factor that could explain the results obtained in this chapter is the resilience that students and other individuals of a community can develop when confronted with harsh conditions over a long period of time. Hopelessness and fear, amongst other tribulations associated to a civil conflict, may produce desolation and feelings of powerlessness amongst students affecting their academic performance. However, there are certain factors that can help students overcome the negative effects of risks (civil conflict), such as individual coping skills, parental support, adult mentoring and community organizations (Fergus and Zimmerman 2005). In fact, according to Waxman et al. (2003) resiliency refers to “those factors and processes that limit the negative behaviors associated with stress and result in adaptative outcomes even in the presence of adversity”. It is possible that those students who came from conflict zones might have developed certain skills that allow them to overcome numerous challenges and face hindering conditions. Furthermore, it is likely that those students from conflict zones benefit from the emergence of protective factors at the family level. It is probable that those students coming from municipalities with a high intensity and/or a high presence of conflict receive special support from their parents. It can be the case that parents are willing to invest as many resources as needed to give their kids a good education that can keep them away from the conflict. As Waxman et al. (2003) have highlighted, the literature that studies resilience has provided a framework to understand the reasons why some students who have been exposed to high levels of risk do not experience problems associated with those risks. Students, who develop coping strategies and successfully generate mechanisms to achieve their objectives, may have a special protection to exposure to a conflict-contaminated environment. In that sense, this literature has provided a possible explanation as to why students exposed to risk have better outcomes than expected (Rutter 1987) (Fergus and Zimmerman 2005).

Students who have been exposed to conflict possibly have more significant improvements in their academic performance than those students who have not been affected because of their ability to adapt to new conditions despite having witnessed traumatic events because of conflict. A second factor that could help us explain the results obtained in this chapter is the importance of future expectations. Some students who come from conflict zones have been exposed to

high degrees of conflict-related violence over a long period of time. However, the willingness of a student to learn and succeed in the educational process does not only depend on past experiences. In fact, the perspectives for the future can possibly play a greater role than past experiences in determining the levels of effort that each student sets forth. In that sense, and despite being exposed to high degrees of violence and experiencing adversity, those students coming from conflict zones might have more significant improvements in their academic performance if they have positive expectations about the future and the viability of successfully pursuing their goals.

A third factor that can explain why students who have been exposed to conflict have more significant improvements in their academic performance than those students who have not been affected is that there might be an auto-selection mechanism in action. It could be the case that those individuals that are mostly affected by conflict are the ones that do not participate in the formal education system. Those individuals that do not have the resources (monetary, psychological, family support) to overcome or adapt to the rough conditions that are prevalent in conflict areas are more likely to drop out from the formal educational system. It is likely that those individuals from conflict zones that continue their formal education have certain resources that allow them to stay in the system. There is auto-selection because those individuals with resources to overcome or adapt to the harsh conditions are most likely to continue with their studies. This second factor can be easily linked to previous research efforts that have studied the relationship between conflict, lower levels of enrollment and school drop-outs at the high school and elementary school levels. Rodriguez and Sanchez (2012) have found that armed conflict reduces the average years of schooling in 8.78% for a pooled sample of Colombian children, and it reduces the average years of schooling in 17.03% for children between sixteen and seventeen years old. Barrera and Ibañez (2004) have also found that after controlling for individual, household and other context variables, violence influences negatively school enrollment. Furthermore, Sánchez and Diaz (2005) show that municipalities which were affected by the activities of illegally armed groups, had slower growth rates of primary and secondary school enrolment compared to those in municipalities that were not affected by illegal activities using a matching mechanism. Wharton and Oyelere (2012) also show that children who live in a municipality with high levels of conflict have a gap in enrolment and accumulation.

The evidence obtained in this chapter can be used to guide the formulation of educational policy targeted specifically for the population affected by the civil conflict. Policy makers should work in the creation of new and better incentives to retain students who come from conflict zones in school with a full time dedication, and if possible, provide financial assistance so that qualified students get the

opportunity to attend university. Moreover, more governmental resources should be spent in order to meet the special needs of those students who have been affected by conflict (i.e. psychological help, school supplies, special remedial tutorials) and to supervise their progress at educational institutions. It should be noted that all these efforts not only help improve the quality of life of those individuals affected by the conflict. As Ostby (2013) has highlighted there is evidence that clearly points to a pacifying effect of education, which has positive repercussions for the entire society. Furthermore, Ostby (2013) has also emphasized that democracies, like Colombia, tend to experience a greater stabilizing effect of education than non-democracies do.

13. Concluding Remarks

Colombia has been suffering from one of the longest internal conflicts ever registered in the world. The availability of reliable information about this non-international armed conflict provides a unique opportunity for research. The objective of this chapter is to get a better understanding of an important channel that links civil conflict with economic and social development: the educational outcomes of students who come from conflict zones. More specifically, this chapter provides estimates of the effect of civil conflict on student achievement gains in standardized examinations. To the best of my knowledge, this is the first attempt to study the relationship between civil conflict and academic achievement measured by cognitive examinations at both high school and university levels.

By considering students who have been exposed to a conflict environment and students who have not been exposed to a conflict environment, we found that Colombian students who had been affected by conflict tend to have greater gains in performance in comparison to those students who have not been affected. This finding was robust to the different measures of conflict (intensity and presence) that we used in the empirical exercise of this chapter. To come up with this conclusion, we used the scores that students obtained in the Colombian high school exit examination (Saber11) and in the Colombian college exit examination (SaberPro) in order to apply a difference in differences estimation strategy (equivalent to value added specification with individual fixed effects). This estimation strategy allowed us to quantify the student's learning gains using information at two points in time. Additionally, this methodology helped us tackle issues associated to self-selection bias and sample selection bias.

There are a number of factors that could be taken into consideration to explain the results obtained in the econometric exercise. It is particularly important to

establish the factors that could explain why those students who came from conflict zones experienced more significant achievement gains in the standardized examinations, and in doing so, help us get a better understanding of the results obtained in this chapter. There are three factors that allowed us to rationalize the results obtained in the econometric exercise: **(i)** resilience, **(ii)** future expectations and **(iii)** self-selection. The literature that studies resilience provides us a framework to understand the reasons why some students who have been exposed to high levels of risk not necessarily experience problems associated with those risks (Rutter 1987) (Fergus and Zimmerman 2005) (Waxman et al. 2003). It is possible that those students who came from conflict zones might have developed certain skills and coping strategies that allow them to overcome numerous challenges and face hindering conditions. Future expectations could also play a very important role in explaining our results. The perspectives for the future can possibly play a greater role than past experiences in determining the levels of effort that each student sets forth. In that sense, and despite being exposed to high degrees of violence and experiencing adversity, those students coming from conflict zones might have more significant improvements in their academic performance if they have positive expectations about the future and the viability of successfully pursuing their goals. A third possible explanation for the results obtained in this chapter is self-selection. It is possible that those individuals who are more affected by conflict do not participate in the formal education system. Those individuals that do not have the resources (monetary, psychological, family support) to overcome or adapt to the rough conditions that are prevalent in conflict areas are more likely to drop out from the formal educational system. Auto-selection could be explaining our results because those individuals with resources to overcome or adapt to the harsh conditions are most likely to continue with their studies.

The evidence obtained in this chapter can be used to guide the formulation of educational policy targeted specifically for the population affected by the civil conflict. Policy makers should work in the creation of new and better incentives to retain students who come from conflict zones in school with a full time dedication, and if possible, provide financial assistance so that qualified students get the opportunity to attend university. Moreover, more governmental resources should be spent in order to meet the special needs of those students who have been affected by conflict (i.e. psychological help, school supplies, special remedial tutorials) and to supervise their progress at educational institutions.

Appendix 1. Descriptive Statistics

Table 21. Descriptive Statistics						
	All (1)		Students Coming from Municipalities with Above Avg. Intensity (2)		Students Coming from Municipalities with Below Avg. Intensity (3)	
	High School	University	High School	University	High School	University
Z-Score Mathematics	0.467 (1.114)	0.343 (1.058)	0.515 (1.120)	0.395 (1.059)	0.356 (1.093)	0.226 (1.044)
Z-Score Language	0.611 (0.984)	0.406 (0.952)	0.674 (0.986)	0.463 (0.943)	0.469 (0.966)	0.278 (0.959)
Age	16.64 (2.00)	23.37 (2.438)	16.64 (1.99)	23.33 (2.368)	16.64 (2.032)	23.464 (2.587)
Public Institution	0.511 (0.500)	0.392 (0.488)	0.431 (0.495)	0.359 (0.480)	0.692 (0.462)	0.467 (0.499)
Municipality Development Indicator	61.975 (15.682)	66.534 (12.893)	65.449 (14.108)	67.568 (12.800)	54.098 (16.213)	64.194 (12.797)
Tuition						
Low Tuition	0.148 (0.355)	0.312 (0.463)	0.155 (0.362)	0.285 (0.452)	0.132 (0.338)	0.371 (0.483)
Mid Tuition	0.322 (0.467)	0.491 (0.500)	0.385 (0.487)	0.494 (0.500)	0.180 (0.384)	0.485 (0.500)
High Tuition	0.050 (0.217)	0.183 (0.387)	0.040 (0.196)	0.209 (0.407)	0.072 (0.258)	0.124 (0.330)
Mother Education						
Low level of education – Partial or Complete Elementary School Education (level 2)	0.273 (0.446)	0.165 (0.371)	0.243 (0.429)	0.141 (0.348)	0.311 (0.463)	0.219 (0.414)
Mid-level of education – Partial or Complete High school and/or technical studies (level 3)	0.510 (0.500)	0.539 (0.499)	0.543 (0.498)	0.549 (0.498)	0.468 (0.499)	0.517 (0.500)
High level of education – Partial or Complete University and/or graduate school (level 4)	0.202 (0.402)	0.290 (0.454)	0.201 (0.401)	0.305 (0.460)	0.203 (0.403)	0.255 (0.436)
Father Education						
Low level of education – Partial or Complete Elementary School	0.282 (0.450)	0.195 (0.396)	0.246 (0.431)	0.162 (0.368)	0.328 (0.470)	0.271 (0.444)

Education (level 2)						
Mid-level of education – Partial or Complete High school and/or technical studies (level 3)	0.458 (0.498)	0.455 (0.498)	0.483 (0.500)	0.461 (0.499)	0.425 (0.495)	0.440 (0.496)
High level of education – Partial or Complete University and/or graduate school (level 4)	0.240 (0.427)	0.328 (0.469)	0.255 (0.436)	0.356 (0.479)	0.221 (0.415)	0.262 (0.440)
Observations	56306	56306	39072	39055	17234	17251

Source: author's calculation using information from CERAC, ICFES, and DNP. Standard deviations are presented in parenthesis.

Table 22. Descriptive Statistics						
	All (1)		Students Coming from Municipalities with Above Avg. Presence (2)		Students Coming from Municipalities with Below Avg. Presence (3)	
	High School	University	High School	University	High School	University
Z-Score Mathematics	0.467 (1.114)	0.343 (1.058)	0.514 (1.125)	0.387 (1.062)	0.320 (1.063)	0.209 (1.034)
Z-Score Language	0.611 (0.984)	0.406 (0.952)	0.668 (0.983)	0.453 (0.944)	0.435 (0.965)	0.263 (0.961)
Age	16.64 (2.00)	23.37 (2.438)	16.64 (1.977)	23.326 (2.382)	16.65 (2.070)	23.519 (2.596)
Public Institution	0.511 (0.500)	0.392 (0.488)	0.446 (0.497)	0.368 (0.482)	0.715 (0.451)	0.467 (0.499)
Municipality Development Indicator	61.975 (15.682)	66.534 (12.893)	64.675 (14.309)	67.105 (12.799)	53.581 (16.757)	64.787 (13.022)
Tuition						
Low Tuition	0.148 (0.355)	0.312 (0.463)	0.154 (0.361)	0.292 (0.455)	0.130 (0.336)	0.371 (0.483)
Mid Tuition	0.322 (0.467)	0.491 (0.500)	0.375 (0.484)	0.493 (0.500)	0.157 (0.364)	0.484 (0.500)
High Tuition	0.050 (0.217)	0.183 (0.387)	0.044 (0.205)	0.202 (0.402)	0.067 (0.251)	0.126 (0.332)
Mother Education						
Low level of education – Partial or Complete Elementary School Education (level 2)	0.273 (0.446)	0.165 (0.371)	0.244 (0.430)	0.144 (0.351)	0.354 (0.478)	0.228 (0.419)
Mid-level of education – Partial or Complete High school and/or technical studies (level 3)	0.510 (0.500)	0.539 (0.499)	0.533 (0.499)	0.545 (0.498)	0.446 (0.497)	0.520 (0.500)
High level of education – Partial or Complete University and/or graduate school (level 4)	0.202 (0.402)	0.290 (0.454)	0.210 (0.407)	0.305 (0.460)	0.181 (0.385)	0.244 (0.429)
Father Education						
Low level of education – Partial or Complete Elementary School Education (level 2)	0.282 (0.450)	0.195 (0.396)	0.248 (0.432)	0.167 (0.373)	0.377 (0.485)	0.281 (0.450)
Mid-level of education – Partial or Complete High school and/or technical studies (level 3)	0.458 (0.498)	0.455 (0.498)	0.476 (0.500)	0.458 (0.498)	0.406 (0.491)	0.443 (0.497)

High level of education – Partial or Complete University and/or graduate school (level 4)	0.240 (0.427)	0.328 (0.469)	0.258 (0.438)	0.354 (0.478)	0.188 (0.391)	0.248 (0.432)
Observations	56306	56306	42602	42442	13704	13864

Source: author's calculation using information from CERAC, ICFES, and DNP. Standard deviations are presented in parenthesis.

Appendix 2. Results Differences in Differences

Table 23. Differences in Differences – Intensity of Conflict Discrete Version (Binary Treatment Variable)					
Above Average Deaths and 5-Year Above Average Deaths					
	Dependent variable: Language				
	(1)	(2)	(3)	(4)	(5)
Period	-0.20*** (0.015)	-0.17*** (0.020)	-0.51*** (0.113)	-0.45*** (0.119)	-0.22*** (0.015)
Above Average Deaths (Dummy)	0.037 (0.071)		0.206 (0.403)		0.025 (0.072)
Period × Above Average Deaths (Dummy)	0.009 (0.012)		0.046 (0.040)		0.016 (0.012)
Period × Above Average Deaths (Dummy) × Gender					
Above 5 Year Average Deaths (Dummy)		0.125*** (0.046)		0.073 (0.149)	
Period × Above 5 Year Average Deaths (Dummy)		0.029 (0.117)		0.046 (0.055)	
Period × Above 5 Year Avg. Deaths (Dummy) × Gender					
Public School					-0.03*** (0.009)
Log (Population)	-0.248* (0.145)	-0.214 (0.145)	-0.124 (0.464)	0.071 (0.456)	-0.283** (0.146)
Municipality Dev. Indicator					
Age			0.030** (0.012)	0.030** (0.012)	
Low Tuition					0.026*** (0.009)
Mid Tuition					0.033*** (0.011)
High Tuition					0.002 (0.014)
Low level Mother Education (level 2)			0.149 (0.117)	0.150 (0.117)	
Mid level Mother Education (level 3)			0.156 (0.123)	0.155 (0.123)	
High level Mother Education (level 4)			0.081 (0.132)	0.081 (0.132)	

Low level Father Education (level 2)			0.004 (0.088)	-0.001 (0.088)	
Mid level Father Education (level 3)			-0.006 (0.091)	-0.007 (0.090)	
High level Father Education (level 4)			0.009 (0.101)	0.007 (0.101)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Charact. Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0669	0.0671	0.1621	0.1621	0.0684
Additional Specifications					
	Dependent variable: Language				
	(6)	(7)	(8)	(9)	(10)
Period	-0.19*** (0.021)	-0.51*** (0.114)	-0.43*** (0.121)	-0.51*** (0.114)	-0.43*** (0.121)
Above Average Deaths (Dummy)		0.175 (0.404)		0.162 (0.403)	
Period × Above Average Deaths (Dummy)		0.052 (0.042)		0.007 (0.045)	
Period × Above Average Deaths (Dummy) × Gender				0.095*** (0.036)	
Above 5 Year Average Deaths (Dummy)	0.117** (0.047)		0.004 (0.151)		0.002 (0.151)
Period × Above 5 Year Average Deaths (Dummy)	0.022 (0.118)		0.054 (0.057)		0.107* (0.059)
Period × Above 5 Year Avg. Deaths (Dummy) × Gender					0.112*** (0.033)
Public School	-0.03*** (0.009)	-0.041 (0.027)	-0.044 (0.027)	-0.042 (0.027)	-0.047* (0.027)
Log (Population)	-0.243* (0.146)	-0.147 (0.475)	0.037 (0.467)	-0.133 (0.474)	0.051 (0.466)
Municipality Dev. Indicator		-0.002 (0.033)	-0.013 (0.031)	-0.003 (0.033)	-0.014 (0.031)
Age		0.029** (0.012)	0.029** (0.012)	0.029** (0.012)	0.029** (0.012)

Low Tuition	0.024** (0.009)	-0.045 (0.035)	-0.051 (0.035)	-0.046 (0.035)	-0.052 (0.035)
Mid Tuition	0.032*** (0.011)	-0.024 (0.039)	-0.027 (0.039)	-0.022 (0.039)	-0.025 (0.039)
High Tuition	0.003 (0.014)	-0.036 (0.037)	-0.035 (0.037)	-0.038 (0.037)	-0.037 (0.037)
Low level Mother Education (level 2)		0.137 (0.119)	0.137 (0.119)	0.144 (0.119)	0.143 (0.119)
Mid level Mother Education (level 3)		0.141 (0.124)	0.141 (0.124)	0.149 (0.124)	0.148 (0.124)
High level Mother Education (level 4)		0.077 (0.134)	0.076 (0.134)	0.086 (0.134)	0.083 (0.134)
Low level Father Education (level 2)		0.010 (0.091)	0.007 (0.091)	0.009 (0.091)	0.004 (0.091)
Mid level Father Education (level 3)		0.001 (0.093)	-0.0001 (0.093)	-0.002 (0.093)	-0.005 (0.093)
High level Father Education (level 4)		-0.002 (0.104)	-0.004 (0.104)	-0.008 (0.104)	-0.011 (0.104)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Charact. Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0685	0.1643	0.1641	0.1658	0.1666

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 24. Differences in Differences – Intensity of Conflict Discrete Version (Binary Treatment Variable)					
10-Year Above Average Deaths and 15-Year Above Average Deaths					
	Dependent variable: Language				
	(1)	(2)	(3)	(4)	(5)
Period	-0.17*** (0.018)	-0.16*** (0.019)	-0.47*** (0.117)	-0.44*** (0.123)	-0.18*** (0.019)
Above 10 year Average Deaths (Dummy)	0.442*** (0.130)		0.149 (0.405)		0.462*** (0.133)
Period × Above 10 year Average Deaths (Dummy)	0.034 (0.116)		0.0203 (0.053)		0.026 (0.116)
Period × Above 10 year Avg. Deaths (Dummy) × Gender					
Above 15 Year Average Deaths (Dummy)		0.437*** (0.135)		0.124 (0.406)	
Period × Above 15 Year Average Deaths (Dummy)		0.036 (0.117)		0.046 (0.062)	
Period × Above 15 Year Avg. Deaths (Dummy) × Gender					
Public School					-0.03*** (0.009)
Log (Population)	-0.254* (0.144)	-0.246* (0.144)	0.023 (0.456)	0.012 (0.452)	-0.282** (0.145)
Municipality Development Indicator					
Age			0.029** (0.012)	0.029** (0.012)	
Low Tuition					0.024*** (0.009)
Mid Tuition					0.032*** (0.011)
High Tuition					0.003 (0.014)
Low level Mother Education (level 2)			0.150 (0.117)	0.149 (0.117)	
Mid level Mother Education (level 3)			0.157 (0.123)	0.156 (0.123)	
High level Mother Education (level 4)			0.081 (0.132)	0.080 (0.132)	
Low level Father Education (level 2)			0.004 (0.088)	0.0008 (0.088)	

Mid level Father Education (level 3)			-0.008 (0.090)	-0.008 (0.090)	
High level Father Education (level 4)			0.007 (0.101)	0.006 (0.101)	
Indiv./Parent Charac. Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0672	0.0672	0.1619	0.1620	0.0686
Additional Specifications					
	Dependent variable: Language				
	(6)	(7)	(8)	(9)	(10)
Period	-0.18*** (0.019)	-0.46*** (0.119)	-0.45*** (0.125)	-0.47*** (0.119)	-0.45*** (0.125)
Above 10 year Average Deaths (Dummy)		0.116 (0.405)		0.096 (0.405)	
Period × Above 10 year Average Deaths (Dummy)		0.016 (0.054)		0.068 (0.056)	
Period × Above 10 year Avg. Deaths (Dummy) × Gender				0.108** (0.033)	
Above 15 Year Average Deaths (Dummy)	0.449*** (0.047)		0.099 (0.407)		0.080 (0.406)
Period × Above 15 Year Average Deaths (Dummy)	0.028 (0.117)		0.034 (0.064)		0.082 (0.059)
Period × Above 15 Year Avg. Deaths (Dummy) × Gender					0.100*** (0.032)
Public School	-0.03*** (0.009)	-0.043 (0.027)	-0.043 (0.027)	-0.047* (0.027)	-0.048* (0.027)
Log (Population)	-0.275* (0.145)	0.007 (0.468)	-0.0022 (0.463)	0.0175 (0.467)	-0.001 (0.463)
Municipality Development Indicator		-0.011 (0.031)	-0.011 (0.031)	0.012 (0.031)	-0.011 (0.031)
Age		0.029** (0.012)	0.029** (0.012)	0.029** (0.012)	0.029** (0.012)
Low Tuition	0.024** (0.009)	-0.049 (0.035)	-0.050 (0.035)	-0.049 (0.035)	-0.050 (0.035)

Mid Tuition	0.031*** (0.011)	-0.026 (0.039)	-0.027 (0.039)	-0.024 (0.039)	-0.025 (0.039)
High Tuition	0.002 (0.014)	-0.035 (0.037)	-0.035 (0.037)	-0.037 (0.037)	-0.037 (0.037)
Low level Mother Education (level 2)		0.138 (0.119)	0.137 (0.119)	0.142 (0.119)	0.142 (0.119)
Mid level Mother Education (level 3)		0.142 (0.124)	0.141 (0.124)	0.147 (0.124)	0.147 (0.124)
High level Mother Education (level 4)		0.077 (0.134)	0.076 (0.134)	0.084 (0.134)	0.084 (0.134)
Low level Father Education (level 2)		0.006 (0.091)	0.006 (0.091)	0.004 (0.091)	0.006 (0.091)
Mid level Father Education (level 3)		-0.001 (0.093)	-0.001 (0.093)	-0.005 (0.093)	-0.003 (0.093)
High level Father Education (level 4)		-0.004 (0.104)	-0.004 (0.104)	-0.011 (0.104)	-0.009 (0.104)
Indiv./Parent Charac. Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0686	0.1640	0.1640	0.1663	0.1661

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 25. Differences in Differences – Intensity of Conflict Discrete Version (Binary Treatment Variable)					
Above Average Deaths and 5-Year Above Average Deaths					
	Dependent variable: Mathematics				
	(1)	(2)	(3)	(4)	(5)
Period	-0.24*** (0.017)	-0.25*** (0.023)	-0.115 (0.149)	-0.181 (0.158)	-0.25*** (0.018)
Above Average Deaths (Dummy)	0.162** (0.081)		0.601 (0.535)		0.145* (0.083)
Period × Above Average Deaths (Dummy)	0.028 (0.114)		0.134** (0.054)		0.037 (0.114)
Period × Above Average Deaths (Dummy) × Gender					
Above 5 Year Average Deaths (Dummy)		0.100* (0.053)		0.198 (0.197)	
Period × Above 5 Year Average Deaths (Dummy)		0.033 (0.120)		0.170** (0.073)	
Period × Above 5 Year Average Deaths (Dummy) × Gender					
Public School					0.021** (0.010)
Log (Population)	1.351*** (0.165)	-0.25*** (0.023)	-1.74*** (0.615)	-1.47** (0.604)	1.312*** (0.166)
Municipality Development Indicator					
Age			0.021 (0.016)	0.021 (0.016)	
Low Tuition					0.059*** (0.011)
Mid Tuition					0.058*** (0.013)
High Tuition					0.002 (0.016)
Low level Mother Education (level 2)			0.259* (0.155)	0.267* (0.155)	
Mid level Mother Education (level 3)			0.391** (0.163)	0.400** (0.163)	
High level Mother Education (level 4)			0.387** (0.175)	0.399** (0.175)	

Low level Father Education (level 2)			0.110 (0.117)	0.097 (0.117)	
Mid level Father Education (level 3)			0.215* (0.120)	0.205* (0.120)	
High level Father Education (level 4)			0.202 (0.134)	0.193 (0.134)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0372	0.0371	0.1320	0.1315	0.0388
Additional Specifications					
	Dependent variable: Mathematics				
	(6)	(7)	(8)	(9)	(10)
Period	-0.27*** (0.024)	-0.126 (0.152)	-0.191 (0.161)	-0.125 (0.152)	-0.191 (0.161)
Above Average Deaths (Dummy)		0.506 (0.536)		0.527 (0.536)	
Period × Above Average Deaths (Dummy)		0.145*** (0.055)		0.073 (0.060)	
Period × Above Average Deaths (Dummy) × Gender				0.150*** (0.048)	
Above 5 Year Average Deaths (Dummy)	0.100* (0.059)		0.231 (0.201)		0.233 (0.201)
Period × Above 5 Year Average Deaths (Dummy)	0.047 (0.120)		0.176** (0.075)		0.103 (0.078)
Period × Above 5 Year Average Deaths (Dummy) × Gender					0.154*** (0.043)
Public School	0.022** (0.010)	0.046 (0.036)	0.046 (0.036)	0.044 (0.036)	0.042 (0.036)
Log (Population)	1.389*** (0.167)	-1.84*** (0.631)	-1.558** (0.621)	-1.82*** (0.630)	-1.539** (0.620)
Municipality Development Indicator		0.060 (0.043)	0.044 (0.043)	0.058 (0.044)	0.043 (0.043)

Age		0.023 (0.016)	0.022 (0.016)	0.022 (0.016)	0.022 (0.016)
Low Tuition	0.059*** (0.011)	0.077* (0.046)	0.080* (0.046)	0.077* (0.046)	0.080* (0.046)
Mid Tuition	0.059*** (0.013)	0.006 (0.052)	0.010 (0.052)	0.010 (0.052)	0.0131 (0.052)
High Tuition	0.005 (0.016)	0.063 (0.049)	0.074 (0.049)	0.061 (0.049)	0.071 (0.049)
Low level Mother Education (level 2)		0.289* (0.158)	0.299* (0.158)	0.301* (0.158)	0.307** (0.158)
Mid level Mother Education (level 3)		0.426** (0.165)	0.436*** (0.165)	0.439*** (0.165)	0.446*** (0.165)
High level Mother Education (level 4)		0.413** (0.178)	0.426** (0.178)	0.428** (0.178)	0.437** (0.178)
Low level Father Education (level 2)		0.078 (0.121)	0.066 (0.121)	0.077 (0.121)	0.062 (0.121)
Mid level Father Education (level 3)		0.194 (0.124)	0.186 (0.124)	0.189 (0.124)	0.179 (0.124)
High level Father Education (level 4)		0.206 (0.138)	0.198 (0.138)	0.196 (0.138)	0.188 (0.138)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0387	0.1373	0.1367	0.1395	0.1395

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 26. Differences in Differences – Intensity of Conflict Discrete Version (Binary Treatment Variable) 10-Year Above Average Deaths and 15-Year Above Average Deaths					
	Dependent variable: Mathematics				
	(1)	(2)	(3)	(4)	(5)
Period	-0.24*** (0.021)	-0.23*** (0.021)	-0.164 (0.155)	-0.109 (0.164)	-0.26*** (0.022)
Above 10 Year Average Deaths (Dummy)	0.086 (0.148)		0.566 (0.536)		0.108 (0.151)
Period × Above 10 Year Average Deaths (Dummy)	0.029 (0.118)		0.167** (0.070)		0.041 (0.118)
Period × Above 10 Yr. Avg. Deaths (Dummy) × Gender					
Above 15 Year Average Deaths (Dummy)		0.125 (0.154)		0.641 (0.539)	
Period × Above 15 Year Average Deaths (Dummy)		0.020 (0.119)		0.080 (0.083)	
Period × Above 15-Yr Avg. Deaths (Dummy) × Gender					
Public School					0.022** (0.010)
Log (Population)	1.384*** (0.165)	1.378*** (0.165)	-1.6*** (0.605)	-1.41** (0.599)	1.355*** (0.166)
Municipality Development Indicator					
Age			0.020 (0.016)	0.021 (0.016)	
Low Tuition					0.060*** (0.011)
Mid Tuition					0.059*** (0.013)
High Tuition					0.004 (0.016)
Low level Mother Education (level 2)			0.269* (0.155)	0.265* (0.155)	
Mid level Mother Education (level 3)			0.402** (0.163)	0.398** (0.163)	
High level Mother Education (level 4)			0.397** (0.175)	0.393** (0.175)	
Low level Father Education			0.101	0.101	

(level 2)			(0.117)	(0.117)	
Mid level Father Education (level 3)			0.206* (0.120)	0.208* (0.120)	
High level Father Education (level 4)			0.193 (0.134)	0.195 (0.134)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
# of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0371	0.0371	0.1319	0.1309	0.0387
Additional Specifications					
	Dependent variable: Mathematics				
	(6)	(7)	(8)	(9)	(10)
Period	-0.26*** (0.022)	-0.171 (0.158)	-0.124 (0.167)	-0.175 (0.158)	-0.129 (0.166)
Above 10 Year Average Deaths (Dummy)		0.473 (0.538)		0.505 (0.537)	
Period × Above 10 Year Average Deaths (Dummy)		0.172** (0.072)		0.092 (0.075)	
Period × Above 10 Yr. Avg. Deaths (Dummy) × Gender				0.166*** (0.043)	
Above 15 Year Average Deaths (Dummy)	0.117 (0.156)		0.539 (0.541)		0.574 (0.539)
Period × Above 15 Year Average Deaths (Dummy)	0.033 (0.119)		0.095 (0.085)		0.008 (0.087)
Period × Above 15-Yr Avg. Deaths (Dummy) × Gender					0.180*** (0.042)
Public School	0.022** (0.010)	0.045 (0.036)	0.043 (0.036)	0.039 (0.036)	0.035 (0.036)
Log (Population)	1.346*** (0.166)	-1.69*** (0.622)	-1.50** (0.616)	-1.67*** (0.621)	-1.499** (0.615)
Municipality Development Indicator		0.055 (0.044)	0.040 (0.043)	0.054 (0.044)	0.040 (0.043)
Age		0.021 (0.016)	0.022 (0.016)	0.022 (0.0161)	0.023 (0.016)
Low Tuition	0.059*** (0.011)	0.081* (0.046)	0.075 (0.047)	0.081* (0.046)	0.075 (0.046)

Mid Tuition	0.058*** (0.013)	0.009 (0.052)	0.007 (0.052)	0.013 (0.052)	0.011 (0.052)
High Tuition	0.004 (0.016)	0.070 (0.049)	0.071 (0.049)	0.067 (0.049)	0.068 (0.049)
Low level Mother Education (level 2)		0.300* (0.158)	0.297* (0.158)	0.306** (0.157)	0.305* (0.157)
Mid level Mother Education (level 3)		0.438*** (0.165)	0.435*** (0.165)	0.447*** (0.165)	0.446*** (0.165)
High level Mother Education (level 4)		0.424** (0.178)	0.421** (0.178)	0.435** (0.177)	0.434 (0.177)
Low level Father Education (level 2)		0.068 (0.121)	0.069 (0.121)	0.064 (0.121)	0.068 (0.121)
Mid level Father Education (level 3)		0.185 (0.124)	0.187 (0.124)	0.179 (0.124)	0.183 (0.124)
High level Father Education (level 4)		0.196 (0.138)	0.198 (0.138)	0.186 (0.138)	0.190 (0.138)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
# of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0386	0.1370	0.1361	0.1402	0.1401

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 27. Differences in Differences – Intensity of Conflict Continuous Version (Continuous Treatment Variable)					
Above Average Deaths and 5-Year Above Average Deaths					
	Dependent variable: Language				
	(1)	(2)	(3)	(4)	(5)
Period	-0.20*** (0.011)	-0.20*** (0.012)	-0.49*** (0.111)	-0.53*** (0.111)	-0.21*** (0.012)
Deaths	0.006 (0.007)		0.014 (0.031)		0.005 (0.007)
Period × Deaths	0.0001 (0.0002)		0.0003 (0.0004)		0.0001 (0.0002)
Period × Deaths × Gender					
5 Year Average Deaths		0.057*** (0.014)		0.052 (0.056)	
Period × 5 Year Average Deaths		0.0002 (0.0003)		0.002** (0.001)	
Period × 5 Year Average Deaths × Gender					
Public School					-0.032*** (0.009)
Log (Population)	-0.247* (0.147)	-0.233 (0.146)	-0.063 (0.461)	-0.103 (0.454)	-0.284* (0.148)
Municipality Development Indicator					
Age			0.029** (0.012)	0.029** (0.012)	
Low Tuition					0.026*** (0.010)
Mid Tuition					0.033*** (0.011)
High Tuition					0.003 (0.014)
Low level Mother Education (level 2)			0.150 (0.117)	0.149 (0.117)	
Mid level Mother Education (level 3)			0.156 (0.123)	0.154 (0.123)	
High level Mother Education (level 4)			0.081 (0.132)	0.079 (0.132)	
Low level Father Education (level 2)			0.003 (0.088)	0.009 (0.088)	

Mid level Father Education (level 3)			-0.006 (0.091)	-0.001 (0.090)	
High level Father Education (level 4)			0.008 (0.101)	0.015 (0.101)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
# of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0669	0.0672	0.1619	0.1631	0.0684
Additional Specifications					
	Dependent variable: Language				
	(6)	(7)	(8)	(9)	(10)
Period	-0.22*** (0.013)	-0.49*** (0.112)	-0.52*** (0.113)	-0.48*** (0.112)	-0.52*** (0.113)
Deaths		0.011 (0.031)		0.010 (0.031)	
Period × Deaths		0.0003 (0.0004)		0.0004 (0.001)	
Period × Deaths × Gender				0.0015*** (0.001)	
5 Year Average Deaths	0.057*** (0.014)		0.035 (0.057)		0.032 (0.057)
Period × 5 Year Average Deaths	0.001* (0.0003)		0.002** (0.001)		0.001 (0.001)
Period × 5 Year Average Deaths × Gender					0.003*** (0.001)
Public School	-0.03*** (0.009)	-0.041 (0.027)	-0.041 (0.027)	-0.044 (0.027)	-0.044 (0.027)
Log (Population)	-0.294** (0.148)	-0.087 (0.473)	-0.129 (0.465)	-0.072 (0.473)	-0.108 (0.465)
Municipality Development Indicator		-0.005 (0.033)	-0.001 (0.033)	-0.006 (0.031)	-0.003 (0.033)
Age		0.029** (0.012)	0.029** (0.012)	0.029** (0.012)	0.028** (0.012)
Low Tuition	0.028*** (0.010)	-0.045 (0.035)	-0.039 (0.035)	-0.049 (0.035)	-0.042 (0.035)
Mid Tuition	0.035*** (0.011)	-0.025 (0.039)	-0.017 (0.039)	-0.026 (0.039)	-0.018 (0.039)

High Tuition	0.003 (0.014)	-0.036 (0.037)	-0.032 (0.037)	-0.040 (0.037)	-0.037 (0.037)
Low level Mother Education (level 2)		0.137 (0.119)	0.136 (0.119)	0.143 (0.119)	0.144 (0.118)
Mid level Mother Education (level 3)		0.141 (0.124)	0.138 (0.124)	0.148 (0.124)	0.146 (0.124)
High level Mother Education (level 4)		0.077 (0.134)	0.073 (0.134)	0.085 (0.134)	0.084 (0.134)
Low level Father Education (level 2)		0.009 (0.091)	0.017 (0.091)	0.006 (0.091)	0.010 (0.091)
Mid level Father Education (level 3)		0.001 (0.093)	0.007 (0.093)	-0.005 (0.093)	-0.0003 (0.093)
High level Father Education (level 4)		-0.003 (0.104)	0.005 (0.104)	-0.010 (0.104)	-0.003 (0.104)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
# of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0687	0.1641	0.1653	0.1656	0.1669

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 28. Differences in Differences – Intensity of Conflict Continuous Version (Continuous Treatment Variable) 10-Year Above Average Deaths and 15-Year Above Average Deaths					
	Dependent variable: Language				
	(1)	(2)	(3)	(4)	(5)
Period	-0.21*** (0.012)	-0.21*** (0.011)	-0.54*** (0.111)	-0.53*** (0.111)	-0.23*** (0.013)
10 Year Average Deaths	0.124*** (0.028)		0.094 (0.093)		0.131*** (0.029)
Period × 10 Year Average Deaths	0.001** (0.0003)		0.003*** (0.001)		0.001*** (0.0004)
Period × 10 Year Average Deaths × Gender					
15 Year Average Deaths		0.125*** (0.031)		0.080 (0.095)	
Period × 15 Year Average Deaths		0.001*** (0.0002)		0.002*** (0.001)	
Period × 15 Year Average Deaths × Gender					
Public School					-0.032*** (0.009)
Log (Population)	-0.248* (0.145)	-0.262* (0.145)	-0.045 (0.452)	-0.030 (0.451)	-0.295** (0.146)
Municipality Development Indicator					
Age			0.029** (0.012)	0.029** (0.012)	
Low Tuition					0.029*** (0.010)
Mid Tuition					0.037*** (0.011)
High Tuition					0.005 (0.014)
Low level Mother Education			0.147 (0.117)	0.147 (0.117)	
Mid level Mother Education			0.152 (0.123)	0.152 (0.123)	
High level Mother Education			0.076 (0.132)	0.076 (0.132)	
Low level Father Education			0.010	0.011	

			(0.088)	(0.088)	
Mid level Father Education			-0.001 (0.090)	0.0001 (0.090)	
High level Father Education			0.016 (0.101)	0.017 (0.101)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
# of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0673	0.0673	0.1641	0.1640	0.0688
Additional Specifications					
	Dependent variable: Language				
	(6)	(7)	(8)	(9)	(10)
Period	-0.22*** (0.012)	-0.53*** (0.113)	-0.53*** (0.112)	-0.53*** (0.113)	-0.53*** (0.112)
10 Year Average Deaths		0.062 (0.094)		0.058 (0.094)	
Period × 10 Year Average Deaths		0.003*** (0.001)		0.002 (0.001)	
Period × 10 Year Average Deaths × Gender				0.003*** (0.001)	
15 Year Average Deaths	0.135*** (0.032)		0.054 (0.095)		0.049 (0.095)
Period × 15 Year Average Deaths	0.001*** (0.0002)		0.002*** (0.001)		0.001 (0.001)
Period × 15 Year Average Deaths × Gender					0.002** (0.001)
Public School	-0.03*** (0.009)	-0.042 (0.027)	-0.041 (0.027)	-0.044* (0.027)	-0.044* (0.027)
Log (Population)	-0.301** (0.146)	-0.076 (0.463)	-0.053 (0.463)	-0.053 (0.463)	-0.033 (0.462)
Municipality Development Indicator		-0.004 (0.032)	-0.005 (0.032)	-0.005 (0.032)	-0.007 (0.032)
Age		0.029** (0.012)	0.029** (0.012)	0.028** (0.012)	0.028** (0.012)
Low Tuition	0.029*** (0.010)	-0.037 (0.035)	-0.036 (0.035)	-0.040 (0.035)	-0.039 (0.035)

Mid Tuition	0.037*** (0.011)	-0.011 (0.039)	-0.010 (0.039)	-0.012 (0.039)	-0.012 (0.039)
High Tuition	0.007 (0.014)	-0.026 (0.037)	-0.025 (0.037)	-0.032 (0.037)	-0.031 (0.037)
Low level Mother Education		0.134 (0.119)	0.134 (0.119)	0.142 (0.118)	0.142 (0.118)
Mid level Mother Education		0.135 (0.124)	0.136 (0.124)	0.143 (0.124)	0.143 (0.124)
High level Mother Education		0.070 (0.134)	0.070 (0.134)	0.079 (0.133)	0.078 (0.133)
Low level Father Education		0.018 (0.091)	0.019 (0.091)	0.011 (0.091)	0.011 (0.091)
Mid level Father Education		0.008 (0.093)	0.009 (0.093)	0.001 (0.093)	0.001 (0.093)
High level Father Education		0.007 (0.104)	0.008 (0.104)	-0.001 (0.104)	-0.001 (0.104)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
# of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0689	0.1663	0.1662	0.1678	0.1676

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 29. Differences in Differences – Intensity of Conflict Continuous Version (Continuous Treatment Variable) Above Average Deaths and 5-Year Above Average Deaths					
	Dependent variable: Mathematics				
	(1)	(2)	(3)	(4)	(5)
Period	-0.239*** (0.013)	-0.27*** (0.014)	-0.120 (0.146)	-0.156 (0.147)	-0.26*** (0.014)
Deaths	0.015** (0.008)		0.043 (0.041)		0.015* (0.008)
Period × Deaths	0.0029*** (0.0004)		0.003*** (0.0006)		0.003*** (0.0002)
Period × Deaths × Gender					
5 Year Average Deaths		0.019 (0.016)		0.024 (0.075)	
Period × 5 Year Average Deaths		0.003*** (0.0003)		0.006 (0.111)	
Period × 5 Year Average Deaths × Gender					
Public School					0.027*** (0.010)
Log (Population)	1.04*** (0.168)	1.118*** (0.167)	-2.09*** (0.610)	-1.77*** (0.600)	0.977*** (0.169)
Municipality Development Indicator					
Age			0.020 (0.016)	0.020 (0.016)	
Low Tuition					0.078*** (0.011)
Mid Tuition					0.072*** (0.013)
High Tuition					0.023 (0.016)
Low level Mother Education (level 2)			0.254* (0.155)	0.257* (0.155)	
Mid level Mother Education (level 3)			0.377** (0.162)	0.385** (0.162)	
High level Mother Education (level 4)			0.378** (0.175)	0.379** (0.175)	
Low level Father Education			0.125	0.130	

(level 2)			(0.117)	(0.117)	
Mid level Father Education (level 3)			0.228* (0.120)	0.230* (0.120)	
High level Father Education (level 4)			0.212 (0.133)	0.223* (0.133)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
# of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0391	0.0387	0.1372	0.1369	0.0410
Additional Specifications					
	Dependent variable: Mathematics				
	(6)	(7)	(8)	(9)	(10)
Period	-0.30*** (0.015)	-0.141 (0.149)	-0.168 (0.150)	-0.138 (0.149)	-0.166 (0.150)
Deaths		0.0345 (0.149)		0.036 (0.041)	
Period × Deaths		0.0035*** (0.001)		0.0024*** (0.001)	
Period × Deaths × Gender				0.0023*** (0.0001)	
5 Year Average Deaths	0.018 (0.016)		0.011 (0.075)		0.015 (0.075)
Period × 5 Year Average Deaths	0.004*** (0.0003)		0.006 (0.111)		0.004*** (0.001)
Period × 5 Year Average Deaths × Gender					0.005*** (0.001)
Public School	0.026** (0.010)	0.055 (0.036)	0.046 (0.036)	0.051 (0.036)	0.042 (0.036)
Log (Population)	1.003*** (0.168)	-2.225*** (0.626)	-1.86*** (0.617)	-2.202*** (0.626)	-1.83*** (0.617)
Municipality Development Indicator		0.087** (0.0437)	0.066 (0.043)	0.085** (0.044)	0.064 (0.043)
Age		0.022 (0.016)	0.021 (0.016)	0.021 (0.016)	0.021 (0.016)
Low Tuition	0.080*** (0.011)	0.098** (0.046)	0.097** (0.046)	0.092** (0.046)	0.092** (0.046)

Mid Tuition	0.077*** (0.013)	0.005 (0.052)	0.026 (0.052)	0.004 (0.052)	0.025 (0.052)
High Tuition	0.004 (0.016)	0.052 (0.049)	0.074 (0.049)	0.046 (0.049)	0.066 (0.049)
Low level Mother Education (level 2)		0.285* (0.157)	0.288* (0.157)	0.293* (0.157)	0.300* (0.157)
Mid level Mother Education (level 3)		0.412** (0.165)	0.419** (0.165)	0.423*** (0.165)	0.432*** (0.165)
High level Mother Education (level 4)		0.404** (0.023)	0.404** (0.177)	0.416** (0.165)	0.419** (0.177)
Low level Father Education (level 2)		0.093 (0.121)	0.099 (0.121)	0.088 (0.121)	0.089 (0.121)
Mid level Father Education (level 3)		0.206* (0.124)	0.210* (0.124)	0.198* (0.124)	0.199* (0.124)
High level Father Education (level 4)		0.215 (0.137)	0.226* (0.138)	0.204 (0.137)	0.213 (0.137)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
# of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0411	0.1433	0.1421	0.1455	0.1444

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 30. Differences in Differences – Intensity of Conflict Continuous Version (Continuous Treatment Variable) 10-Year Above Average Deaths and 15-Year Above Average Deaths					
	Dependent variable: Mathematics				
	(1)	(2)	(3)	(4)	(5)
Period	-0.257*** (0.014)	-0.25*** (0.013)	-0.133 (0.147)	-0.123 (0.147)	-0.28*** (0.015)
10 Year Average Deaths	0.037 (0.032)		0.032 (0.123)		0.042 (0.033)
Period × 10 Year Average Deaths	0.003*** (0.0004)		0.006 (0.111)		0.004*** (0.0004)
Period × 10 Year Average Deaths × Gender					
15 Year Average Deaths		0.049 (0.035)		0.082 (0.125)	
Period × 15 Year Average Deaths		0.0024*** (0.0003)		0.004*** (0.001)	
Period × 15 Year Average Deaths × Gender					
Public School					0.025** (0.010)
Log (Population)	1.246*** (0.166)	1.247*** (0.166)	-1.554** (0.598)	-1.499** (0.597)	1.175*** (0.167)
Municipality Development Indicator					
Age			0.021 (0.016)	0.020 (0.016)	
Low Tuition					0.073*** (0.011)
Mid Tuition					0.073*** (0.013)
High Tuition					0.009 (0.016)
Low level Mother Education (level 2)			0.254* (0.155)	0.255* (0.155)	
Mid level Mother Education (level 3)			0.386** (0.163)	0.386** (0.163)	
High level Mother Education (level 4)			0.377** (0.175)	0.377** (0.175)	
Low level Father Education			0.121	0.124	

(level 2)			(0.117)	(0.117)	
Mid level Father Education (level 3)			0.222* (0.120)	0.225* (0.120)	
High level Father Education (level 4)			0.217* (0.134)	0.218* (0.134)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
# of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0381	0.0385	0.1355	0.1357	0.0402
Additional Specifications					
	Dependent variable: Mathematics				
	(6)	(7)	(8)	(9)	(10)
Period	-0.27*** (0.014)	-0.141 (0.150)	-0.132 (0.149)	-0.139 (0.149)	-0.129 (0.149)
10 Year Average Deaths	0.055 (0.037)	0.011 (0.124)		0.017 (0.124)	
Period × 10 Year Average Deaths	0.003*** (0.0002)	0.006 (0.111)		0.004** (0.001)	
Period × 10 Year Average Deaths × Gender				0.005*** (0.002)	
15 Year Average Deaths			0.060 (0.126)		0.068 (0.126)
Period × 15 Year Average Deaths			0.004*** (0.001)		0.002** (0.001)
Period × 15 Year Average Deaths × Gender					0.003*** (0.001)
Public School	0.025** (0.010)	0.043 (0.036)	0.044 (0.036)	0.039 (0.036)	0.039 (0.036)
Log (Population)	1.194*** (0.167)	-1.642*** (0.616)	-1.590** (0.614)	-1.608*** (0.615)	-1.558** (0.614)
Municipality Development Indicator		0.052 (0.043)	0.049 (0.043)	0.050 (0.043)	0.046 (0.043)
Age		0.022 (0.016)	0.021 (0.016)	0.021 (0.016)	0.021 (0.016)
Low Tuition	0.070*** (0.011)	0.093** (0.046)	0.095** (0.046)	0.088* (0.046)	0.089* (0.046)

Mid Tuition	0.072*** (0.013)	0.030 (0.052)	0.033 (0.052)	0.028 (0.052)	0.030 (0.052)
High Tuition	0.014 (0.016)	0.084* (0.049)	0.086* (0.049)	0.074 (0.049)	0.075 (0.049)
Low level Mother Education (level 2)		0.285* (0.157)	0.286* (0.157)	0.297* (0.157)	0.297* (0.157)
Mid level Mother Education (level 3)		0.420** (0.165)	0.420** (0.165)	0.431*** (0.165)	0.431*** (0.165)
High level Mother Education (level 4)		0.402** (0.177)	0.401** (0.177)	0.415** (0.177)	0.415** (0.177)
Low level Father Education (level 2)		0.090 (0.121)	0.093 (0.121)	0.080 (0.121)	0.081 (0.121)
Mid level Father Education (level 3)		0.204* (0.124)	0.206* (0.124)	0.193 (0.124)	0.194 (0.124)
High level Father Education (level 4)		0.220* (0.138)	0.222* (0.138)	0.209 (0.138)	0.209 (0.138)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
# of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0404	0.1401	0.1407	0.1425	0.1428

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 31. Differences in Differences – Presence of Conflict Binary Version (Binary Treatment Variable) Above Average Deaths and 5-Year Above Average Deaths					
	Dependent variable: Language				
	(1)	(2)	(3)	(4)	(5)
Period	-0.16*** (0.018)	-0.16*** (0.020)	-0.45*** (0.128)	-0.40*** (0.125)	-0.18*** (0.019)
Above Average Events (Dummy)	0.006 (0.059)		0.129 (0.408)		0.008 (0.059)
Period × Above Average Events (Dummy)	0.037 (0.116)		0.038 (0.071)		0.031 (0.116)
Period × Above Average Events (Dummy) × Gender					
Above 5 Year Average Events (Dummy)		0.138* (0.079)		0.084 (0.406)	
Period × Above 5 Year Average Events (Dummy)		0.034 (0.117)		0.086 (0.062)	
Period × Above 5 Yr. Avg. Events (Dummy) × Gender					
Public School					-0.033*** (0.009)
Log (Population)	-0.247* (0.145)	-0.257* (0.145)	0.006 (0.452)	-0.050 (0.452)	-0.275* (0.146)
Municipality Development Indicator					
Age			0.029** (0.012)	0.029** (0.012)	
Low Tuition					0.024** (0.009)
Mid Tuition					0.031*** (0.011)
High Tuition					0.0018 (0.014)
Low level Mother Education (level 2)			0.151 (0.117)	0.150 (0.117)	
Mid level Mother Education (level 3)			0.157 (0.123)	0.155 (0.123)	
High level Mother Education (level 4)			0.082 (0.132)	0.079 (0.132)	
Low level Father Education			0.0004	0.001	

(level 2)			(0.088)	(0.088)	
Mid level Father Education (level 3)			-0.008 (0.090)	-0.006 (0.090)	
High level Father Education (level 4)			0.006 (0.101)	0.007 (0.101)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
# of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0670	0.0670	0.1619	0.1622	0.0684
Additional Specifications					
	Dependent variable: Language				
	(6)	(7)	(8)	(9)	(10)
Period	-0.18*** (0.021)	-0.46*** (0.130)	-0.41*** (0.127)	-0.46*** (0.130)	-0.41*** (0.127)
Above Average Events (Dummy)		0.109 (0.409)		0.094 (0.408)	
Period × Above Average Events (Dummy)		0.022 (0.072)		0.073 (0.324)	
Period × Above Average Events (Dummy) × Gender				0.105*** (0.032)	
Above 5 Year Average Events (Dummy)	0.130* (0.079)		0.065 (0.407)		0.051 (0.406)
Period × Above 5 Year Average Events (Dummy)	0.024 (0.118)		0.068 (0.063)		0.118 (0.173)
Period × Above 5 Yr. Avg. Events (Dummy) × Gender					0.104*** (0.032)
Public School	-0.03*** (0.009)	-0.043 (0.027)	-0.044 (0.027)	-0.047* (0.027)	-0.048* (0.027)
Log (Population)	-0.279* (0.146)	-0.008 (0.463)	-0.049 (0.464)	0.0045 (0.463)	-0.038 (0.463)
Municipality Development Indicator		-0.010 (0.031)	-0.009 (0.031)	-0.0112 (0.031)	-0.0095 (0.031)
Age		0.029** (0.012)	0.029** (0.0122)	0.029** (0.012)	0.029** (0.0121)
Low Tuition	0.024** (0.009)	-0.049 (0.035)	-0.053 (0.035)	-0.049 (0.035)	-0.053 (0.035)

Mid Tuition	0.031*** (0.011)	-0.026 (0.039)	-0.029 (0.039)	-0.024 (0.039)	-0.028 (0.039)
High Tuition	0.0015 (0.014)	-0.035 (0.037)	-0.038 (0.037)	-0.038 (0.037)	-0.041 (0.037)
Low level Mother Education (level 2)		0.138 (0.119)	0.137 (0.119)	0.143 (0.119)	0.144 (0.119)
Mid level Mother Education (level 3)		0.142 (0.124)	0.140 (0.124)	0.148 (0.124)	0.148 (0.124)
High level Mother Education (level 4)		0.078 (0.134)	0.075 (0.134)	0.085 (0.134)	0.084 (0.134)
Low level Father Education (level 2)		0.007 (0.091)	0.007 (0.091)	0.006 (0.091)	0.006 (0.091)
Mid level Father Education (level 3)		-0.001 (0.093)	0.0003 (0.093)	-0.003 (0.093)	-0.002 (0.093)
High level Father Education (level 4)		-0.004 (0.104)	-0.004 (0.104)	-0.008 (0.104)	-0.008 (0.104)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
# of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0684	0.1640	0.1642	0.1663	0.1664

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 32. Differences in Differences – Presence of Conflict Binary Version (Binary Treatment Variable)					
Above Average Deaths and 5-Year Above Average Deaths					
	Dependent variable: Mathematics				
	(1)	(2)	(3)	(4)	(5)
Period	-0.24*** (0.021)	-0.24*** (0.023)	-0.12 (0.170)	-0.081 (0.166)	-0.26*** (0.022)
Above Average Events (Dummy)	0.069 (0.067)		0.629 (0.246)		0.065 (0.067)
Period × Above Average Events (Dummy)	0.026 (0.018)		0.087 (0.095)		0.034* (0.018)
Period × Above Average Events (Dummy) × Gender					
Above 5 Year Average Events (Dummy)		0.032 (0.090)		0.672 (0.539)	
Period × Above 5 Year Average Events (Dummy)		0.029 (0.020)		0.045 (0.082)	
Period × Above 5 Yr. Avg. Events (Dummy) × Gender					
Public School					0.022** (0.010)
Log (Population)	1.391*** (0.165)	1.397*** (0.165)	-1.40** (0.599)	-1.36** (0.600)	1.360*** (0.166)
Municipality Development Indicator					
Age			0.021 (0.016)	0.021 (0.016)	
Low Tuition					0.059*** (0.011)
Mid Tuition					0.058*** (0.013)
High Tuition					0.004 (0.016)
Low level Mother Education (level 2)			0.263* (0.155)	0.264* (0.155)	
Mid level Mother Education (level 3)			0.396** (0.163)	0.396** (0.163)	
High level Mother Education (level 4)			0.391** (0.175)	0.392** (0.175)	
Low level Father Education (level 2)			0.100 (0.117)	0.100 (0.117)	

Mid level Father Education (level 3)			0.207* (0.120)	0.207* (0.120)	
High level Father Education (level 4)			0.195 (0.134)	0.195 (0.134)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
# of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0371	0.0371	0.1309	0.1308	0.0387
Additional Specifications					
	Dependent variable: Mathematics				
	(6)	(7)	(8)	(9)	(10)
Period	-0.26*** (0.024)	-0.120 (0.173)	-0.091 (0.169)	-0.124 (0.172)	-0.091 (0.169)
Above Average Events (Dummy)		0.540 (0.543)		0.563 (0.542)	
Period × Above Average Events (Dummy)		0.089 (0.096)		0.010 (0.098)	
Period × Above Average Events (Dummy) × Gender				0.164*** (0.042)	
Above 5 Year Average Events (Dummy)	0.032 (0.091)		0.575 (0.541)		0.596 (0.540)
Period × Above 5 Year Average Events (Dummy)	0.040* (0.020)		0.055 (0.084)		0.019 (0.087)
Period × Above 5 Yr. Avg. Events (Dummy) × Gender					0.152*** (0.042)
Public School	0.022** (0.010)	0.043 (0.036)	0.043 (0.036)	0.037 (0.036)	0.037 (0.036)
Log (Population)	1.371*** (0.167)	-1.49** (0.616)	-1.44** (0.617)	-1.473** (0.615)	-1.43** (0.616)
Municipality Development Indicator		0.040 (0.043)	0.036 (0.043)	0.038 (0.043)	0.035 (0.043)
Age		0.022 (0.016)	0.022 (0.016)	0.022 (0.016)	0.022 (0.016)
Low Tuition	0.060*** (0.011)	0.073 (0.046)	0.073 (0.047)	0.073 (0.046)	0.072 (0.046)
Mid Tuition	0.059*** (0.013)	0.005 (0.052)	0.006 (0.052)	0.008 (0.052)	0.008 (0.052)

High Tuition	0.005 (0.016)	0.070 (0.049)	0.070 (0.049)	0.065 (0.049)	0.066 (0.049)
Low level Mother Education (level 2)		0.294* (0.158)	0.295* (0.158)	0.302* (0.158)	0.305* (0.158)
Mid level Mother Education (level 3)		0.433*** (0.165)	0.433*** (0.165)	0.442*** (0.165)	0.445*** (0.165)
High level Mother Education (level 4)		0.418** (0.178)	0.419** (0.178)	0.429** (0.178)	0.432** (0.178)
Low level Father Education (level 2)		0.068 (0.121)	0.068 (0.122)	0.066 (0.121)	0.066 (0.121)
Mid level Father Education (level 3)		0.186 (0.124)	0.186 (0.124)	0.183 (0.124)	0.183 (0.124)
High level Father Education (level 4)		0.198 (0.138)	0.198 (0.138)	0.191 (0.138)	0.191 (0.138)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
# of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0387	0.1360	0.1359	0.1393	0.1387

Source: authors calculations using information from CERAC, ICSES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 33. Differences in Differences – Presence of Conflict Continuous Version (Continuous Treatment Variable) Above Average Deaths and 5-Year Above Average Deaths					
	Dependent variable: Language				
	(1)	(2)	(3)	(4)	(5)
Period	-0.198*** (0.012)	-0.21*** (0.013)	-0.56*** (0.112)	-0.54*** (0.112)	-0.21*** (0.013)
Events	0.011* (0.006)		0.032 (0.057)		0.012** (0.006)
Period × Events	0.00002 (0.0001)		0.002*** (0.001)		0.0002 (0.0002)
Period × Events × Gender					
5 Year Average Events		0.053*** (0.013)		0.063 (0.067)	
Period × 5 Year Average Events		0.0002 (0.0002)		0.002** (0.001)	
Period × 5 Year Avg. Events × Gender					
Public School					-0.03*** (0.009)
Log (Population)	-0.231 (0.146)	-0.211 (0.145)	0.080 (0.451)	-0.028 (0.452)	-0.280* (0.147)
Municipality Development Indicator					
Age			0.029** (0.012)	0.029** (0.012)	
Low Tuition					0.027*** (0.010)
Mid Tuition					0.034*** (0.011)
High Tuition					0.002 (0.014)
Low level Mother Education (level 2)			0.149 (0.117)	0.149 (0.117)	
Mid level Mother Education (level 3)			0.154 (0.123)	0.153 (0.123)	
High level Mother Education (level 4)			0.079 (0.132)	0.079 (0.132)	
Low level Father Education (level 2)			0.0134 (0.088)	0.010 (0.088)	

Mid level Father Education (level 3)			0.002 (0.090)	0.001 (0.090)	
High level Father Education (level 4)			0.0181 (0.101)	0.0159 (0.101)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
# of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0669	0.0672	0.1638	0.1631	0.0684
Additional Specifications					
	Dependent variable: Language				
	(6)	(7)	(8)	(9)	(10)
Period	-0.23*** (0.014)	-0.55*** (0.114)	-0.53*** (0.114)	-0.55*** (0.114)	-0.53*** (0.113)
Events		0.027 (0.057)		0.023 (0.057)	
Period × Events		0.002*** (0.001)		0.001 (0.001)	
Period × Events × Gender				0.002*** (0.001)	
5 Year Average Events	0.058*** (0.013)		0.043 (0.067)		0.040 (0.067)
Period × 5 Year Average Events	0.001*** (0.00018)		0.002** (0.001)		0.001 (0.001)
Period × 5 Year Avg. Events × Gender					0.002*** (0.001)
Public School	-0.03*** (0.009)	-0.041 (0.027)	-0.041 (0.027)	-0.044 (0.027)	-0.044 (0.027)
Log (Population)	-0.248* (0.146)	0.068 (0.463)	-0.050 (0.463)	0.091 (0.463)	0.091 (0.463)
Municipality Development Indicator		-0.012 (0.032)	-0.005 (0.032)	-0.014 (0.032)	-0.007 (0.032)
Age		0.029** (0.012)	0.029 (0.012)	0.028** (0.012)	0.028** (0.012)
Low Tuition	0.029*** (0.010)	-0.037 (0.035)	-0.039 (0.035)	-0.040 (0.035)	-0.042 (0.035)
Mid Tuition	0.035*** (0.011)	-0.013 (0.039)	-0.017 (0.039)	-0.014 (0.039)	-0.017 (0.039)

High Tuition	0.001 (0.014)	-0.029 (0.037)	-0.031 (0.037)	-0.035 (0.037)	-0.037 (0.037)
Low level Mother Education (level 2)		0.136 (0.119)	0.136 (0.119)	0.144 (0.118)	0.143 (0.118)
Mid level Mother Education (level 3)		0.138 (0.124)	0.137 (0.124)	0.145 (0.124)	0.145 (0.124)
High level Mother Education (level 4)		0.073 (0.134)	0.074 (0.134)	0.082 (0.133)	0.083 (0.134)
Low level Father Education (level 2)		0.020 (0.091)	0.018 (0.091)	0.013 (0.091)	0.011 (0.091)
Mid level Father Education (level 3)		0.010 (0.093)	0.008 (0.093)	0.002 (0.093)	0.0003 (0.093)
High level Father Education (level 4)		0.008 (0.104)	0.006 (0.104)	-0.0001 (0.104)	-0.003 (0.104)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
# of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0687	0.1660	0.1653	0.1680	0.1672

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

Table 34. Differences in Differences – Presence of Conflict Continuous Version (Continuous Treatment Variable) Above Average Deaths and 5-Year Above Average Deaths					
	Dependent variable: Mathematics				
	(1)	(2)	(3)	(4)	(5)
Period	-0.25*** (0.014)	-0.268*** (0.015)	-0.181 (0.149)	-0.186 (0.148)	-0.281*** (0.015)
Events	0.007 (0.007)		0.086 (0.076)		0.007 (0.007)
Period × Events	0.001*** (0.0002)		0.004*** (0.001)		0.0014*** (0.0001)
Period × Events × Gender					
5 Year Average Events		0.017 (0.015)		0.023 (0.089)	
Period × 5 Year Average Events		0.0013*** (0.0001)		0.005*** (0.001)	
Period × 5 Yr. Avg. Events × Gender					
Public School					0.026*** (0.010)
Log (Population)	1.245*** (0.166)	1.341*** (0.165)	-1.215** (0.598)	-1.55*** (0.597)	1.156*** (0.167)
Municipality Development Indicator					
Age			0.021 (0.016)	0.020 (0.016)	
Low Tuition					0.076*** (0.011)
Mid Tuition					0.071*** (0.013)
High Tuition					0.005 (0.016)
Low level Mother Education (level 2)			0.259* (0.155)	0.256* (0.155)	
Mid level Mother Education (level 3)			0.389** (0.163)	0.385** (0.162)	
High level Mother Education (level 4)			0.383** (0.175)	0.380** (0.175)	
Low level Father Education (level 2)			0.126 (0.117)	0.133 (0.117)	

Mid level Father Education (level 3)			0.228* (0.120)	0.233** (0.120)	
High level Father Education (level 4)			0.220* (0.134)	0.226* (0.133)	
Individual/Parental Characteristics Controls	No	No	Yes	Yes	No
School Characteristics Controls	No	No	No	No	Yes
Gender Specific Effects	No	No	No	No	No
Number of Groups (Students)	56252	56252	55903	55903	56251
Observations	112504	112504	111795	111795	111795
R-Squared	0.0377	0.0379	0.1355	0.1372	0.0398
Additional Specifications					
	Dependent variable: Mathematics				
	(6)	(7)	(8)	(9)	(10)
Period	-0.307*** (0.016)	-0.189 (0.151)	-0.199 (0.151)	-0.187 (0.151)	-0.197 (0.150)
Events		0.074 (0.076)		0.078 (0.076)	
Period × Events		0.004*** (0.0008)		0.003*** (0.001)	
Period × Events × Gender				0.002** (0.001)	
5 Year Average Events	0.019 (0.015)		0.008 (0.089)		0.012 (0.089)
Period × 5 Year Average Events	0.0019*** (0.0002)		0.005*** (0.0009)		0.003*** (0.001)
Period × 5 Yr. Avg. Events × Gender					0.003*** (0.001)
Public School	0.025** (0.010)	0.045 (0.036)	0.047 (0.036)	0.042 (0.036)	0.043 (0.036)
Log (Population)	1.283*** (0.166)	-1.311** (0.615)	-1.63*** (0.614)	-1.286** (0.614)	-1.60*** (0.614)
Municipality Development Indicator		0.033 (0.043)	0.054 (0.043)	0.031 (0.043)	0.051 (0.043)
Age		0.022 (0.016)	0.022 (0.016)	0.021 (0.016)	0.021 (0.016)
Low Tuition	0.076*** (0.011)	0.090** (0.046)	0.098** (0.046)	0.087* (0.046)	0.093** (0.046)

Mid Tuition	0.072*** (0.013)	0.026 (0.052)	0.027 (0.052)	0.026 (0.052)	0.026 (0.052)
High Tuition	-0.001 (0.016)	0.077 (0.049)	0.076 (0.049)	0.072 (0.049)	0.068 (0.049)
Low level Mother Education (level 2)		0.290* (0.157)	0.287* (0.157)	0.297* (0.157)	0.297* (0.157)
Mid level Mother Education (level 3)		0.423*** (0.165)	0.419** (0.165)	0.430*** (0.165)	0.429*** (0.165)
High level Mother Education (level 4)		0.407** (0.177)	0.404** (0.177)	0.417** (0.177)	0.417** (0.177)
Low level Father Education (level 2)		0.095 (0.121)	0.102 (0.121)	0.087 (0.121)	0.092 (0.121)
Mid level Father Education (level 3)		0.209* (0.124)	0.214* (0.124)	0.201* (0.124)	0.203* (0.124)
High level Father Education (level 4)		0.224* (0.138)	0.230* (0.138)	0.215 (0.138)	0.217 (0.137)
Individual/Parental Characteristics Controls	No	Yes	Yes	Yes	Yes
School Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Gender Specific Effects	No	No	No	Yes	Yes
Number of Groups (Students)	56251	55577	55577	55577	55577
Observations	111795	111795	111795	111795	111795
R-Squared	0.0402	0.1405	0.1424	0.1418	0.1444

Source: authors calculations using information from CERAC, ICFES and DNP. All the specifications include student and municipal fixed effects. **Significance levels:** *** p<0.01, ** p<0.05, * p<0.1

***I also estimated additional specifications that use a 10-year average and a 15-year average of the conflict events as a measure of the presence of conflict. The results for these specifications are also available upon request.

WORKS CITED

- Akbulut-Yuksel, M. (2009). *Children of war: The long-run effects of large-scale physical destruction and warfare on children*. IZA.
- Akbulut-Yuksel, M. (2014). Children of war: The long-run effects of large-scale physical destruction and warfare on children. *The Journal of Human Resources*. 49: 634-662.
- Akresh, R., and de Walque, D. (2008). Armed conflict and schooling: evidence from the 1994 Rwandan genocide.
- Balcázar, C., and Nopo, H. (2015). Broken gears: the value added of higher education on teachers' academic achievement. *World Bank Policy Research Working Paper*, (7168).
- Barrera, F., and Ibáñez, A. M. (2004). Does Violence Reduce Investment in Education?: A Theoretical and Empirical Approach: UNIVERSIDAD DE LOS ANDES-CEDE.
- Belley, P., and Lochner, L. (2007). *The changing role of family income and ability in determining educational achievement* (No. w13527). National Bureau of Economic Research.
- Bellows, J., and Miguel, E. (2006). War and institutions: New evidence from Sierra Leone. *The American economic review*, 96(2), 394-399.
- Benefield, P. and Tomlinson., K. (2005). Education and Conflict: Research and Research Possibilities: National Foundation for Educational Research (NFER).
- Bogoya, J. D., and Bogoya, J. M. (2013). An academic value-added mathematical model for higher education in Colombia: Caso de la educación superior en Colombia. *Ingeniería e Investigación*, 33(2), 76-81.
- Brück, T., Justino, P., Verwimp, P. and Avdeenko, A. (2010). Identifying Conflict and Violence in Micro-Level Surveys. *IZA Discussion Paper*, (No. 5067). <http://ssrn.com/abstract=1648352>
- Bundervoet, T. (2012). War, Health, and Educational Attainment: A Panel of Children during Burundi's Civil War: Households in Conflict Network.

- (CERAC), Centro de Recursos para el Análisis del Conflicto (2013). *Conflict data*.
- Colombian National Information System for Higher Education - SNIES (2015). Estadísticas de Docentes de Educación Superior. Retrieved February 11, 2015, from <http://www.mineduacion.gov.co/sistemasdeinformacion/1735/w3-article-212400.html>
- Cunha, J. M., and Miller, T. (2014). Measuring value-added in higher education: Possibilities and limitations in the use of administrative data. *Economics of Education Review*, 42, 64-77.
- Dabalen, A., and Paul, S. (2012). Estimating the causal effects of conflict on education in Cote d'Ivoire. *World Bank Policy Research Working Paper* (6077).
- (DANE), Departamento Administrativo Nacional de Estadísticas (2010). Ficha Metodológica de la Investigación de Educación Formal. Vol. 6, pp. 10. Bogotá: Departamento Administrativo Nacional de Estadística (DANE).
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294.
- DeShazo, P., Primiani, T., and McLean, P. (2007). *Back from the brink: evaluating progress in Colombia, 1999-2007*: CSIS.
- Departamento Nacional de Planeación - DNP (2014). *Bases del Plan Nacional de Desarrollo 2014 – 2018*. Versión Preliminar para Discusión del Consejo Nacional de Planeación.
- Departamento Nacional de Planeación (DNP). *Indicador de Desarrollo Municipal (INDEMUN)*. Retrieved February 20, 2015, from: <https://www.dnp.gov.co/programas/desarrollo-territorial/evaluacion-y-seguimiento-de-la-descentralizacion/Paginas/documentos-de-evaluacion.aspx>
- Domingue, B. (2012). *Measuring Effects of Colombian Postsecondary Institutions on Student Learning*. ICFES: November 2012.
- Duncan, T. E., and Duncan, S. C. (2009). The ABC's of LGM: An introductory guide to latent variable growth curve modeling. *Social and personality psychology compass*, 3(6), 979-991.

- Fergus, S., and Zimmerman, M. A. (2005). Adolescent resilience: A framework for understanding healthy development in the face of risk. *Annu. Rev. Public Health*, 26, 399-419.
- Fundación Ideas para la Paz (2014). *El Conflicto en Colombia*. Retrieved February 10, 2015, from: <http://www.ideaspaz.org/>
- Galdo, J. (2013). The long-run labor-market consequences of civil war: Evidence from the Shining Path in Peru. *Economic Development and Cultural Change*, 61(4), 789-823.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., and Vermeersch, C. M. (2011). *Impact evaluation in practice*. World Bank Publications.
- Häkkinen, I., Kirjavainen, T., and Uusitalo, R. (2003). School resources and student achievement revisited: new evidence from panel data. *Economics of Education Review*, 22(3), 329-335.
- Haveman, R., and Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of economic literature*, 1829-1878.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153-161.
- Hess, B. (1999). *Using Latent Growth Modeling in Program Evaluation: A Primer for the Evaluator*. ERIC.
- Hox, J., and Stoel, R. D. (2005). Multilevel and SEM approaches to growth curve modeling. *Wiley StatsRef: Statistics Reference Online*.
- Ibáñez, A. M., and Moya, A. (2010). Do conflicts create poverty traps? Asset losses and recovery for displaced households in Colombia. *The Economics of Crime: Lessons for and from Latin America* (pp. 137-172): University of Chicago Press.
- (ICFES), Instituto Colombiano para la Evaluación de la Educación (2013). *FTP Databases*. from Instituto Colombiano para la Evaluación de la Educación (ICFES)

- Isáziga-David, C. H., Coello, J. G., and Vásquez-Rizo, F. E. (2014). La intervención académica en la construcción de una sociedad con calidad: análisis del valor agregado en el proceso formativo colombiano. *Hallazgos*, (22).
- Jimenez, O. R. R. (2014, December). The value-added in engineer programs in Colombia value-added in Colombia. In *Interactive Collaborative Learning (ICL), 2014 International Conference on* (pp. 410-412). IEEE.
- Justino, P. (2010). War and Poverty. *Handbook of the Economics of Peace and Security*.
- Kecmanovic, M. (2012). The Short-run Effects of the Croatian War on Education, Employment, and Earnings. *Journal of Conflict Resolution*.
- Khandker, S. R., Koolwal, G. B., and Samad, H. A. (2010). Handbook on impact evaluation: quantitative methods and practices. World Bank Publications.
- Kim, H., and Lalancette, D. (2013). Literature Review on the Value-Added Measurement in Higher Education.
- Klein, S. P., Kuh, G., Chun, M., Hamilton, L., and Shavelson, R. (2005). An Approach To Measuring Cognitive Outcomes Across Higher Education Institutions. *Research in Higher Education*, 46(3), 251-276.
- Lai, B. and C. Thyne. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, 44(3), 277-292. doi: 10.1177/0022343307076631
- LeGrand, C. C. (2001). The Colombian crisis in historical perspective. *Canadian Journal of Latin American and Caribbean Studies*, 28(55-56), 165-209.
- Liu, O. L. (2011a). Value-added assessment in higher education: A comparison of two methods. *Higher Education*, 61(4), 445-461.
- Liu, O. L. (2011b). Measuring value-added in higher education: conditions and caveats—results from using the Measure of Academic Proficiency and Progress (MAPP™). *Assessment and Evaluation in Higher Education*, 36(1), 81-94.
- Melo, L. A., Ramos, J. E., and Hernández, P. O. (2014). La Educación Superior en Colombia: Situación Actual y Análisis de Eficiencia. *Banco de la República de Colombia*, 3.

- Merrouche, O. (2006). The human capital cost of landmine contamination in Cambodia: Households in Conflict Network.
- Ministerio de Educación Nacional (2014). Estadísticas de Educación Superior. Retrieved February 11, 2015, from http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-212350_Estadisticas_de_Educacion_Superior_.pdf
- OECD (2012). Reviews of National Policies for Education: Tertiary Education in Colombia. Organisation for Economic Cooperation and Development.
- Østby, G. (2013). Inequality and political violence: A review of the literature. *International Area Studies Review*, 16(2), 206-231.
- Plecki, M. L., Elfers, A. M., and Nakamura, Y. (2012). Using Evidence for Teacher Education Program Improvement and Accountability An Illustrative Case of the Role of Value-Added Measures. *Journal of Teacher Education*, 63(5), 318-334.
- Quinn, J. M., Mason, T. D., and Gurses, M. (2007). Sustaining the peace: Determinants of civil war recurrence. *International Interactions*, 33(2), 167-193.
- Restrepo, J. A., Spagat, M., and Vargas, J. F. (2006). El conflicto en Colombia: ¿quién hizo qué a quién? Un enfoque cuantitativo *Nuestra guerra sin nombre: Transformaciones del conflicto en Colombia*: Editorial Norma.
- Rodríguez, C., and Sánchez, F. (2010, 2012). Books and guns: the quality of schools in conflict zones. Documento CEDE (38).
- Rodríguez, O. R. (2014). The value-added in engineer programs in Colombia value-added in Colombia. In *Interactive Collaborative Learning (ICL), 2014 International Conference on* (pp. 410-412). IEEE.
- Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American journal of orthopsychiatry*, 57(3), 316.
- Saavedra, A. R., and Saavedra, J. E. (2011). Do colleges cultivate critical thinking, problem solving, writing and interpersonal skills?. *Economics of Education Review*, 30(6), 1516-1526.

- Sánchez, F., and Díaz, A. M. (2005). Los efectos del conflicto armado en el desarrollo social colombiano, 1990-2002. *CEDE, Centro de Estudios sobre Desarrollo Económico*, Facultad de Economía, Universidad de los Andes.
- Sanchez, F., Quirós, M., Reverón, C., and Rodríguez, A. (2002). Equidad Social en el acceso y permanencia en la Universidad Pública determinantes y factores asociados. *CEDE, Centro de Estudios sobre Desarrollo Económico*.
- Shemyakina, Olga N. Armed conflict, education and the marriage market: evidence from Tajikistan. *ProQuest*, 2007.
- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, 95(2), 186-200.
- Stoel, R. D., van den Wittenboer, G., and Hox, J. (2004). Methodological issues in the application of the latent growth curve model. In *Recent developments on structural equation models* (pp. 241-261). Springer Netherlands.
- Swee, E. L. (2009). On war and schooling attainment: The case of Bosnia and Herzegovina. *Households in Conflict Network Working Paper*, 57.
- Todd, P. E., and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485), F3-F33.
- Verwimp, P., and Van Bavel, J. (2013). Schooling, violent conflict, and gender in Burundi. *The World Bank Economic Review*.
- Waxman, H. C., Gray, J. P., and Padron, Y. N. (2003). *Review of Research on Educational Resilience*. Research Report.
- Wharton, K. and Oyelere, R. (2012). *Conflict and Its Impact on Educational Accumulation and Enrollment in Colombia: What We Can Learn from Recent IDPs*. IZA Discussion Paper. IZA. Retrieved from <http://ssrn.com/abstract=1922011>
- Winters, M. A., Dixon, B. L., and Greene, J. P. (2012). Observed characteristics and teacher quality: impacts of sample selection on a value added model. *Economics of Education Review*, 31(1), 19-32.

Wößmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2), 117-170.

CHAPTER 11

CONCLUSIONS

The importance of having a better understanding of the relationship between education and conflict has been highlighted by the United Nations General Assembly report *Impact of Armed Conflict on Children* (1996). According to that report, “education is particularly important at times of armed conflict [...] while all around may be in chaos, schooling can represent a state of normalcy [...] school children have the chance to be with friends and enjoy their support and encouragement [...] the ability to carry on schooling in the most difficult circumstances demonstrates a confidence in the future: communities that still have a school feel they have something durable and worthy of protection” (Machel et al 1996). However, this important issue has been insufficiently addressed in the literature⁵⁹. The limited availability of academic research in this topic is partly due to the difficulties involved in the systematic collection of evidence on why, where and how attacks occur⁶⁰. The availability of a rich database about conflict events in Colombia, maintained by the Resource Centre for Conflict Analysis (CERAC), has opened new possibilities for the analysis of the relationship between conflict and education. In that sense, the Colombian case offers a unique opportunity for research in this topic.

The aim of this doctorate thesis is precisely to get a better understanding of the relationship between civil conflict and educational achievement, and quantify the magnitude in which civil conflict affects the accumulation of human capital measured by cognitive exams. The approach taken in this dissertation, which takes into consideration educational achievement and not academic desertion as a measure of human capital accumulation, is of relevance given the implications of obtaining high marks in national exams in terms of future academic and labour opportunities⁶¹. The research presented in this dissertation concentrates on those students residing in conflict and non-conflict zones who take part in the formal education system, as these are the only students that can be observed in the available micro data. Hence, the specific objective of this dissertation is to better

⁵⁹ Benefield and Tomlinson (2005) has pointed out that education in conflict and post-conflict situations, as a recognized field of research, is in its infancy. According to Bruck et al. (2010): “the current state of art of empirical research on violent conflict offers considerable opportunity for improving our knowledge of violent conflict itself, its functions and dynamics, as well as the impact of conflict on behaviour, welfare and overall development.”

⁶⁰ According to Justino (2010), high quality monitoring data and systematic reporting of events are very limited.

⁶¹ e.g. Chetty et al. (2011), have documented that test scores predict long-run socioeconomic outcomes (e.g., earnings, employment, arrests, etc.)

understand the relationship between educational achievement and civil conflict amongst those students that participate in the formal educational system.

This dissertation contributes to the existing literature that studies the relationship between civil conflict and education, including the contributions of Swee (2008), Akresh and de Walque (2008), Merrouche (2006), Lai and Thyne (2007), Shemyakina (2011), Debalen and Paul (2012), Bundervoet (2012), Bellows and Miguel (2006), Rodríguez and Sánchez (2010), amongst others⁶². These research efforts have focused on the relationship between school enrollment, gender, years of education and civil conflict, but have not studied the relationship between civil conflict and academic achievement measured by cognitive examinations as we do. The dissertation also contributes to the literature that studies the impact of internal displacement on the welfare of civilians, including the work of Ibañez and Velez (2008), Ibañez and Moya (2010a), Ruiz and Vargas (2013), Calderón and Ibañez (2009), Kirschhoff and Ibañez (2002), Kondylis (2008 and 2010), and Fiala (2009).

Main Findings

In the following, I briefly summarize the main findings from each chapter. This section brings together the results of each empirical chapter in order to draw general conclusions regarding the relationship between civil conflict and the levels of achievement in standardized examinations.

Chapter 7 explores the relationship between civil conflict and educational achievement in Colombia through the use of multilevel analysis. Multilevel analysis is employed in this chapter in order to deal adequately with the hierarchical structure of the data by simultaneously modeling variables at different levels without resorting to aggregation or disaggregation of the dataset. The results obtained in this chapter show that the relationship between the intensity of the conflict and the performance in exit examinations might not be as straightforward as it may seem at first glance. For the year 2001 we found a positive relationship between our contemporaneous measure of the intensity of the conflict and the performance in the mathematics and language examinations, but for the year 2002 we found a negative relationship between the intensity of the conflict and the performance in the language examination. It should be acknowledged that the results obtained in this chapter show a smaller effect than expected. These results were contextualized using the contributions from other researchers, including Rodríguez and Sanchez (2012), Sánchez and Diaz (2005), Barrera and Ibañez (2004) and Wharton and Oyelere (2012), to help us establish

⁶² This literature was reviewed in detail in Chapter 4.

possible factors through which the intensity of the conflict could affect the performance of students in exit examinations, and in doing so, help us get a better understanding of the results. We also incorporated into the analysis a historical perspective of the evolution of the Colombian conflict in order to get a more complete understanding of the results.

Due to the lack of longitudinal data, in Chapter 8 we construct a pseudo-panel of schools from cross-sectional data sets in order to get a better understanding of the relationship between civil conflict and education. In fact, to the best of my knowledge this is the first effort to try to understand the relationship between conflict and achievement in standardized exams through the use of a pseudo panel of schools. The results obtained in this chapter provide evidence of the negative repercussions that a civil conflict can have on human capital accumulation. Our results show a negative and significant relation between our two measures of the intensity of the conflict and the levels of performance of schools in the mathematics and language examinations. The results also show a negative and significant relation between our second measure of the presence of conflict (dummy variable) and the levels of performance of schools in the mathematics and language examinations. However, the magnitude of this relationship was smaller than expected. There are a number of factors that were taken into consideration to explain the results obtained in the econometric exercise. Following Lupton (2004), Williams (2004), O'Malley (2007), Sammons (1995) and Thrupp and Lupton (2006) some of these factors were discussed in this chapter.

In *Chapter 9* we used two methodologies to estimate the effects of forced displacement on the performance of Colombian students in the national high school exit examination: cluster robust inference and an instrumental variables approach. In cluster robust inference observations are grouped into clusters, with model errors uncorrelated across clusters but correlated within clusters. However, this is our naïve approach in the sense that we were not considering the possibility of endogeneity. Following Calderon and Ibañez (2009), we used an instrumental variables approach as our second method of analysis in order to deal with the possibility of endogeneity as a source of bias. In fact, the instrumental variables approach was our preferred method of estimation in that chapter. The main finding of this chapter is that being displaced has a negative and significant effect on the levels of achievement of students in mathematical and language exit examinations. However, the results obtained using the instrumental variables approach show a larger effect of displacement on the performance in the examinations. The results presented are not surprising if we take into consideration the findings of a number of researchers who have studied the effects of displacement. The literature has shown that there are a number of needs and vulnerabilities associated to the displacement status that affect the well-being of

individuals, and in doing so, might also be affecting the performance of students in standardized tests. According to the literature, some of the factors that generate and heighten existing vulnerabilities are: loss of property (Ibañez and Moya 2010a); reduced consumption levels (Ibañez and Moya 2010b); early entrance to the labor market (Rodriguez and Sanchez 2012); fear and psychological trauma; and separation from family members, social networks and communities (Ferris and Winthrop 2010). All of these factors can be seen as channels through which the displacement status could have potentially affected educational results.

Chapter 10 estimates the effect of civil conflict on student achievement gains in standardized examinations. By considering students who have been exposed to a conflict environment and students who have not been exposed to a conflict environment, we want to find out whether Colombian students affected by conflict have differential gains or losses in performance in comparison to those students who have not been affected, using the results from the Colombian high school exit examination (Saber11) and the Colombian college exit examination (SaberPro). To the best of my knowledge, this is the first attempt to study the relationship between civil conflict and academic achievement measured by cognitive examinations at both high school and university levels. To tackle this research question we used the theoretical framework employed in educational value added models. Following this theoretical backbone, we applied a difference in differences estimation strategy in order to quantify the student's learning gains using information at two points in time (high school and university). The main finding of this chapter is that students who have been exposed to conflict have more significant improvements in their academic performance in comparison to those students who have not been affected. This finding is robust to the different measures of conflict (intensity and presence) that were used in the empirical exercise of this chapter. There are a number of factors that could be taken into consideration to explain the results obtained in the econometric exercise. The literature that studies resilience provides us a framework to understand the reasons why some students who have been exposed to high levels of risk not necessarily experience problems associated with those risks (Rutter 1987) (Fergus and Zimmerman 2005) (Waxman et al. 2003).

Contextualization of the Results

It should be acknowledged that the results obtained in this dissertation show a smaller than expected effect of conflict on achievement in standardized examinations. There are two reasons why we believe this is the case: (i) the protective role of education and (ii) a self-selection mechanism.

The Protective Role of Education

It is very important to call attention to the fact that the findings of this thesis can be indicative of the protective effect that can result from attending school. Indeed, the literature has shown that attending school can have a very important role in protecting vulnerable children and young adults. It is well known that children and young adults who live in conflict zones are vulnerable and in need of special protection. Schools can deliver that protection by providing them a safe place to play, offering an alternative to destructive conducts, providing access to healthy and nutritious meals, and offering guidance from counselors and teachers (Nicolai and Triplehorn 2003). Nicolai and Triplehorn (2003) have pointed out that there are benefits to the psychosocial health of students living in conflict zones related to education. That is the case because attending school provides students an opportunity for self-expression and interaction with other kids. This interaction is particularly important because it boosts their sense of identity and inclusion (Nicolai and Triplehorn 2003). Keeping kids in school can be a protective mechanism because it gives students a sense of self-worth, allows students to develop social networks, and provides a structured program of activities (Nicolai and Triplehorn 2003). Additionally, as Smith (2010) highlights, schools provide critical information and problem-solving skills that can help protect students from exploitation and harm, abduction, and child soldiering.

Sommers (2002) has highlighted that despite the traumas caused by civil war, some children do manage to become healthy, vibrant, contributing adults. In fact, this researcher points out that it is possible to foster resiliency amongst those children and young adults who have been affected by the civil conflict (Sommers 2002). Attending school is crucial because it can help kids recuperate hope by providing better prospects for the future. Nicolai and Triplehorn (2003) highlight that by setting goals such as completing homework, preparing for exams or completing a school certificate, regular assignments and tests, schools can provide students with achievable objectives that can be seen as reasons to continue fighting for a better future. Education indeed plays a critical role in building bridges and in helping break down stigma.

Yet, there are certain aspects that should be taken into consideration in order to ensure that schools can act as protective spaces for children. In fact, Bird (2007) points out that it is particularly important to “build strong community-school partnerships to ensure safe havens for schools/teachers, building up protection mechanisms and integrating these into the disarmament, demobilisation and reintegration (DDR) processes to prevent re-recruitment”. Furthermore, it is imperative for governments to negotiate with illegal armed groups so that it is

possible to provide a guarantee that schools are indeed protected and are therefore safe learning spaces (Bird 2007).

Education is fundamental in order to achieve stability in countries that are being affected by a civil conflict. Providing young adults with a sense of normalcy during a civil conflict, by promoting their active participation at school, can impart a sense of self-confidence and hope for the future. This sense of normalcy plays a vital role in the development of individuals because it can help break important barriers that could be limiting school performance (Justino 2014). In fact, as Justino (2014) has highlighted: “education systems in conflict-affected countries can play important roles in supporting stability and economic resilience when children and young people are educated or trained to support positive social change, and when they work in tandem with interventions aimed at addressing specific poverty, vulnerabilities and aspirations of individuals and household affected by violence”. Furthermore, it is particularly important to understand that in countries where peace processes are underway, like Colombia, education can play a vital role by contributing to the transformation of the conflict (Smith 2014). This can be done by educating people about potential new arrangements for political representation and justice that can reduce confrontation between different groups of the society (Smith 2014).

Self Selection

Self-selection is a second factor that could explain why the results obtained in this dissertation show a smaller than expected influence of conflict on achievement in standardized examinations. It is important to emphasize that the empirical studies presented in this thesis concentrate only on those students residing in conflict and non-conflict zones who take part in the formal education system. These students are the ones who get evaluated in the exit examinations, and therefore, they are the ones that we can observe in the micro data. However, it is possible that those individuals who are more affected by conflict do not participate in the formal education system, and therefore they are not observable in our dataset. In fact, the literature has shown that those individuals that do not have the resources (monetary, psychological, family support) to overcome or adapt to the rough conditions that are prevalent in conflict areas are more likely to drop out from the formal educational system. In that sense, self-selection could also be explaining our results because it is more likely that those individuals with resources to overcome or adapt to the harsh conditions have the opportunity to continue with their studies.

For the Colombian case there are a number of studies that have documented the relationship between civil conflict, low levels of enrollment and school dropout. Rodriguez and Sanchez (2012) have found that armed conflict reduces the average years of schooling in 8.78% for a pooled sample of Colombian children, and it reduces the average years of schooling in 17.03% for children between sixteen and seventeen years old. Barrera and Ibañez (2004) have also found that after controlling for individual, household and other context variables, violence influences negatively school enrollment. Furthermore, Sánchez and Diaz (2005) show that municipalities that were affected by the activities of illegally armed groups had slower growth rates of primary and secondary school enrolment compared to those in municipalities that were not affected by illegal activities using a matching mechanism. Wharton and Oyelere (2012) also show that children who live in a municipality with high levels of conflict have a gap in enrolment and accumulation.

Furthermore, Sommers (2002) highlights that young adults who drop out of the formal education system have a higher probability of experiencing increased instability and violence in their lives and communities. This researcher further highlights that taking part in the formal education system provides important alternatives to child soldiering, social and cultural alienation, violence, and self-destruction (Sommers 2002).

Formulation of Policy

The evidence obtained in this thesis can be used to guide the formulation of educational policy. There are a number of policy recommendations that could be implemented in order to assure that education plays the protective role discussed above. Policy makers should work in the creation of new and better incentives to retain students who reside in conflict zones in school with a full time dedication, and if possible, provide financial assistance so that qualified students get the opportunity to attend university. It is possible that those students who are facing difficulties associated to the internal armed conflict are developing high levels of resilience that allows them to continue with their studies successfully, but they need a special follow up. The design of effective policies to guarantee that those students who stay in conflict zones have the adequate incentives and resources to continue with their studies is an essential task that should be prioritized in the agenda of Colombian policy makers.

More governmental resources should be spent in order to meet the special needs of those students who have been affected by conflict (i.e. psychological help, school supplies, special remedial tutorials) and to supervise their progress at educational

institutions. The government and development agencies should provide funding in order to assure that schools in conflict zones provide a high quality education that is inclusive. Low levels of investment on education for young adults who are in danger can make peace a more difficult task to achieve. It should be noted that all these efforts not only help improve the quality of life of those individuals affected by the conflict. As Ostby (2013) has highlighted there is evidence that clearly points to a pacifying effect of education, which has positive repercussions for the entire society. Furthermore, Ostby (2013) has also emphasized that democracies, like Colombia, tend to experience a greater stabilizing effect of education than non-democracies do.

Special efforts should be made at conflict zones to strengthen the capacity of the education authorities, and to promote organizational development at schools. Additionally, policy makers should design policies to develop the different factors that are conducive to effective schools (i.e. shared vision and goals, a learning environment, concentration on teaching and learning, purposeful teaching, high expectations, home-school partnership, and learning organization). Policy makers should design effective policies to guarantee that those schools in conflict zones have the adequate incentives and resources to continue developing these factors. It is particularly important to provide sufficient training to teachers and school authorities on how to deal with the special circumstances associated to the conflict. For instance, teachers should be prepared to guide students that need special protection (e.g. displaced students, risk of recruitment).

The public authorities should make the necessary arrangements in order to protect schools. This is particularly important in order to guarantee that schools are safe learning spaces where students can feel protected. Furthermore, these spaces should provide specialized guidance for kids so that they receive adequate support for learning and psychosocial development. Authorities should also monitor school attendance. This is particularly important in order to have an early detection system of those students who are more likely to dropout from school. As shown previously, those students who dropout from the formal education system might be the ones with the highest levels of affectation.

Policy makers should work in the creation of new and better incentives to keep displaced students in school with a full time dedication, and delay their entrance into the labor market. If school age individuals can dedicate their time and effort exclusively to schooling, the results in the examinations will likely be better. Most of the transmission channels through which internal displacement could have potentially affected the performance of students in High School exit examinations were related to the low levels of income amongst displaced households. Consequently, the authorities should implement programs to help displaced

families earn a higher income (i.e. employment programs, conditional cash transfer programs, and other forms of subsidies) in order to reduce some of the needs and vulnerabilities associated to the displacement status that affect the well-being of students. Moreover, more governmental resources should be spent in order to meet the special needs of displaced students at schools (i.e. psychological help, school supplies, special remedial tutorials) and to supervise their progress at educational institutions.

Limitations and Future Research

The limitations of this dissertation are illustrative of the empirical, theoretical and methodological challenges that this area of research faces. For instance, it is important to acknowledge that given the nature of the datasets under analysis, self-selection bias and sample selection bias might have been a concern in the empirical exercises presented in this thesis. However, in order to tackle these issues (address the possibility of bias), we used different methodologies including an instrumental variables approach and a difference in differences methodology.

Another limitation that should be acknowledged is related to the specification of the education production function. An important problem that comes up in the specification of an education production function to be estimated empirically is the non-existence of a valid measure of innate abilities. The absence of a valid measure of innate abilities can be problematic because it can bias the estimated regression coefficients. The size of the bias depends on how strong is the effect of the variable on achievement and also on the correlation of the omitted variable with the other variables that are included in the model. Omitting the endowed capacity of the student from the empirical specification will likely cause an upward bias in the estimated impact of the other variables on achievement (Hanushek 1979). The possibility of this upward bias should not be ignored, and therefore, it should be considered carefully in the interpretation of the results of the empirical exercises presented in this thesis.

Throughout this dissertation we pointed out a number of factors that could help us explain the results obtained in the empirical exercises. However, in some cases given data restrictions we were not able to empirically test the importance of these factors in explaining our results. In order to analyse most of the factors pointed out we would need to use mixed methods (i.e. qualitative and quantitative analysis). The application of a qualitative approach to our research would be of great help to get a better understanding of the underlying factors that explain the quantitative results that were obtained using micro-data. The use of qualitative techniques in the analysis of the relationship between civil conflict and education achievement constitutes an important part of my future research agenda.

While this thesis examined the relationship between civil conflict and educational achievement by quantifying the magnitude in which civil conflict affects the accumulation of human capital measured by cognitive exams, more research can be done to further improve our understanding of education in conflict and post-conflict situations. The limited amount of research in this area is largely explained by the limited availability and reliability of micro data about civil conflict events and standardized test results. The availability of a rich database about conflict events in Colombia, maintained by the Resource Centre for Conflict Analysis (CERAC) and a government (ICFES) kept database containing the results of standardized tests at different levels of schooling, opens new possibilities for the analysis of the relationship between conflict and education. In that sense, the Colombian case offers a unique opportunity for research in this topic. It is important to take advantage of the very rich information available in these datasets to carry out more research in this topic. Without a doubt there is much more that can be learned about civil conflict and education. In fact, a very important topic that should be further explored is the relationship between academic achievement at a young age (elementary school) and civil conflict. This topic will be explored in my future research agenda.

WORKS CITED

- Akresh, R., and de Walque, D. (2008). Armed conflict and schooling: evidence from the 1994 Rwandan genocide.
- Barrera, F., and Ibáñez, A. M. (2004). Does Violence Reduce Investment in Education?: A Theoretical and Empirical Approach: UNIVERSIDAD DE LOS ANDES-CEDE.
- Bellows, J., and Miguel, E. (2006). War and institutions: New evidence from Sierra Leone. *The American economic review*, 96(2), 394-399.
- Benefield, P. and Tomlinson, K. (2005). Education and Conflict: Research and Research Possibilities: National Foundation for Educational Research (NFER).
- Bird, L. (2007). Children in Crisis: Education rights for children in conflict affected and fragile states.
- Brück, T., Justino, P., Verwimp, P. and Avdeenko, A. (2010). Identifying Conflict and Violence in Micro-Level Surveys. IZA Discussion Paper, (No. 5067). <http://ssrn.com/abstract=1648352>
- Bundervoet, T. (2012). War, Health, and Educational Attainment: A Panel of Children during Burundi's Civil War: Households in Conflict Network.
- Calderón, V., and Ibáñez, A. M. (2009). *Labor market effects of migration-related supply shocks: Evidence from internally displaced populations in Colombia*: CEDE, Centro de Estudios sobre Desarrollo Económico, Facultad de Economía, Universidad de los Andes.
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, 126(4), 1593-1660.
- Debalen, A., and Paul, S. (2012). Estimating the causal effects of conflict on education in Cote d'Ivoire. *World Bank Policy Research Working Paper* (6077).
- Fergus, S., and Zimmerman, M. A. (2005). Adolescent resilience: A framework for understanding healthy development in the face of risk. *Annu. Rev. Public Health*, 26, 399-419.

- Ferris, E., and Winthrop, R. (2010). Education and Displacement: Assessing Conditions for Refugees and Internally Displaced Persons affected by Conflict. *Background paper for the EFA Global Monitoring Report 2011. The hidden crisis: Armed conflict and education.*
- Fiala, N. (2009). The consequences of forced displacement in northern Uganda: Households in Conflict Network.
- Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources*, 14(3), 351-388.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 24(3), 1141-1177.
- Hanushek, E. A. (1997). Assessing the Effects of School Resources on Student Performance: An Update. *Educational evaluation and policy analysis*, 19(2), 141-164. doi: 10.3102/01623737019002141
- Hoeffler, A., and Reynal-Querol, M. (2003). Measuring the costs of conflict. Washington, DC: World Bank.
- Ibáñez, A. M., and Moya, A. (2010a). Vulnerability of victims of civil conflicts: empirical evidence for the displaced population in Colombia. *World Development*, 38(4), 647-663.
- Ibáñez, A. M., and Moya, A. (2010b). Do conflicts create poverty traps? Asset losses and recovery for displaced households in Colombia *The Economics of Crime: Lessons for and from Latin America* (pp. 137-172): University of Chicago Press.
- Ibáñez, A. M., and Velez, C. E. (2008). Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia. *World Development*, 36(4), 659-676. doi: 10.1016/j.worlddev.2007.04.013
- Justino, P. (2010). How Does Violent Conflict Impact on Individual Educational Outcomes? The Evidence So Far. *background paper for the Education For All Global Monitoring Report.*
- Justino, P. (2014). Barriers to Education in Conflict-Affected Countries and Policy Opportunities. *Paper commissioned for Fixing the Broken Promise of Education*

for All: Findings from the Global Initiative on Out-of-School Children (UIS/UNICEF, 2015), Montreal: UNESCO Institute for Statistics (UIS).

- Kirchhoff, S., and Ibáñez, A. M. (2001). Displacement due to violence in Columbia: Determinants and consequences at the household level: ZEF Discussion Papers on Development Policy.
- Kondylis, F. (2008). Agricultural outputs and conflict displacement: Evidence from a policy intervention in Rwanda. *Economic Development and Cultural Change*, 57(1), 31-66.
- Kondylis, F. (2010). Conflict displacement and labor market outcomes in post-war Bosnia and Herzegovina. *Journal of Development Economics*, 93(2), 235-248.
- Lai, B. and Thyne, C.. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, 44(3), 277-292. doi: 10.1177/0022343307076631
- Merrouche, O. (2006). The human capital cost of landmine contamination in Cambodia: Households in Conflict Network.
- Nicolai, S., and Triplehorn, C. (2003). The role of education in protecting children in conflict. *Network Paper: Humanitarian Practice Network (HPN)*, 42, 1-36.
- Østby, G. (2013). Inequality and political violence: A review of the literature. *International Area Studies Review*, 16(2), 206-231.
- Rodríguez, C., and Sánchez, F. (2010). Books and guns: the quality of schools in conflict zones. *Documento CEDE* (38).
- Rodriguez, C., and Sanchez, F. (2012). Armed conflict exposure, human capital investments, and child labor: Evidence from Colombia. *Defence and peace economics*, 23(2), 161-184.
- Ruiz, I., and Vargas-Silva, C. (2013). The economics of forced migration. *The Journal of Development Studies*, 49(6), 772-784.
- Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American journal of orthopsychiatry*, 57(3), 316.
- Sánchez, F., and Díaz, A. M. (2005). *Los efectos del conflicto armado en el desarrollo social colombiano, 1990-2002*: CEDE, Centro de Estudios sobre Desarrollo Económico, Facultad de Economía, Universidad de los Andes.

- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, 95(2), 186-200. doi: 10.1016/j.jdeveco.2010.05.002
- Smith, A. (2010). EFA Global Monitoring Report 2011: Education and Conflict.
- Smith, A. (2014). Contemporary Challenges for Education in Conflict Affected Countries. *Journal of International and Comparative Education*, 3(1), 113-125.
- Sommers, M. (2002). Children, Education and War: Reaching Education for All (EFA) Objectives in Countries Affected by Conflict. *Conflict Prevention and Reconstruction Unit Working Paper*.
- Swee, E. L. (2009). On war and schooling attainment: The case of Bosnia and Herzegovina. *Households in Conflict Network Working Paper*, 57.
- Waxman, H. C., Gray, J. P., and Padron, Y. N. (2003). *Review of Research on Educational Resilience*. Research Report.
- Wharton, K. and Oyelere, R. (2012). *Conflict and Its Impact on Educational Accumulation and Enrollment in Colombia: What We Can Learn from Recent IDPs*. IZA Discussion Paper. IZA. Retrieved from <http://ssrn.com/abstract=1922011>

ADDENDUM ON VALORISATION

This section discusses the valorisation opportunities offered by this doctorate thesis. Following the Regulation Governing the Attainment of Doctoral Degrees established by Maastricht University, these opportunities are examined in terms of the social and economic relevance of the dissertation, the potential target groups, and the methodological innovations applied for the analysis.

A civil war has significant economic, social and political implications that potentially interfere with the development process of the countries affected by it. The objective of this doctorate thesis is to get a better understanding of the relationship between civil conflict and educational achievement, and quantify the magnitude in which civil conflict affects educational achievement measured by the results in exit examinations. To the best of my knowledge, only one effort has been made to try to understand the relationship between conflict and achievement in standardized exams. Such effort was made by Rodríguez and Sánchez (2010). However, this doctorate thesis contributes significantly to the literature because it introduces substantial improvements to Rodríguez and Sanchez's (2010) work in terms of methodology and data quality.

The four empirical papers presented in this doctorate thesis (*Chapters 7, 8, 9 and 10*) contribute to advance our understanding of the relationship between civil conflict and educational achievement. The social and economic implications of the results obtained in these chapters are of considerable importance for the formulation of policy to improve the wellbeing of the individuals living in countries affected by civil conflict. In particular, the evidence obtained in this thesis can be used to guide the formulation of policies to improve educational outcomes in countries that are affected or have been affected by civil conflict.

The approach adopted in this dissertation, which takes into consideration educational achievement and not academic desertion as a measure of human capital accumulation, is of relevance given the implications of obtaining high marks in national exams in terms of future academic and labour opportunities. The results of standardized exit examinations are used in many countries as a filter to be admitted to higher levels of education, to qualify for scholarships and to be accepted in job applications. Additionally, recent papers, including Chetty et al. (2011), have documented that test scores predict long-run socioeconomic outcomes (e.g., earnings, employment, arrests, etc.).

This study focuses on those students residing in conflict and non-conflict zones who take part in the formal education system. It is relevant to explore the particular case of the students who stay in the conflict zones because these students might be adapting (or not) to the rough conditions that are prevalent in these areas, and perhaps they have a different educational gap than policy makers think. Although at first glance it might seem obvious that the relationship between civil conflict and education is negative, there are reasons to think that this relationship is not as straightforward as it seems. In fact, the existing literature that studies the relationship between educational enrolment and civil conflict has shown mixed results. For instance, Swee (2008) and Bellows and Miguel (2006) find no significant effects of civil war on enrolment, whereas Lai and Thyne (2007) and Akresh and de Walque (2008) find evidence to show that civil war has a negative effect on enrolment rates. The lack of consistency in the existing literature justifies the need for further research given the social and economic relevance of this topic.

It is also particularly relevant to study the case of those students who have been displaced from their hometown. The available literature for developing countries has shown that there are serious consequences of forced migration in terms of consumption, employment, and health conditions (Ruiz and Vargas 2013). Nevertheless, the lack of research studying the effects of displacement, explained in large part by the unavailability of data to conduct proper analysis, has limited our understanding of this phenomenon (Ruiz and Vargas 2013). It is notable that to the best of my knowledge the empirical exercise presented in *Chapter 9* of this dissertation is the first effort to try to understand and quantify the impact of forced displacement on educational achievement.

The main conclusions of this thesis are of interest not only to the academic community but also to policy makers. In particular, it can provide insights to policy makers who are interested on the formulation of educational policy for conflict and post-conflict scenarios. There are a number of policy recommendations that can be derived from this research and could be implemented in order to assure that education plays a protective role in conflict regions. A general conclusion of this doctorate thesis is that policy makers should work in the creation of new and better incentives to retain students who reside in conflict zones in school with a full time dedication, and if possible, provide financial assistance so that qualified students get the opportunity to attend university. The results obtained in the empirical chapters of this dissertation point to the possibility that those students who are facing difficulties associated to the internal armed conflict are developing high levels of resilience that allows them to continue with their studies successfully, but they need a special follow up. As a

consequence, the design of effective policies to guarantee that those students who stay in conflict zones have the adequate incentives and resources to continue with their studies is an essential task that should be prioritized in the agenda of Colombian policy makers. More governmental resources should be spent in order to meet the special needs of those students who have been affected by conflict (i.e. psychological help, school supplies, special remedial tutorials) and to supervise their progress at educational institutions. It should be noted that the literature has shown that these governmental efforts not only help improve the quality of life of those individuals directly affected by the conflict. As Ostby (2013) has highlighted there is evidence that clearly points to a pacifying effect of education, which has positive repercussions for the entire society. Furthermore, Ostby (2013) has also emphasized that democracies, like Colombia, tend to experience a greater stabilizing effect of education than non-democracies do.

Another policy implication that can be inferred from the results of this PhD thesis is linked to the need for safe learning spaces. The public authorities should make the necessary arrangements in order to protect schools. This is particularly important in order to guarantee that schools are safe learning spaces where students can feel protected. Furthermore, these spaces should provide specialized guidance for kids so that they receive adequate support for learning and psychosocial development. Authorities should also monitor school attendance. This is particularly important in order to have an early detection system of those students who are more likely to dropout from school.

Last but not least, special efforts should be made at conflict zones to strengthen the capacity of the education authorities, and to promote organizational development at schools. Policy makers should design policies to develop the different factors that are conducive to effective schools (i.e. shared vision and goals, a learning environment, concentration on teaching and learning, purposeful teaching, high expectations, home-school partnership, and learning organization). It is particularly important to provide sufficient training to teachers and school authorities on how to deal with the special circumstances associated to the conflict. For instance, teachers should be prepared to guide students that need special protection (e.g. displaced students, risk of recruitment).

Regarding the innovativeness of the research methods applied, this dissertation proposes novel methodologies for the analysis of the relationship between conflict and educational outcomes. This thesis addresses several research questions econometrically using a number of techniques, which include: multilevel analysis, cluster robust inference, instrumental variables, difference in differences approach and a pseudo panel approach. The use of a pseudo panel approach is particularly novel in this area of research. In this thesis I construct a pseudo panel of schools in

order to get a better understanding of the relationship between civil conflict and education. According to Deaton (1985), a pseudo-panel can be constructed in order to track 'cohorts', i.e. a group with fixed membership, just as if panel data were available. In this case, the individual results are grouped in school 'cohorts'. To the best of my knowledge this is the first effort to try to understand the relationship between conflict and achievement in standardized exams through the use of a pseudo panel of schools. The use of a difference-in-differences approach (equivalent to a value added specification with individual fixed effects) to quantify the student's learning gains using information at two points in time is also a novel approach in this literature. The idea behind using this methodology is to unravel the effect of institutions from other factors (e.g. individual characteristics, parental characteristics) that also contribute to student achievement.

In terms of the diffusion of the results of the research, all the empirical chapters in this dissertation are going to be published in peer-reviewed journals. **Chapter 8** is currently under review for publication in the *International Journal of Educational Development* and the remaining chapters are currently being reshaped for submission to peer-reviewed journals. Additionally, the results of the empirical chapters have been presented in the GPAC2 workshops. The GPAC2 workshops offer a valuable space to disseminate the results of this research and to get feedback and suggestions to improve the research.

WORKS CITED

- Akresh, R., and de Walque, D. (2008). Armed conflict and schooling: evidence from the 1994 Rwandan genocide.
- Bellows, J., and Miguel, E. (2006). War and institutions: New evidence from Sierra Leone. *The American economic review*, 96(2), 394-399.
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, 126(4), 1593-1660.
- Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of econometrics*, 30(1), 109-126.
- Lai, B. and Thyne, C.. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, 44(3), 277-292. doi: 10.1177/0022343307076631
- Østby, G. (2013). Inequality and political violence: A review of the literature. *International Area Studies Review*, 16(2), 206-231.
- Rodríguez, C., and Sánchez, F. (2010). Books and guns: the quality of schools in conflict zones. *Documento CEDE* (38).
- Ruiz, I., and Vargas-Silva, C. (2013). The economics of forced migration. *The Journal of Development Studies*, 49(6), 772-784.
- Swee, E. L. (2009). On war and schooling attainment: The case of Bosnia and Herzegovina. *Households in Conflict Network Working Paper*, 57.

ABOUT THE AUTHOR

Silvia C. Gómez Soler was born in Bogotá, Colombia in 1984. She holds a BSc and a MSc in Economics (highest University honors) from Pontificia Universidad Javeriana (Bogotá, Colombia), and a MA in International Development (Development Economics) from American University (Washington DC). Silvia has worked as instructor of Economics and as researcher at Universidad Javeriana (Colombia), Universidad de los Andes (Colombia), and American University (Washington DC). She has also worked in the financial sector as an economic researcher at Corficolombiana (Grupo Aval - one of the largest economic organizations in Colombia). Silvia is currently working as Assistant Professor of Economics at Universidad de La Sabana (Chía, Colombia) and also holds a position as Associated Researcher at the Resource Centre for Conflict Analysis (CERAC).

Her research interests are focused on economic development, economics of education, civil conflict, and program evaluation techniques.

MGSoG DISSERTATION SERIES

Paula Nagler

Occupational Choice in the Developing World

MGSoG Dissertation Series, nr 63 (2015)

Jasmin Kientzel

Determinants of Professional Commitment to Environmental Sustainability

MGSoG Dissertation Series, nr 62 (2015)

Mehmet Güney Celbiş

Regional Policies: Convergence, Trade, and the Allocation of Public Capital

MGSoG Dissertation Series, nr 61 (2015)

Florian Henning

Living Up to Standards

Interoperability Governance and Standards Adoption in Government Information Networks

MGSoG Dissertation Series, nr 60 (2015)

Niels P. Groen

The Never-ending Project

Understanding E-Government Project Escalation

MGSoG Dissertation Series, nr 59 (2015)

Derek Copp

Teacher-based Reactivity to Provincial Large-scale Assessment in Canada

MGSoG Dissertation Series, nr 58 (2015)

Michaela Vanore

Family-Member Migration and the Psychosocial Health Outcomes of Children in Moldova and Georgia

MGSoG Dissertation Series, nr 57 (2015)

Sonja Fransen

The Economic and Social Effects of Remittances and Return Migration in Conflict-Affected Areas:

The Case of Burundi

MGSOG Dissertation Series, nr 56 (2015)

Ibrahim Khalil Conteh

The Impact of Floods on Primary School Education in Zambia

MGSOG Dissertation Series, nr 55 (2015)

Richard Bluhm

Growth Dynamics and Development

Essays in Applied Econometrics and Political Economy

MGSOG Dissertation Series, nr 54 (2015)

Nevena P. Zhelyazkova

Work-Family Reconciliation and Use of Parental Leave in Luxembourg

Empirical Analysis of Administrative Records

MGSOG Dissertation Series, nr 53 (2015)

Sachin Kumar Badkas

Metachoice and Metadata

Innovating with Environmental Policy Analysis in Europe

MGSOG Dissertation Series, nr 52 (2014)

Irina S. Burlacu

An Evaluation of Tax-Benefit Systems Impact on the Welfare of Frontier Workers

The Case of Luxembourg and Belgium

MGSOG Dissertation Series, nr 51 (2014)

Özge Bilgili

Simultaneity in Transnational Migration Research:

Links Between Migrants' Host and Home Country Orientation

MGSOG Dissertation Series, nr 50 (2014)

Yulia Privalova Krieger

Reshaping the Big Agenda: Transnational Politics and Domestic Resistance

Financial crisis and social protection reform in Bosnia and Herzegovina

MGSOG Dissertation Series, nr 49 (2014)

Marieke van Houte

Moving Back or Moving Forward?

Return migration after conflict

MGSOG Dissertation Series, nr 48 (2014)

Oxana Slobozhan

Global Governance in the Management of Natural Resources

The Case of the Extractive Industries Transparency Initiative (EITI)

MGSOG Dissertation Series, nr 47 (2014)

Luis Bernardo Mejia Guinand

The Changing Role of the Central Planning Offices in Latin America:

A Comparative Historical Analysis Perspective (1950-2013)

MGSOG Dissertation Series, nr 46 (2014)

Cheng Boon Ong

Ethnic Segregation in Housing, Schools and Neighbourhoods in the Netherlands

MGSOG Dissertation Series, nr 45 (2014)

Luciana V. Cingolani

Bureaucracies for Development: Oxymoron or Reality?

Studies on State Capacity in Challenging Governance Contexts

MGSOG Dissertation Series, nr 44 (2014)

Carlos Cadena Gaitán

Green Politics in Latin American Cities - Sustainable Transport Agendas

MGSOG Dissertation Series, nr 43 (2014)

Katie Kuschminder

Female Return Migration and Reintegration Strategies in Ethiopia

MGSOG Dissertation Series, nr 42 (2014)

Metka Hercog

Highly-Skilled Migration and New Destination Countries

MGSOG Dissertation Series, nr 41 (2014)

Margaret Agaba Rugadya

Can Remittances Influence the Tenure and Quality of Housing in Uganda?

MGSOG Dissertation Series, nr 40 (2014)

Ilire Agimi

New Governance Under Limited Statehood

The Case of Local Government Reform in Kosovo

MGSOG Dissertation Series, nr 39 (2014)

Kristine Farla

Empirical Studies on Institutions, Policies and Economic Development

MGSOG Dissertation Series, nr 38 (2013)

Marina Petrovic

Social Assistance and Activation in the Pursuit of Happiness:

Shedding New Light on Old Policy Solutions to Social Exclusion

MGSOG Dissertation Series, nr 37 (2013)

Laura Torvinen

Assessing Governance Assessments; The Case of Mozambique

Governance Assessments in the Context of Aid Effectiveness Discourse

MGSOG Dissertation Series, nr 36 (2013)

Biniam Egu Bedasso

Institutional Change in the Long Shadow of Elites

Essays on Institutions, Human Capital and Ethnicity in Developing Countries

MGSOG Dissertation Series, nr 35 (2013)

Sepideh Yousefzadeh Faal Deghati

Childhoods Embargoed

Constructing and Reconstructing Multidimensional Child Poverty in Iran 1984-2009

MGSoG Dissertation Series, nr 34 (2013)

Robert Bauchmüller

Investing in Early Childhood Care and Education:

The Impact of Quality on Inequality

MGSoG Dissertation Series, nr 33 (2013)

Martin Rehm

Unified Yet Separated

Empirical Study on the Impact of Hierarchical Positions within Communities of Learning

MGSoG Dissertation Series, nr 32 (2013)

Dorcas Mbuvi

Utility Reforms and Performance of the Urban Water Sector in Africa

MGSoG Dissertation Series, nr 31 (2012)

Lina Salanauskaite

Distributional Impacts of Public Policies:

Essays in Ex-Ante and Ex-Post Evaluation

MGSoG Dissertation Series, nr 30 (2012)

Esther Schüring

To Condition or not – is that the Question?

An Analysis of the Effectiveness of Ex-Ante and Ex-Post Conditionality in Social Cash Transfer Programs

MGSoG Dissertation Series, nr 29 (2012)

Joe Abah

Strong Organisations in Weak States

Atypical Public Sector Performance in Dysfunctional Environments

MGSoG Dissertation Series, nr 28 (2012)

Zina Samih Nimeh

Social Citizenship Rights: Inequality and Exclusion

MGSoG Dissertation Series, nr 27 (2012)

Lenka Eisenhamerová
Legitimacy of 'Humanitarian Military Intervention'
MGSoG Dissertation Series, nr 26 (2011)

Sonila Tomini
Informal Payments for Health Care Services in Albania
MGSoG Dissertation Series, nr 25 (2011)

Jinjing Li
Dynamic Microsimulation in Public Policy Evaluation
MGSoG Dissertation Series, nr 24 (2011)

Aziz Atamanov
*Rural Nonfarm Employment and International Migration as Alternatives to Agricultural Employment:
The Case of Kyrgyzstan*
MGSoG Dissertation Series, nr 23 (2011)

Frieda Vandeninden
Poverty Alleviation: Aid and Social Pensions
MGSoG Dissertation Series, nr 22 (2011)

Juliana Nyasha Tirivayi
*The Welfare Effects of Integrating AIDS Treatment with Food Transfers:
Evidence from Zambia*
MGSoG Dissertation Series, nr 21 (2011)

Agnieska Ewa Sowa
Who's Left Behind? Social Dimensions of Health Transition and Utilization of Medical Care in Poland
MGSoG Dissertation Series, nr 20 (2011)

Emmanaouil Sfakianakis
The Role of Private Actors in the Provision of Public Goods with Applications to Infrastructure and Financial Stability
MGSoG Dissertation Series, nr 19 (2011)

Siu Hing Lo

White Collars Green Sleeves

An Interorganizational Comparison of Determinants of Energy-Related Behaviors among Office Workers

MGSOG Dissertation Series, nr 18 (2011)

Treena Wu

Constraints to Human Capital Investment in Developing Countries:

Using the Asian Financial Crisis in Indonesia as a Natural Experiment

MGSOG Dissertation Series, nr 17 (2011)

Henry Espinoza Peña

Impact Evaluation of a Job-Training Programme for Disadvantaged Youths:

The Case of Projoven

MGSOG Dissertation Series, nr 16 (2011)

Florian Tomini

Between Family and Friends

Understanding the Interdependency of Private Transfers

MGSOG Dissertation Series, nr 15 (2010)

Michał Polalowski

The Institutional Transformation of Social Policy in East Central Europe:

Poland and Hungary in comparative and historical perspective

MGSOG Dissertation Series, nr 14 (2010)

Maha Ahmed

Defining, Measuring and Addressing Vulnerability:

The Case of Post Conflict Environments

MGSOG Dissertation Series, nr 13 (2010)

Pascal Beckers

Local Space and Economic Success

The role of spatial segregation of migrants in the Netherlands

MGSOG Dissertation Series, nr 12 (2010)

Victor Cebotari

Complicating Demands in Ethnically Diverse Societies
Ethnopolitical Contention and Identity Values in Europe
MGSoG Dissertation Series, nr 11 (2010)

Dennis Gyllensporre

Competing and Complementary Perspectives on the EU as a Crisis Management Actor:
An Examination of the Common Security and Defence Policy through the Lenses of
Idealism and Realism
MGSoG Dissertation Series, nr 10 (2010)

Judit Vall Castello

Business Cycle and Policy Effects on Labour Market Transitions of Older and Disabled
Workers in Spain
MGSoG Dissertation Series, nr. 9 (2010)

Keetie Roelen

False Positives or Hidden Dimensions: the definition and measurement of child poverty
MGSoG Dissertation Series, nr. 8 (2010)

Denisa Maria Sologon

Earning Dynamics in Europe
MGSoG Dissertation Series, nr. 7 (2010)

Melissa Siegel

Money and Mobility: Migration and Remittances
MGSoG Dissertation Series, nr. 6 (2010)

Jessica S. Hagen-Zanker

Modest Expectations: Causes and effects of migration on migrant households in
source countries
MGSoG Dissertation Series, nr. 5 (2010)

Mirtha R. Muniz Castillo

Human Development and Autonomy in Project Aid: Experiences from four

bilateral projects in Nicaragua and El Salvador
MGSOG Dissertation Series, nr. 4 (2009)

Christiane Arndt
Governance Indicators
MGSOG Dissertation Series, nr. 3 (2009)

Britta Augsburg
Microfinance – Greater Good or Lesser Evil?
MGSOG Dissertation Series, nr. 2 (2009)

Geranda Notten
Measuring and Managing Poverty Risks
MGSOG Dissertation Series, nr. 1 (2008)