

GROWTH DYNAMICS AND DEVELOPMENT

Essays in Applied Econometrics and
Political Economy

Richard Bluhm

ISBN 978 90 8666 363 7

Copyright © Richard Bluhm, 2015

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.

Publisher: Boekenplan, Maastricht

Growth Dynamics and Development: Essays in Applied
Econometrics and Political Economy

DISSERTATION

to obtain the degree of Doctor at
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 Thursday, 26 March 2015, at 16:00 hours

by

Richard Bluhm

Promoter

Prof. Dr. Adam Szirmai

Co-Supervisor

Dr. Denis de Crombrughe

Assessment Committee

Prof. Dr. Franz C. Palm (chair)

Prof. Dr. Pierre Mohnen

Dr. Toke Aidt (Cambridge University)

Prof. Dr. Jakob de Haan (Rijksuniversiteit Groningen, De Nederlandsche Bank)

Acknowledgements

“Why don’t you apply in Maastricht?” These six words, spoken by Yannic Franken at the Canadian Embassy in Berlin in early 2009, started it all. I moved to Maastricht for my masters. After a brief detour to Geneva and thanks to the generous offer of a position from Chris de Neubourg as well as encouragement from Michael Cichon, I returned in the fall of 2010 to embark on a PhD.

Many would consider themselves blessed if they had one engaged promoter or supervisor. I had the immense pleasure to learn from, work with, and intellectually spar with three on a regular basis. First and foremost, I would like to thank my promoter Eddy Szirmai. I remember a time when I needed to convince him to supervise me. I am glad he finally agreed, shaped this book with me, and taught me the ins and outs of academia in the process. Sitting directly next to Eddy, my eyes would usually find Denis de Crombrughe. Without his guidance, support and the many challenges he gave me, I could not imagine having developed such a knack for quantitative economics. I also wish to thank my “informal co-supervisor”, Kaj Thomsson, who inspired me to work on a political economy theory paper with him. All three gave me the greatest gift there is for any graduate student: seemingly infinite amounts of their time.

I would also like to express gratitude to my assessment committee: Franz Palm, Pierre Mohnen, Toke Aidt, and Jakob de Haan. I look up to their accomplishments in so many ways. Receiving their comments on this book, together with their approval, has been an enormous joy.

I owe a deep debt to my colleagues and friends at the Agence Française de Développement (AFD): Nicolas Meisel, Thomas Roca, Cyrille Bellier and Véronique Sauvat. This dissertation would have never happened without their financial and intellectual support. Working with them in the ‘Institutions and Long-term Growth’ project has been nothing but a rewarding experience. They also introduced me to another set of wonderful people, such as John Wallis, Mushtaq Khan, Lawrence Haddad, Hiroshi Kato, Lawrence Chandy, Charles Kenny, Steven Webb, and countless others during workshops in Paris and elsewhere.

I also benefited from frequent mini-workshops together with the other project members in Maastricht, among them were Bart Verspagen (who never failed to ask the toughest questions), Franziska Gassmann, Thomas Ziesemer, and my extended supervisory team of three. Furthermore, I had many exciting discussions with the other junior researchers who graduated before me: Biniam Bedasso, Samyukta Bhupatiraju, Luciana Cingolani, and Kristine Farla. Now it is time to join your ranks.

Along the way, I met many fantastic people who took the time to discuss my papers or comment on my presentations, among them were Ajay Shah (thanks for inviting me to India), Jan Egbert-Sturm (thanks for inviting me to Zurich), Jeffrey Wooldridge, D.S. Prasada Rao, Gani Aldashev, Carmine Guerriero, Steven van de Walle, Sascha Becker, Oded Galor, Stefan Klasen, Agustin Casas, Florent Bresson and William Poulliot. I owe

a special thank you to Melanie Krause, who has gone through nearly all of my papers with Argus' eyes and never failed to find yet another typo or spot where they could be improved. Thank you Théophile Azomahou, Raouf Boucekine, Pierre Mohnen and Bart Verspagen for inviting one of my papers to appear in a special issue of Macroeconomic Dynamics.

Several people at the school made my life interesting and easier in many ways: Franziska Gassmann, Lutz Krebs, Zina Nimeh, Robin Cowan, Eveline in de Braek, Mindel van de Laar, Susan Roggen, Howard Hudson, Janneke Knaapen and Herman Pijpers. Depending on their role, they have supported me, taught me during the first year of the PhD, helped coordinate my teaching, and helped with project websites, or (as in the case of Susan) never failed to smile when presented with yet another administrative question or travel claim. Thank you!

I am also indebted to Martin Gassebner who made my transition into the (post-doctoral) "afterlife" so seamless by offering me a great job. I am glad he is joining the festivities celebrating the conclusion of this dissertation.

A huge thank you needs to be extended to my paranymphs: Andrea Franco-Correa and Elisa Calza. These two have been there in all circumstances and played many roles in my life. They have been my friends, housemates, and intellectual sparring partners. They readily provided convenient distractions and much needed motivation. Together with Alison, they are known as "the aunts". They are joined by many in the UM/UNU-MERIT building who have been great people to have a coffee with, many of which I am glad to call my friends: Paula, Serdar, Guney, Tobias, Hampton, Ayo, Saba, Mary, Martin, Elvis, Nana, Irina, Muid, Luciana, Kristine, Juan Carlos, Ibrahima, Maty, Florian, Jennifer, Oxana and many more. Then there is my original cohort (Andrea, Paula, Serdar, Patricia, Mahmut, Corinne, Marta, Hampton, Omar, Florence, Yulia, Valery and Michaela) and my many office mates throughout these years (you know who you are).

I would have not achieved any of this without my family: Katharina, Harald, Christel and Stefan. Together they showered me with love and support, but also instilled in me a deep understanding of what it means to be an academic. They were always ready to give the best of advice.

Last but not least, I would like to thank a special someone. Meeting Gerda in the finishing stretches of this dissertation apparently had the following effect: "I have never seen anyone who is so relaxed about finishing their PhD" (Elisa Calza, fall 2014). It seems sometimes hard work *is* followed by the loveliest of rewards.

Richard Bluhm
Berlin, February 11, 2015

Contents

Acknowledgements	v
-------------------------	----------

1 Introduction	1
-----------------------	----------

Part I Political institutions and economic crises

2 The dynamics of stagnation	15
1 Introduction	16
2 Related literature	17
3 Growth episodes and long-run growth	18
4 Explanatory variables	22
5 Empirical strategy	25
6 Results and discussion	30
7 Concluding remarks	38
3 Do weak institutions prolong crises?	45
1 Introduction	46
2 Identifying slumps	47
3 Data and characteristics of slumps	52
4 The duration of declines	56
5 Concluding remarks	67
4 Ethnic divisions, political institutions and the duration of declines	79
1 Introduction	80
2 Empirical motivation	80
3 Related literature	82
4 Theory	84
4.1 Basic setup	84
4.2 Extensions: asymmetric and multigroup settings	90
5 Empirics and Discussion	92
6 Concluding remarks	101

Part II Growth, distribution and poverty

5 The pace of poverty reduction	111
1 Introduction	112
2 Modeling poverty and elasticities	113
2.1 Traditional approaches: linear models of poverty changes	113

2.2	Alternative approaches: non-linear models of poverty	116
2.3	Econometrics of fractional response models	117
3	Data	120
4	Results	122
4.1	Fractional response models	122
4.2	Projecting poverty	130
5	Concluding remarks	133
6	Poor trends	147
1	Introduction	148
2	Drawing the line: international poverty lines	149
3	Taking stock: poverty reduction over the past three decades	152
4	Going forward: poverty projections until 2030	158
5	Concluding remarks and policy recommendations	165
7	Concluding remarks	173
	Valorization	179
	Samenvatting	183
	About the author	189

Chapter 1

Introduction

Is there some action a government of India could take that would lead the Indian economy to grow like Indonesia's or Egypt's? If so, what, exactly? If not, what is it about the 'nature of India' that makes it so? The consequences for human welfare involved in questions like these are simply staggering: Once one starts to think about them, it is hard to think about anything else.

Robert Lucas (1988, p. 5)

ROBERT LUCAS' celebrated quote is often paraphrased as “once you start thinking about economic growth, it is hard to think about anything else”. While Lucas would certainly select two different countries for his example today and many parts of the world have seen large welfare increases since the late 1980s, the quote has not lost its relevance. Understanding why some countries are poor and others are rich, and how this gap can be closed, remains the most fundamental problem in development economics, macroeconomics and economic history.

The answers economists have offered to this question varied over the decades and can be characterized by a progression from proximate to ultimate causality (Maddison, 1988; Szirmai, 2012). Early neoclassical growth theory provided important insights into the proximate sources of growth; that is, the fundamental role played by capital accumulation, labor and productivity in determining the level of income per capita (Solow, 1956). Endogenous growth theory added increasing returns to scale by incorporating human capital and innovation capacity (Romer, 1986, 1990; Lucas, 1988). Yet these contributions only explain part of the puzzle. The way they look at developing countries usually takes well-functioning markets and efficient governments as a given; features of modern economies that are usually absent, incomplete or under construction in developing countries.

Economic historians have long searched for more “ultimate” sources of economic growth to explain the onset of the industrial revolution or the ‘great divergence’ (Pritchett, 1997; Pomeranz, 2009). New Institutional Economics (NIE) combined historical analysis with the Coasian transaction cost approach and unified the analysis of institutions, rules and norms with the neoclassical perspective (e.g. North and Thomas, 1973; Williamson, 1985; North, 1990; Engerman and Sokoloff, 2000). This rediscovery of the primacy of institutions is also chiefly responsible for reintegrating political economy into mainstream micro and macroeconomics, both through advances in formal theory

(e.g. Milgrom et al., 1990; Persson and Tabellini, 2000; Acemoglu and Robinson, 2006; Greif, 2006; Besley and Persson, 2011b) and a flurry of empirical studies highlighting the effects of (political) institutions on economic growth (e.g. Mauro, 1995; Knack and Keefer, 1995; Hall and Jones, 1999; Acemoglu et al., 2001; Acemoglu and Johnson, 2005). As a result, NIE also represents a major analytical break with its antecedents – early institutionalism (e.g. Veblen, 1899; Commons, 1936; Mitchell, 1910a,b) and post-WWII institutionalism (e.g. Gruchy, 1947; Hirschman, 1958; Myrdal, 1968).

Today, few economists would disagree that political and economic institutions are among the fundamental sources of long-run growth. We may call this paradigm shift the ‘institutional turn’ in economics. It refers to a convergence around the idea that inclusive or open access institutions promote development in the long run, while extractive or limited access institutions retard growth or at best allow it to occur temporarily (Acemoglu and Robinson, 2006; North et al., 2009). According to this perspective, fundamental changes in institutions are rare and only occur at ‘critical junctures’ which, as in the case of colonialism, can potentially reverse the relative position of countries in the world income distribution (Engerman and Sokoloff, 2000; Acemoglu et al., 2002). However, the causes of long-run development remain a vibrant and disputed field of research. Many other – often complementary, sometimes contradictory – sources of growth are being championed in the empirical and theoretical literatures. The suggestions range from culture (Weber, 1905; Tabellini, 2008) over geography (Diamond, 1997; Gallup et al., 1999) and technology (Schumpeter, 1934; Nelson and Winter, 1982) to genetic diversity (Spolaore and Wacziarg, 2009; Ashraf and Galor, 2013), but interactions among these various sources of growth are only poorly understood.

Disappointing growth experiences, policy experiments and academic thought interacted throughout the 1980s, 1990s and early 2000s. The failure of the Washington consensus – with its narrow focus on macroeconomic stability, trade liberalization and price stabilization – to generate sustained growth put institutions and political economy considerations back on the map. Two reactions emerged to this perceived failure: *i*) the reforms did not go far enough (Krueger, 2002; Kuczynski and Williamson, 2003, but also see the discussion in Rodrik, 2006), and *ii*) the reform agenda should be broadened to include neostructuralist ideas, second-best solutions and more socially-oriented reforms (e.g. Stiglitz, 2008). The first reaction brought about the term ‘good governance’ within the international financial institutions (IFIs) and coincided with the institutional turn. In addition, the renewed attention from policy circles quickly revealed that early studies of the economic effects of institutions did not include much in the way of “useful ideas on how to implement institutional reforms” (Naím, 2000, p. 94), or, helpful suggestions with respect to which particular institutions require reform, for that matter.

The instability of growth itself received renewed attention as the developed world experienced the ‘great moderation’ from the mid-1980s until the early 2000s, while economic volatility remained high in the developing world and it was noted that there is little persistence of growth across decades (Easterly et al., 1993). Such stylized facts spurred a body of research that questioned the established practice of analyzing growth determinants by looking at levels or average growth rates of GDP per capita (including the focus on conditional or unconditional convergence as in Barro, 1991). Instead, it highlighted the uneven nature of growth in the developing world which can be much more easily characterized by sequences of qualitatively different episodes, such as accelerations, collapses, stagnation and recovery (e.g. Pritchett, 2000; Hausmann et al., 2005; Jones and Olken, 2008). While this certainly expanded our understanding of the dynamics of

growth, most of the literature did not explore the theoretical underpinnings of unstable growth. Interestingly, only a few lines before his famous quote, Robert Lucas also reflected on two growth collapses: the declines in Angola and in Iran from the 1960s to the 1970s. He then asserted that we do not need an “economic theory for an account of either of these declines” (Lucas, 1988, p.4). However, we do need a political economy theory of uneven growth to be in a better position to understand what separates developed country business cycles from the immense welfare losses the developing world is experiencing frequently. Could it be that it is not the lack of rapid growth which creates the dividing line between catching up and falling behind, but the failure to achieve sustained growth and to preserve earlier welfare gains because of social conflict and mismanagement? If so, then is fundamentally a political and economic problem.

Taking a political economy perspective when trying to explain uneven growth paths turns Lucas’ quote on its head. The question is no longer *what* a government should do, but *why* is it that the government of India *does not* undertake an action that would otherwise lead the Indian economy to grow like Egypt’s? Or even worse, why do some governments take actions that clearly precipitate economic declines? And, how can we change or influence the behavior of such governments? While it is now well-established that institutions matter, many of these related questions have not yet been convincingly answered. We know very little about the (political) economics of switching between different growth regimes, or how political and economic institutions affect contemporary growth rates vis-à-vis their strong effect on long-run development (Acemoglu et al., 2001). Similarly, the welfare consequences of such interrupted growth paths are only beginning to be explored.

Alluding to Moses Abramovitz (1989), “thinking about growth” is not equivalent to thinking about human welfare or the lack thereof. If we narrowly define poverty as income or consumption poverty, then the poverty of nations is strongly associated with the poverty of people. The relationship between growth, inequality and poverty is inherently non-linear and within-country inequalities play an important role. There is substantial heterogeneity in how different countries deal with (or have historically dealt with) the type of growth that occurs and the distributional concerns it raises. Understanding this process thoroughly is both an empirical challenge, due to paucity of data and the complexity of income distributions, and a theoretical challenge, for there is not much known about how political institutions indirectly influence poverty through economic growth and, more importantly, changes in distribution.

This dissertation is about the empirical analysis and political economics of growth and inequality. It contributes to fragments of the larger questions posed in this introduction. It addresses different aspects of the development puzzle in two parts: understanding how political institutions affect economic crises, and understanding the relationship between growth, inequality and poverty reduction. Part I is a collection of three essays focusing on the empirics and theory of political institutions and economic crises. The first essay empirically examines the role of institutions and macroeconomic factors in how countries switch in and out of stagnation episodes. The second essay focuses on the econometric identification of the duration of economic declines and shows that high ethnic diversity in combination with weak political institutions is associated with longer declines. The third and final essay of Part I outlines and tests a political economy theory of delayed cooperation during economic slumps which can both explain the findings of the second essay and generate interesting additional insights. Part II consists of two empirical essays on growth, distribution and poverty. The fourth essay is a contribution to the

poverty decomposition literature. It is an applied econometrics paper that outlines a new framework for estimating poverty elasticities and predicting poverty headcount ratios with important substantive implications. The last essay then uses this framework to study a policy question, namely: can extreme poverty be ended by 2030?

The remainder of this introduction outlines the arguments made in the following chapters, focusing on the main research questions, the contributions to the respective literatures and the common themes.

Part I: Political institutions and economic crises

In the spirit of Douglas North (1990), we understand institutions as the humanly devised constraints that structure human interaction; that is, *institutions are the rules of the game*. To a large extent, institutions determine the scope and degrees of freedom for policy making. Together with policies and culture, they provide the incentives which guide the behavior of economic actors. This definition of institutions is deliberately open-ended and encompasses formal, as well as informal institutions, norms and customs, but not culture as is sometimes done in other characterizations (e.g. Greif, 2006).

Throughout the development process, politics and economics are first intimately linked and then become separated through the formalization of rights, the creation of independent organizations and a diffusion of political and economic power (North et al., 2009). There is some disagreement in the literature on the primacy of economic institutions (e.g. Engerman and Sokoloff, 2000) versus political institutions (e.g. Acemoglu and Johnson, 2005; North et al., 2009). By focusing on the role of political institutions, this dissertation implicitly relies on the ‘hierarchy of institutions’ outlined by Acemoglu and Johnson (2005) which states that political institutions define the structure and scope of economic institutions. Power is constrained by political institutions and exercised through them. The economic theory of democratization is a classic example of this hierarchy. In non-democracies, ruling elites may not be able to credibly promise to implement economic institutions that are favorable to the citizens and may instead choose to stave off a threat of revolution by democratizing (Acemoglu and Robinson, 2006). Thus only a change of political institutions brings about more egalitarian economic institutions.

The overarching research question for the chapters in Part I is: how do political institutions affect economic crises? The answer necessarily remains partial. The first two essays examine two different aspects of related phenomena: *i*) the transition into and out of economic stagnation and *ii*) the duration of economic declines. The central insight derived from the empirical analysis is that institutional quality, which changes slowly, cannot be convincingly linked to the timing of transitions into and out of stagnation episodes. However, it does affect the duration of the decline segment; that is, the time until a recovery starts. Furthermore, being able to identify an effect of institutions on crises hinges on the definition of a crisis and what types of crises are being examined. The overarching themes that reoccur throughout these chapters are: the econometric identification and analysis of crises, non-linearities in the growth process, and how political institutions interact with different types of ethno-political heterogeneity.

The first essay in Chapter 2 borrows a very simple definition of an economic stagnation episode from Hausmann et al. (2008). A stagnation episode begins with a contraction in GDP per capita at a time when GDP per capita was higher than ever before and ends when it is again at or above its pre-stagnation level. The chapter then proceeds to

analyze the transitions in and out of such periods of economic stagnation in a panel of countries. We address two questions within the wider research agenda on the instability of growth. First, we ask if institutional characteristics and political shocks determine the incidence of stagnation, and then we ask how these effects compare to standard macroeconomic explanations. In terms of political variables this involves positive and negative regime changes, sudden exits of leaders, and outright wars or civil conflict. In terms of macroeconomic variables, this refers to inflation, financial openness or trade openness, among others. Second, we analyze if any of the included variables have a different impact on the onset of stagnation than on its continuation. In other words, we examine if the variables associated with a higher or lower probability of *falling into* stagnation are the same as those associated with a higher or lower probability of *continuing to be* in stagnation.

Econometrically, we study the determinants of stagnation episodes using dynamic linear and non-linear models; that is, fixed effects models, Generalized Method of Moments (system GMM), fixed effects logit, and a dynamic random effects probit estimator proposed by Wooldridge (2005). We treat stagnation spells as a dynamic problem, subject to state dependence, and pay special attention to the estimation of partial effects of interactions in non-linear (dynamic) models. The main findings are that inflation, negative regime changes, real exchange rate undervaluation, financial openness, and trade openness have statistically and economically significant effects on both the onset and the continuation of stagnation. Only for trade openness there is robust evidence of a differential impact: open economies have a significantly lower probability of falling into stagnation, but once in stagnation they do not recover faster. Overall, predicting the onset of stagnation with macroeconomic factors works rather well, while institutional factors other than negative regime changes are not robustly related to the incidence of stagnation episodes. This finding is not too surprising, since institutions change slowly while we can observe countries moving in and out of stagnation from year to year.

Motivated by the conclusions and lessons learned in the previous chapter, Chapter 3 takes a different approach. In this chapter, we characterize economic crises by a trend break or shift in the growth regime with a restricted pattern. The main idea is twofold. On the one hand, we want to exclude business cycle recessions and instead focus on large, unexpected and negative departures from a previously positive trend in GDP per capita. We call these episodes economic slumps. On the other hand, we now focus exclusively on the decline phase, as the dynamics of decline and recovery may be very different (both empirically and theoretically). We address three larger research questions. First, how can we identify large economic slumps empirically? Second, is there any evidence of institutional change when slumps occur? Third, conditional on the occurrence of a slump, do weak institutions prolong the duration of the decline phase?

To identify slumps and their associated declines, we use a variant of a restricted structural change approach (Papell and Prodan, 2012) together with a bootstrap significance test (Diebold and Chen, 1996) and then date the empirical trough using a simple rule. These techniques draw on the insights of a large literature on sequential and multiple testing for structural breaks in econometrics (e.g. Bai, 1997, 1999; Bai and Perron, 1998, 2003) which inspired a number of studies in the growth literature (Hausmann et al., 2005; Jones and Olken, 2008; Berg et al., 2012). We identify 58 slumps in GDP per capita in 138 countries between 1950 and 2008. Three findings stand out. First, slumps occur frequently and in many cases the decline phase lasts for a long time. Second, we find systematic evidence of weak political institutions before slumps hit and

positive institutional change during and in the immediate aftermath of slumps. Third, the duration of declines decreases with stronger political institutions but increases with greater degrees of ethnic cleavages (ethno-linguistic fractionalization). We also provide evidence of a more subtle non-linear effect suggesting that strong political institutions can potentially help to overcome the adverse effects of ethnic heterogeneity. It turns out that the onset of a slump may be brought about by many factors which are not necessarily related to a country's political institutions or level of social cohesion, but the duration of declines is strongly associated with these two variables. Could it be that the duration of declines depends on ethno-political groups agreeing on coordinated responses and on the political institutions governing this process?

Chapter 4 provides a political economy theory of delayed recovery that can explain the empirical findings from Chapter 3 and then empirically evaluates the theory. The focus is now on the process of agreement on a policy response during the decline phase of a slump. In this chapter, we draw on insights from the literatures on policy reform (Alesina and Drazen, 1991; Fernandez and Rodrik, 1991), political institutions (e.g. Acemoglu and Robinson, 2006; Besley and Persson, 2011a), and the effects of ethnic diversity on growth (e.g. Easterly and Levine, 1997; Alesina and Ferrara, 2005) and suggest a novel answer to the question why economic declines in some parts of the globe last so much longer than in others. Our main contribution is to illustrate a simple mechanism of how weak constraints on the political executive can lead to longer declines in ethnically heterogeneous countries. We emphasize a commitment problem among potential winners and losers of the recovery process. Uncertain post-recovery incomes and changes in the political power distribution may lead to a 'winner-take-all' effect that renders a cooperative equilibrium inaccessible. Absent constraints on the political executive, groups might find it in their interest to hold out and delay their cooperation in order to reduce uncertainty about the future distribution of income. Placing strong constraints on the executive solves this commitment problem by reducing the uncertainty and limiting the potential losses, which brings about cooperation earlier on than under weak constraints.

We derive several theoretical and empirical predictions from the model. First, delay can occur in equilibrium. Weak constraints on the executive are a political economy friction in ethnically diverse countries that can lead to large social inefficiencies. Second, stronger political institutions can entirely resolve this issue and bring about cooperation early on. Third, the commitment problem is getting worse when the number of groups increases. We then show empirically that the partial correlations are consistent with the proposed theory. The effect of executive constraints on the length of declines is very large in heterogeneous countries, but practically disappears in ethnically homogeneous societies. More subtle predictions of the model are confirmed as well. Finally, we suggest that these results are particularly relevant for understanding economic declines in Africa where the longest and deepest declines tend to occur, where politics are shaped by ethnicity, and where weak institutions govern the political executive.

Part II: Growth, distribution and poverty

The second part of this dissertation examines how poverty reacts to income growth and changes in inequality. Before summarizing the individual contributions, it is useful to settle on a definition of poverty that is used throughout the dissertation. While poverty is undeniably a multidimensional phenomenon, this dissertation focuses exclusively on monetary poverty, specifically income or consumption poverty. Monetary poverty is still the benchmark for global poverty measurement, can be considered a multidimensional index capturing all things money can buy, and is well-defined by a wealth of literature dealing with the mathematical and statistical properties of income distributions. There are many measures of poverty which typically vary in their degree of sensitivity to distributional concerns. The most commonly used are the Foster-Greer-Thorbecke (FGT) class of poverty measures (Foster et al., 1984). We limit our attention to the simplest of all: the fraction of the population that is poor, or the poverty headcount ratio.

Looking at this research question from the perspective of development economics means that the emphasis is on absolute poverty, where the poverty line is typically fixed at comparatively low values, and not relative poverty, where the poverty line varies with a measure of the general standard of living such as the mean or median (see also Foster, 1998). Absolute poverty is still pervasive in the developing world – in 2010 more than 20% of the world population were living below \$1.25 a day (in 2005 PPPs) – whereas relative poverty also occurs in the developed world (Chen and Ravallion, 2013). Both chapters deal with global poverty rates and thus rely on the well-established, although controversial, \$1.25 and \$2 a day international poverty lines (in 2005 PPPs). This places the following discussion firmly in the tradition of contributions like Ravallion et al. (1991), Ravallion et al. (2009), and Chen and Ravallion (2010).

The main research question in this part is: how much does the poverty headcount ratio react to changes in average income and distribution? The essays look at this issue in two complementary ways: *i*) how have income (consumption) growth and changes in income (consumption) inequality affected poverty reduction in the past (from 1981 to 2010), and *ii*) what are the likely trends in the future, and what are the implications for the formulation of global development goals. The main lessons are also twofold. First, using a new econometric framework we detect a stable relationship between these three variables in the past and argue that semi-elasticities identify a larger role for redistribution than previously suggested. Second, moving this relationship forward under sensible assumptions suggests that extreme poverty is not likely to end by 2030, even with improvements in distribution. Geographically, poverty will be heavily concentrated in Sub-Saharan Africa and South Asia. Other key themes that reoccur in these chapters are: the econometrics of limited dependent variables (specifically fractions), the estimation and interpretation of income and inequality elasticities of poverty versus semi-elasticities, and the complex issues surrounding poverty measurement.

Chapter 5 revisits the debate on the pace of poverty reduction through growth versus redistribution – a key question in development economics. Earlier contributions have generated important insights. For example, we know that poverty is linked to income and distribution through a decomposition identity and that poverty reduction depends on the initial levels of income and inequality (Bourguignon, 2003). Regional heterogeneity in these initial levels also explains most of the regional heterogeneity in income elasticities (Kalwij and Verschoor, 2007) and about two thirds of poverty reduction in the last decades can be attributed to growth (Kraay, 2006). In this chapter, we suggest that the previous

literature (Ravallion and Chen, 1997; Bourguignon, 2003; Kalwij and Verschoor, 2007) misses an important point, namely that the poverty headcount ratio is a fraction. These studies usually rely on linear specifications and present estimates of income and inequality elasticities that are both inaccurate away from the overall mean and imprecisely estimated over the cross-country distribution of income and inequality. Our main contribution is to directly build in the inherent boundedness of the poverty headcount ratio. To this end, we propose a ‘fractional response approach’ to estimating poverty elasticities, which is able to very closely approximate the distribution of income or consumption near the poverty line without making specific distributional assumptions.

We then estimate income and inequality (semi-)elasticities for the \$2 and \$1.25 a day poverty lines using the largest data set to-date. Econometrically, this is a challenging enterprise as there are countless differences in how household surveys are conducted across countries, both incomes and poverty rates are likely to suffer from measurement error, and surveys are undertaken at very infrequent intervals. Building on Papke and Wooldridge (1996, 2008) and Wooldridge (2010), we present extensions of the fractional response approach that solve these issues by allowing for unobserved heterogeneity, measurement error, and unbalancedness of the panel. We then show that we can precisely estimate these quantities of interest over their entire distribution with our framework, and highlight the relevance of focusing on semi-elasticities, not elasticities, for poverty reduction policies. In addition, the chapter shows that this method can be used as direct way of projecting poverty rates into the future.

Chapter 6 is a policy-oriented contribution. In this chapter, we apply the empirical framework from Chapter 5 to study the feasibility of the World Bank’s new goal of reducing the \$1.25 a day global poverty rate to 3% by 2030 (from about 20% in 2010). We begin by discussing the origins of the dollar-a-day poverty line, conduct a sensitivity analysis of this line, and describe the historical trends in poverty and inequality around the world. This serves two purposes. First, we want to highlight that the extreme poverty line of \$1.25 a day is not a fixture but subject to a considerable degree of uncertainty. Many other lines are plausible. Second, and more importantly, we emphasize that precisely those countries that have historically contributed most to poverty reduction are not likely to do so in the future; they experienced GDP (consumption) growth at unusually fast rates in the last two decades. As a result, this changing geography of world poverty has profound implications for the pace of poverty reduction in the future.

We then proceed to empirically evaluate a pessimistic, moderate and an optimistic growth scenario, as well as a pro-poor, distribution neutral and pro-rich scenario for the evolution of inequality. This analysis leads us to two main findings. First, the pace of poverty reduction at \$1.25 a day is likely to slow down. Even our most optimistic scenarios suggest a global poverty rate of 8% to 9% in 2030, far short of the World Bank’s new target of 3%. Second, rapid progress can be maintained at \$2 a day, with an additional one billion people crossing that line by 2030. Both of these results include significant reductions in within-country inequality. On this basis, we propose two new ‘twin targets’ which are aspirational but also achievable. We suggest reducing the proportion of the population living below \$1.25 to 8% by 2030 and reducing the proportion of the population living below \$2 a day to 18% by 2030. Reaching these targets requires an additional acceleration of growth or significant improvements in distribution.

Chapter 7 concludes the dissertation. It answers the two overarching research questions, looks back on the lessons learned and looks forward by discussing open issues.

Bibliography

- Abramovitz, M. (1989). *Thinking about Growth*. Cambridge University Press.
- Acemoglu, D. and S. Johnson (2005). Unbundling institutions. *Journal of Political Economy* 113(5), 949–95.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5), 1369–1401.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics* 117(4), 1231–1294.
- Acemoglu, D. and J. A. Robinson (2006). *Economic origins of dictatorship and democracy*. Cambridge University Press.
- Alesina, A. and A. Drazen (1991). Why are stabilizations delayed? *American Economic Review* 81(5), 1170–1188.
- Alesina, A. and E. L. Ferrara (2005). Ethnic diversity and economic performance. *Journal of Economic Literature* 43(3), 762–800.
- Ashraf, Q. and O. Galor (2013). The “Out of Africa” hypothesis, human genetic diversity, and comparative economic development. *American Economic Review* 103(1), 1–46.
- Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory* 13, 315–352.
- Bai, J. (1999). Likelihood ratio tests for multiple structural changes. *Journal of Econometrics* 91(2), 299–323.
- Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.
- Bai, J. and P. Perron (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18(1), 1–22.
- Barro, R. J. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics* 106(2), 407–443.
- Berg, A., J. D. Ostry, and J. Zettelmeyer (2012). What makes growth sustained? *Journal of Development Economics* 98(2), 149–166.
- Besley, T. and T. Persson (2011a). The logic of political violence. *Quarterly Journal of Economics* 126(3), 1411–1445.
- Besley, T. and T. Persson (2011b). *Pillars of prosperity: The political economics of development clusters*. Princeton University Press.
- Bourguignon, F. (2003). The growth elasticity of poverty reduction: Explaining heterogeneity across countries and time periods. In T. S. Eicher and S. J. Turnovsky (Eds.), *Inequality and Growth: Theory and Policy Implications*, pp. 3–26. Cambridge, MA: MIT Press.
- Chen, S. and M. Ravallion (2010). The developing world is poorer than we thought, but no less successful in the fight against poverty. *Quarterly Journal of Economics* 125(4), 1577–1625.
- Chen, S. and M. Ravallion (2013). More relatively-poor people in a less absolutely-poor world. *Review of Income and Wealth* 59(1), 1–28.
- Commons, J. R. (1936). Institutional economics. *American Economic Review* 26(1), 237–249.
- Diamond, J. (1997). *Guns, germs, and steel: The fates of human societies*. W.W. Norton & Company.
- Diebold, F. X. and C. Chen (1996). Testing structural stability with endogenous breakpoint a size comparison of analytic and bootstrap procedures. *Journal of Econometrics* 70(1), 221–241.
- Easterly, W., M. Kremer, L. Pritchett, and L. Summers (1993). Good policy or good luck? *Journal of Monetary Economics* 32(3), 459–483.
- Easterly, W. and R. Levine (1997). Africa’s growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics* 112(4), 1203–1250.
- Engerman, S. L. and K. L. Sokoloff (2000). Institutions, factor endowments, and paths of development in the new world. *Journal of Economic Perspectives* 14(3), 217–232.
- Fernandez, R. and D. Rodrik (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *American Economic Review* 81(5), 1146–1155.
- Foster, J., J. Greer, and E. Thorbecke (1984). A class of decomposable poverty measures. *Econometrica* 52(3), 761–766.
- Foster, J. E. (1998). Absolute versus relative poverty. *American Economic Review* 88(2), 335–341.
- Gallup, J. L., J. D. Sachs, and A. D. Mellinger (1999). Geography and economic development.

- International Regional Science Review* 22(2), 179–232.
- Greif, A. (2006). *Institutions and the path to the modern economy: Lessons from medieval trade*. Cambridge University Press.
- Gruchy, A. G. (1947). *Modern economic thought: The American contribution*. Prentice-Hall.
- Hall, R. E. and C. I. Jones (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* 114(1), 83–116.
- Hausmann, R., L. Pritchett, and D. Rodrik (2005). Growth accelerations. *Journal of Economic Growth* 10(4), 303–329.
- Hausmann, R., F. Rodriguez, and R. Wagner (2008). Growth collapses. In C. Reinhart, C. Vegh, and A. Velasco (Eds.), *Money, Crises and Transition*, pp. 376–428. Cambridge, Mass.: MIT Press.
- Hirschman, A. O. (1958). *The strategy of economic development*. Yale University Press.
- Jones, B. F. and B. A. Olken (2008). The anatomy of start-stop growth. *Review of Economics and Statistics* 90(3), 582–587.
- Kalwij, A. and A. Verschoor (2007). Not by growth alone: The role of the distribution of income in regional diversity in poverty reduction. *European Economic Review* 51(4), 805–829.
- Knack, S. and P. Keefer (1995). Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics & Politics* 7(3), 207–227.
- Kraay, A. (2006). When is growth pro-poor? Evidence from a panel of countries. *Journal of Development Economics* 80(1), 198–227.
- Krueger, A. O. (2002). *Economic policy reform: The second stage*. University of Chicago Press.
- Kuczynski, P.-P. and J. Williamson (2003). *After the Washington Consensus: Restarting growth and reform in Latin America*. Peterson Institute.
- Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics* 22(1), 3–42.
- Maddison, A. (1988). Ultimate and proximate growth causality: A critique of Mancur Olson on the rise and decline of nations. *Scandinavian Economic History Review* 36(2), 25–29.
- Mauro, P. (1995). Corruption and growth. *Quarterly Journal of Economics* 110(3), 681–712.
- Milgrom, P. R., D. C. North, and B. R. Weingast (1990). The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. *Economics & Politics* 2(1), 1–23.
- Mitchell, W. C. (1910a). The rationality of economic activity: I. *Journal of Political Economy* 18(2), 97–113.
- Mitchell, W. C. (1910b). The rationality of economic activity: II. *Journal of Political Economy* 18(3), 197–216.
- Myrdal, G. (1968). *An inquiry into the poverty of nations*, Volume 1-3. Pantheon.
- Naím, M. (2000). Washington consensus or washington confusion? *Foreign Policy* (118), 86–103.
- Nelson, R. R. and S. G. Winter (1982). *An evolutionary theory of economic change*. Harvard University Press.
- North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.
- North, D. C. and R. P. Thomas (1973). *The rise of the western world: A new economic history*. Cambridge University Press.
- North, D. C., J. J. Wallis, and B. R. Weingast (2009). *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge and New York: Cambridge University Press.
- Papell, D. H. and R. Prodan (2012). The Statistical Behavior of GDP after Financial Crises and Severe Recessions. *The B.E. Journal of Macroeconomics* 12(3), 1–31.
- Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics* 11(6), 619–632.
- Papke, L. E. and J. M. Wooldridge (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics* 145(1–2), 121–133.
- Persson, T. and G. Tabellini (2000). *Political Economics: Explaining Economic Policy*. The MIT Press.
- Pomeranz, K. (2009). *The great divergence: China, Europe, and the making of the modern world economy*. Princeton University Press.
- Pritchett, L. (1997). Divergence, big time. *Journal of Economic Perspectives* 11(3), 3–17.
- Pritchett, L. (2000). Understanding patterns of economic growth: searching for hills among plateaus,

- mountains, and plains. *World Bank Economic Review* 14(2), 221–250.
- Ravallion, M. and S. Chen (1997). What can new survey data tell us about recent changes in distribution and poverty? *World Bank Economic Review* 11(2), 357–382.
- Ravallion, M., S. Chen, and P. Sangraula (2009). Dollar a day revisited. *World Bank Economic Review* 23(2), 163–184.
- Ravallion, M., G. Datt, and D. van de Walle (1991). Quantifying absolute poverty in the developing world. *Review of Income and Wealth* 37(4), 345–361.
- Rodrik, D. (2006). Goodbye Washington Consensus, hello Washington confusion? A review of the World Bank’s economic growth in the 1990s: Learning from a decade of reform. *Journal of Economic Literature* 44(4), 973–987.
- Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy* 94(5), 1002–1037.
- Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy* 98(5), S71–S102.
- Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*. Transaction Publishers.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70(1), 65–94.
- Spolaore, E. and R. Wacziarg (2009). The diffusion of development. *Quarterly Journal of Economics* 124(2), 469–529.
- Stiglitz, J. E. (2008). Is there a post-washington consensus consensus? In N. Serra and J. E. Stiglitz (Eds.), *The Washington Consensus Reconsidered*, Chapter 5, pp. 41–56. Oxford University Press.
- Szirmai, A. (2012). Proximate, intermediate and ultimate causality: Theories and experiences of growth and development. MERIT Working Papers 032, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).
- Tabellini, G. (2008). Presidential address institutions and culture. *Journal of the European Economic Association* 6(2-3), 255–294.
- Veblen, T. (1899). *The theory of the leisure class*. Macmillan.
- Weber, M. (1905). *The Protestant Ethic and the Spirit of Capitalism*. Routledge Classic, 2001.
- Williamson, O. E. (1985). *The economic institutions of capitalism*. The Free Press.
- Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics* 20(1), 39–54.
- Wooldridge, J. M. (2010). Correlated random effects models with unbalanced panels. Manuscript.

Part I

Political institutions and economic crises

Chapter 2

The dynamics of stagnation

*A panel analysis of the onset and continuation of stagnation**

Abstract

This chapter analyzes periods of economic stagnation in a panel of countries. We test whether stagnation can be predicted by institutional characteristics and political shocks, and compare the impacts of such variables with those of traditional macroeconomic variables. We examine the determinants of stagnation episodes using dynamic linear and non-linear models. In addition, we analyze whether the effects of the included variables on the onset of stagnation differ from the effects on the continuation of stagnation. We find that inflation, negative regime changes, real exchange rate undervaluation, financial openness, and trade openness have significant effects on both the onset and the continuation of stagnation. Only for trade openness there is robust evidence of a differential impact. Open economies have a significantly lower probability of falling into stagnation, but once in stagnation they do not recover faster.

Keywords: growth episodes, stagnation, institutions, dynamic panel data

JEL Classification: O11, O43, C25

*This chapter is based on Bluhm, R., de Crombrugge, D. and A. Szirmai. 2012. Explaining the dynamics of stagnation: An empirical examination of the North, Wallis and Weingast approach, MERIT Working Papers 040, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT), forthcoming in *Macroeconomic Dynamics*.

1 Introduction

Since the 1950s, many countries across the globe have experienced substantial increases in GDP per capita brought about by years of economic growth. However, while these gains are mainly the result of steady positive growth rates in the developed world (at least prior to 2008), growth in developing countries has often been erratic and volatile. Most emerging economies have experienced periods of economic stagnation between positive growth spurts and for several countries the absence of sustained growth has proved to be a persistent phenomenon, often lasting for several years or even decades. Explaining why some countries experience more periods of stagnation than others may thus prove essential to the understanding of contemporary differences in levels of development.

Rather than focusing on differences in average growth rates, recent research increasingly aims to unveil the specific characteristics of different growth episodes such as accelerated growth, growth collapses, recoveries, or stagnation. We address two questions within this wider research agenda. First, we ask if institutional characteristics and external or internal political shocks determine the incidence of stagnation, and how these effects compare to standard macroeconomic explanations. Second, we analyze if any of the included variables have a different impact on the onset of stagnation than on its continuation. In other words, we examine if the factors affecting the probability of falling into stagnation are the same as those affecting the probability of continuing to be in stagnation.

Most of the empirical literature on growth episodes uses static models to study factors that are correlated with the onset of a growth spell and, more recently, began to examine factors associated with the duration of growth episodes. Our contribution is to analyze stagnation spells as a dynamic problem, subject to state dependence and interactions between the lagged state and the independent variables. This approach allows the probability of stagnation to depend on whether a country was already in stagnation in the preceding year (state dependence). It lets the data decide whether the included variables have a different effect on the onset of a stagnation episode than its continuation. We estimate the dynamic models using linear probability models, GMM, fixed effects logit, and a dynamic random effects probit estimator proposed by Wooldridge (2005).

Our results indicate that political regime shifts towards autocracy have strong positive effects on the incidence of stagnation (onset and continuation), but other proxies for institutions and political shocks do not have significant effects. Macroeconomic factors explain the onset of stagnation rather well. Higher inflation positively predicts stagnation, while financial openness, trade openness, and real exchange rate undervaluation are associated with a reduced likelihood of stagnation. We find little evidence that the effects of these variables differ between the onset and continuation of stagnation spells. Only trade openness has robustly different effects. It substantially reduces the chances of falling into stagnation, but the ‘protective’ effect of openness vanishes once a country is in a stagnation spell. In addition, we find that stagnation spells exhibit a moderate degree of state dependence, which is consistent with other results in the literature on the duration of growth collapses.

The sequel is organized as follows. Section 2 briefly reviews the literature on institutions and growth, and discusses applications of the growth episodes approach. Section 3 defines stagnation episodes and explores their correlations with GDP levels and institutions. Section 4 describes the variables and data. Section 5 outlines the empirical strategy. Section 6 discusses the results. Section 7 concludes.

2 Related literature

An increasingly large body of literature in economics argues that differences in institutional characteristics are the key to understanding the differences in long-run economic performance of nations. While modern institutional theory has many antecedents, it started from the hypothesis that one explanation for the historical rise of the West is well-developed property rights (e.g. North and Thomas, 1973). Since the 1990s, this literature has been extended to view growth-promoting institutions less narrowly. More recent contributions argue, for example, that institutions for growth are multifaceted (Rodrik, 2000), interact with geography and inequality (Engerman and Sokoloff, 1997), develop semi-endogenously (Greif, 2006), and are embedded in informal arrangements (North, Wallis, and Weingast, 2009).¹

In terms of econometric evidence, several papers have suggested that differences in institutions explain a large part – if not most – of the cross-country variation in *levels* of GDP per capita.² However, many of these studies have also been criticized for their underlying assumptions (e.g. Glaeser, La Porta, Lopez-de Silanes, and Shleifer, 2004) and do generally not establish a link between institutions and growth rates (Crombrugge and Farla, 2012). Potentially bridging this gap in theory, several authors have recently suggested that there is a link between institutional susceptibility to various external or internal shocks and different growth outcomes. North et al. (2009), for example, identify two distinct types of social orders. Open access orders are economically and politically highly developed, experience relatively smooth patterns of economic growth, have active civil societies, many long-lived organizations, heavily formalized rules, and strong rule of law. Large segments of the population have access to political and economic organizations. Limited access orders, on the contrary, are dominated by elites that exclude large parts of the population from access to economic and political organizations. The rents created in this process are then distributed among members of the ruling coalition, in order to achieve a basic degree of social stability and control over violence. Limited access orders typically experience volatile growth patterns and are characterized by politics without broad democratic consent, few organizations, informal rules, weak and unequally enforced rule of law, insecure property rights, and high levels of inequality.

North et al. (2009) suggest that limited access orders are inflexible and less able to cope with shocks, thus causing a higher propensity towards growth collapses and stagnation. Rodrik (1999) links negative growth experiences to terms of trade shocks, latent social conflict, and the ability of institutions to contain conflict and absorb the destructive potential of such shocks. A key question for this chapter is to what extent an empirical analysis of stagnation episodes supports these theories. Therefore, we hypothesize that (a) institutional characteristics play an important role in explaining the onset of stagnation and (b) weak institutions prolong the incidence of stagnation spells.

As Pritchett (2000) pointed out, a problem in traditional panel studies of growth rates is that they focus on average trends over a fixed period, while in reality growth is often erratic and may be contingent on very different growth regimes. This conjecture gave birth to a rapidly growing literature, which since has analyzed growth differentials across decades (Rodrik, 1999), growth accelerations (Hausmann, Pritchett, and Rodrik, 2005),

¹For a review of the debates see Bluhm and Szirmai (2012) and for an earlier survey see Aron (2000).

²The list of empirical studies investigating this issue is long and growing but the seminal papers are Knack and Keefer (1995), Hall and Jones (1999), Acemoglu, Johnson, and Robinson (2001, 2002), and Rodrik, Subramanian, and Trebbi (2004).

switching among multiple growth regimes (Jerzmanowski, 2006), the duration of growth collapses (Hausmann, Rodriguez, and Wagner, 2008), start and stop growth (Jones and Olken, 2008), real income stagnation (Reddy and Minoiu, 2009), and the duration of growth accelerations (Berg, Ostry, and Zettelmeyer, 2012).

This chapter relates most to studies focusing on negative growth experiences. Rodrik (1999) provides first evidence that growth collapses are linked to terms of trade shocks, latent conflict, and the conflict management capacity of institutions. Hausmann et al. (2008) examine the onset and duration of growth collapses. They mainly find that weak export performance and high inflation coincide with the onset of stagnation, but downturns also occur together with wars, sudden stops, and political transitions. However, most of these factors have little influence on the duration of collapses, which is only correlated with a measure of the flexibility of a country's export basket. Last, Reddy and Minoiu (2009) investigate the incidence of stagnation spells (periods of negative growth) and find that these are correlated with weak export performance, low investment, primary commodity exports, and weak institutions.

The study of stagnation spells and other negative growth episodes is also related to the business cycle literature and the literature on economic crises. Although the focus of this chapter is primarily on longer-run growth episodes and not on short-run fluctuations, these literatures provide relevant insights and hypotheses (e.g. see Diebold et al. 1993, on duration dependence, Cerra and Saxena 2008, on post-crisis growth, and Bussière and Fratzscher 2006, on recession probabilities).

Most papers in the growth episodes literature use a methodology that can be summarized in two steps. First, a rule-based or statistical filter is applied to the data to identify a single (or multiple) turning point(s) in the GDP series. If the filter is rule-based, then it often includes a criterion implicitly or explicitly defining the length of the spell. If the filter is statistical, then it may find more than one break in the data and thus lead to the identification of distinct growth episodes. In the second step, correlates of these episodes are examined either by testing for differences in means of potentially correlated variables (before and during), or by estimating probit models. The unexamined assumption in these studies is that factors affecting the onset of an episode are the same as those determining if an episode will continue. Further, most studies of growth episodes take very few measures to account for the possible endogeneity of the included regressors.

3 Growth episodes and long-run growth

Defining the growth episodes

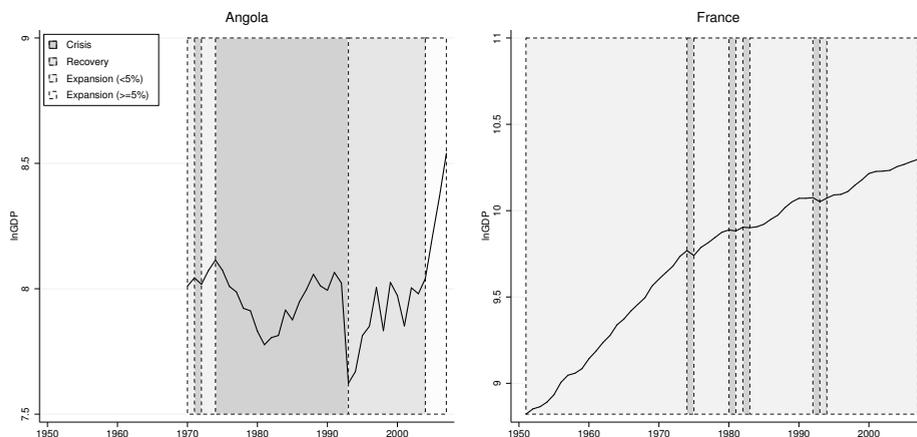
Our classification of growth episodes is a modification of the approach to growth collapses in Hausmann et al. (2008). We define a *stagnation episode* (or stagnation spell) as an event that begins with a contraction of GDP per capita at a time it was higher than ever before, and ends when GDP per capita is again at or above its pre-stagnation level. We denote (for the purpose of this section) the log of GDP per capita in country i in year t by Y_{it} ($i = 1, \dots, N; t = 1, \dots, T_i$). Defined formally, a stagnation episode begins when $Y_{it} < Y_{i,t-1}$ and $Y_{i,t-1} \geq \max_{x=1}^{t-1} Y_{i,x}$, and lasts as long as $Y_{i,t+p} < Y_{i,t-1}$ ($p = 1, 2, \dots$). When $Y_{i,t+p} \geq Y_{i,t-1}$, the stagnation spell is over. Conversely, we define all years when a country is not stagnating as *expansion* years. In other words, an expansion episode begins the first year a country has left (or has not yet experienced) a stagnation spell and lasts until the beginning of the next stagnation spell.

Apart from being very simple, these definitions have several desirable properties. A completed stagnation episode has a net effect of zero on the level of GDP per capita, since it includes both the downturn and the associated recovery. Conversely, the effect of an expansion episode on the level of GDP per capita is always positive. This definition of an expansion explicitly excludes growth that is merely restoring what was lost in past crises, as such growth does not account for long-run increases in GDP per capita. Some commonly used filters, such as the Hausmann et al. (2005) growth accelerations filter, do not make this distinction between recoveries and expansions. Thus some of their growth accelerations include recoveries. (See also Bussière and Fratzscher, 2006, on ‘post-crisis bias’.)

An episode has a minimum duration of one year but can actually last for the entire length of the sampled period (1951–2007). Based on this definition, we can identify long stagnation episodes that may include recurring short-run recessions with incomplete recoveries – incomplete in the sense that the maximum level of GDP per capita prior to the crisis has not been recovered. Stagnation episodes thus deliberately subsume many short-run business cycle fluctuations.

We further differentiate expansions and periods of stagnation into two sub-spells each. For stagnation episodes, we distinguish between *crises*, lasting from the beginning of the stagnation episode to the trough, and *recoveries*, lasting from the year after the trough until the end of the stagnation spell. We define the *trough* to occur at the minimum level of output occurring during a stagnation episode. For expansions, we distinguish between *moderate expansions* with an average growth rate up to 5% per annum, and *rapid expansions* with an average growth rate surpassing 5% per annum.³ In the – rare – case where growth in the first recovery year is so rapid that pre-crisis output is regained in one year, we consider that year part of an expansion, and exclude it from the stagnation spell.

Figure 1 – Examples of growth episodes: Angola and France



³More precisely, we measure the growth rate across an expansion as: $\bar{g}_{t,t+q} = q^{-1}[\ln Y_{i,t+q} - \ln Y_{it}]$, where q is the duration of the expansion. We classify an episode as rapid if $\bar{g}_{t,t+q} > 0.05$, and slow to moderate if $\bar{g}_{t,t+q} \leq 0.05$.

We apply these definitions to GDP per capita data from the *Penn World Table 6.3* (Heston, Summers, and Aten, 2009). Excluding countries with less than one million inhabitants at the latest recorded year as well as countries with less than 20 years of data, we observe 127 countries for a period of at least 20 years between 1950 and 2007. Within this sample, hence before the beginning of the 2007 financial upheaval, we find a total of 578 stagnation episodes, or 3,276 country-years of stagnation.

Figure 1 illustrates how our filter works graphically as applied to Angola and France. These two examples are typical for the different growth experiences of developed and developing countries and show that the filter works reasonably well in identifying episodes of interest. While Angola has had many years of positive growth throughout the sample period, we find only three short expansion spells of which only the last is a rapid expansion. Instead, most of the time, Angola was in one protracted stagnation episode lasting from 1975 until the end of 2004, including significant volatility in between. On the contrary, the French economy grew steadily since 1951 and is characterized by protracted periods of moderate expansion, which are only temporarily interrupted by very short stagnation spells. In the light of these two stylized cases, we propose that the incidence of stagnation spells may explain a large part of the difference in long-run levels of GDP per capita.

Growth profiles

Before focusing on the dynamics of moving in and out of stagnation spells, we first take a more detailed look at the distribution of growth episodes across countries and time. Do developing countries spend more of their time in crisis or stagnation than advanced economies? Are they more prone to experience crisis and stagnation? Using the previously defined growth episodes, Table 1 addresses these issues in more detail.

Table 1 – Growth episodes by income levels in 2007

% Country-years in ...	Low	Low-Mid	Mid-High	High	Total
<i>Panel A: Two episode types</i>					
Expansion	22.12	41.33	54.97	73.14	48.31
Stagnation	77.88	58.67	45.03	26.86	51.69
Total	100.00	100.00	100.00	100.00	100.00
<i>Panel B: Four episode types</i>					
Expansion (above 5%)	10.21	11.52	23.18	22.66	16.99
Expansion (5% or less)	11.91	29.81	31.79	50.49	31.32
Crisis	49.90	30.63	23.81	17.42	30.18
Recovery	27.98	28.04	21.22	9.44	21.51
Total	100.00	100.00	100.00	100.00	100.00

Notes: 127 countries, number of observations 6,338, percentages calculated over 1951 to 2007.

Table 1 groups the relative incidence of each type of growth episode from 1951 to 2007 by quartiles of GDP per capita in 2007. The table shows that low-income countries spend most of their time in stagnation, upper middle-income countries almost half the time and high-income countries only about a quarter. In other words, this suggests that the different propensity to experience stagnation spells is closely linked to development outcomes today. Further, using the finer classification of four distinct growth episodes

– crisis, recovery, expansion, rapid expansion – we find that a high proportion of time spent in crises at low income and lower-middle income levels is driving this relationship. Once we exclude recoveries from the positive growth experiences, there is little indication that lower income countries experience rapid growth more frequently than high income countries during expansions (as unconditional convergence would imply). In fact, the opposite seems to be the case. While countries in the lowest income group spend relatively more of their expansions growing rapidly ($10.21/22.12 \approx 46.15\%$), upper-middle income and high-income countries spend more time growing rapidly in total. Table 1 confirms that currently poor countries have experienced fewer years of positive growth than rich countries. It contradicts the assertion that once poor countries grow, they do so more rapidly than the rich.⁴

As mentioned earlier, North et al. (2009) suggest lasting institutional differences between limited access orders and open access orders as one possible explanation for the lack of generalized convergence among economies. Developing countries with limited access orders are less adaptive, less able to adjust to various external and internal shocks, and more prone to economic crises and stagnation.

Table 2 – Growth episodes by institutional indicators

% Country-years in ...	Low	Low-Mid	Mid-High	High	Total
<i>Panel A: Formalization of regulations</i>					
Expansion	30.49	40.92	54.61	76.41	51.06
Stagnation	69.51	59.08	45.39	23.59	48.94
Total	100.00	100.00	100.00	100.00	100.00
Expansion (above 5%)	14.49	16.15	20.86	18.38	17.47
Expansion (5% or less)	16.00	24.78	33.75	58.03	33.60
Crisis	40.29	37.13	21.80	14.51	28.29
Recovery	29.22	21.95	23.59	9.08	20.65
Total	100.00	100.00	100.00	100.00	100.00
<i>Panel B: Control and intervention</i>					
Expansion	54.98	68.17	50.56	30.55	51.06
Stagnation	45.02	31.83	49.44	69.45	48.94
Total	100.00	100.00	100.00	100.00	100.00
Expansion (above 5%)	25.36	13.77	20.68	11.67	17.47
Expansion (5% or less)	29.62	54.40	29.87	18.88	33.60
Crisis	26.81	18.13	28.39	39.96	28.29
Recovery	18.21	13.70	21.05	29.49	20.65
Total	100.00	100.00	100.00	100.00	100.00

Notes: 47 countries in 1951, 107 countries in 2007, total number of observations 5,405 (Panel A and Panel B), percentages calculated on the basis of all years between 1951 and 2007.

Table 2 links the conjecture of North et al. (2009) and similar theories to the approach developed in this chapter by cross-tabulating the different growth episodes with two indices of institutional characteristics; namely, institutional formalization of regulations and degree of control and intervention by the state. These indices are derived from Crombrugge and Farla (2012), who aggregate a large number of indicators from

⁴We constructed a similar table classifying countries by their GDP per capita in 1960 rather than at the end of the period. Though there were some differences, the basic finding that rich countries spend less of their time in crisis than poor countries is confirmed.

the Institutional Profiles Database (IPD) 2009 using principal components analysis.⁵ Similarly to the income classification used before, we group the scores on each component into quartiles ranked from low to high. The upper panel in Table 2 shows the results for the first component and the lower panel the results for the second.

There is a moderate negative correlation (about -0.5 for 2007) between the index of institutional formalization of regulations and the incidence of stagnation episodes. The countries belonging to the highest quartile on this index are in stagnation less than 25% of the sample period, while those ranked in the lowest quartile stagnate almost 70% of the time. In many ways these results resemble those using income groups. For example, fast expansions occur more frequently in the upper middle quartile and crises occur gradually less often at higher quartiles of the index. This suggests that higher institutional formalization of regulations is associated with fewer stagnation spells and increasingly steady growth. However, there is a strong correlation (about 0.8) between GDP per capita and the formalization index, so the direction of causality remains indeterminate.

The bottom panel of Table 2 gives a more differentiated picture. The second principal component, which can be interpreted as the degree of the state's involvement in the private economy but also as its degree of authoritarianism, is associated with more frequent stagnation spells. The lowest incidence of stagnation spells (31.83%) occurs within the group of countries scoring in the lower middle quartile of the index, whereas countries in the highest quartile stagnated during nearly 70% of the sample period. As Crombrughe and Farla (2012, p. 17) point out, "Western European countries, the USA, Canada, and Australia are at neither extreme of the [index]", which suggests that very low scores represent weak states and very high scores represent mostly authoritarian regimes. This explains why the most stable growth profile is located in the lower middle quartile rather than at either end of the spectrum.

This brief overview of different growth episodes between 1950 and 2007 highlights two points. First, the incidence of stagnation spells is much higher in lower and middle income countries than in high income countries. Second, weak institutions and especially a lack of formalized rules and regulations could be driving less steady growth and more frequent stagnation, but this aspect requires further analysis.

4 Explanatory variables

In the preceding section, we have reported how we defined and obtained our sample of stagnation spells and examined their distribution. In this section, before we start modeling the incidence of stagnation, we briefly outline the sources for and construction of explanatory variables. These broadly belong to two categories: macroeconomic indicators and variables describing political institutions as well as external or internal shocks to these institutions. Table A-1 in Appendix A provides an overview.

Macroeconomic variables. We include a wide range of variables that are typically associated with sound macroeconomic management. Most of these variables have been found to significantly affect growth performance in traditional panel studies using annual, quinquennial (5-year) or decennial (10-year) growth rates.

⁵For more details on the construction of the indices see Crombrughe and Farla (2012). The IPD 2009 is publicly accessible at www.afd.fr/lang/en/home/recherche/bases-ipd.

In order to control for the level of development, we include the lagged log of GDP per capita ($\text{Log } GDPC_{(t-1)}$) in nearly all models. Its expected effect is negative, considering that richer countries tend to experience fewer and shorter stagnation spells. Controlling for the level of GDP also serves a practical purpose. As indicated in the previous section, indices measuring the quality of institutions and GDP are strongly correlated, so that including both will avoid erroneously attributing effects of one to the other.

Maintaining price stability is a core task of central banks and its importance is emphasized in the related literature (e.g. Berg et al., 2012). We expect high inflation to be positively correlated with the onset of stagnation spells. However, the role of inflation is likely to be ambiguous for the continuation of stagnation as it could be instrumental – together with the exchange rate – in bringing about devaluation and regaining competitiveness. Our measure of inflation is 100 times the log of 1 plus the annual inflation rate. This measure is close to the actual inflation rate when that rate is small, but also reduces the influence of larger values (e.g. rare periods of hyperinflation). The annual inflation data is from the IMF's *International Financial Statistics* (IFS), complemented with data from the *World Development Indicators* (WDI) whenever the former is missing.

We also measure whether the exchange rate is overvalued or undervalued in real terms. Recent research finds that depreciations are beneficial for growth accelerations (Hausmann et al., 2005) and stimulate growth in general (Rodrik, 2008). This positive effect may operate through many channels, but is most commonly linked to export-led growth and the relative price of manufactured products. On the negative side, abrupt movements of the exchange rate can also be an omen of excessive volatility and an upcoming currency crisis. If the positive consequences are dominant, exchange rate undervaluation may diminish the likelihood of stagnation spells. To capture this effect, we follow Rodrik (2008) in constructing an index of exchange rate undervaluation ($\text{RER Value}_{(t-1)}$).⁶ The index is centered at zero, with higher values indicating exchange rate undervaluation and lower values indicating overvaluation.

We include two measures of trade performance. First, we measure the price of exports relative to imports, the terms of trade ($\Delta \text{ToT}_{(t-1)}$), as the annual log difference in the net barter terms of trade from the WDI, and supplement this series with data from the IFS when there are gaps in the WDI series. Terms of trade growth, declines and shocks have been linked to growth collapses (Rodrik, 1999; Hausmann et al., 2008), accelerations (Hausmann et al., 2005), and the premature end of fast growth spells (Berg et al., 2012). Second, we also estimate the effects of changes in the value of real merchandise exports ($\Delta \text{Real Exports}_{(t-1)}$), which we measure as the annual log difference in the exports volume index from the WDI, appended with data from the IFS to extend coverage. Growth in real exports has been found to significantly reduce the probability of the onset of a stagnation spell (Hausmann et al., 2008). We examine if this is also the case in the presence of dynamics.

Further, the growth literature has identified *de jure* financial and trade openness as two key policy variables that positively influence growth outcomes. To account for the former, we include the Chinn and Ito (2006) index of financial openness ($\text{Fin. Openness}_{(t-1)}$). This index is the first principal component of the inverses of four variables measuring restrictions on external accounts, based on the IMF's *Annual Report on Exchange Arrangements and Exchange Restrictions* (AREAER). To account for the latter, we use a dummy measure for economic liberalization ($\text{Trade Openness}_{(t-1)}$)

⁶This index is based on the PWT 6.3 and its construction is described in more detail in Table A-1.

developed by Sachs and Warner (1995) and extended by Wacziarg and Welch (2008). This indicator is coded as one in years a given country is completely open to trade and zero otherwise. While the index's authors have linked their respective measures to average growth rates, the growth episodes literature has found financial openness to precede growth accelerations (Hausmann et al., 2005) and trade liberalization to reduce the risk that a fast growth spell ends (Berg et al., 2012). Financial liberalization can lead to both increasing capital inflows and financial deepening but also enable capital flight and generally volatile capital flows. The sign of its effect is not clear ex ante. On the contrary, we expect trade openness to unequivocally reduce the probability of stagnation.

Last, we include a measure for income inequality after taxes and transfers ($Inequality_{(t-1)}$). Net income inequality is not only an economic variable but just as much influenced by a country's political institutions. Most of the growth episodes literature does not systematically analyze the role of inequality, with the exception of an early study by Rodrik (1999) and recent evidence of a negative effect on the length of positive growth spells (Berg et al., 2012). The effect of inequality on average growth rates remains disputed in panel studies.⁷ However, parts of neo-institutional theory (Engerman and Sokoloff, 1997) and earlier work on the interaction of inequality and growth collapses (Rodrik, 1999) suggest a negative sign for inequality, while earlier theories suggests that inequality rises alongside rapid development and falls again at higher income levels (Kuznets, 1955). Our data for net income inequality is taken from Solt (2009), who appends, benchmarks, and standardizes data from UNU-WIDER's *World Income Inequality Database* (WIID).

Institutional and 'shock' variables. This set of variables aims to capture some observable cross-country heterogeneity that can be attributed to institutions, as well as various shocks which require a response from policy makers and economic actors within the constraints of the political and institutional structure. These shocks may be external or internal but have in common that they pose a challenge to the prevailing regime and/or a country's institutional set-up.

Cross-sectional studies of GDP levels find strong support that institutions explain large parts of long-run growth (Knack and Keefer, 1995; Hall and Jones, 1999; Acemoglu et al., 2001, 2002) and also provide evidence that growth-enhancing institutions (e.g. property rights or executive constraints) contribute to lower growth volatility (Acemoglu et al., 2003). We expect that more open and democratically constrained institutions will reduce the probability of experiencing a stagnation spell. Our measure of political institutions is the revised combined polity score ($Polity2_{(t-1)}$) from the Polity IV project (Marshall and Jaggers, 2010). This measure is the difference between a country's score on the aggregate *institutionalized democracy* index and the score on the *institutionalized autocracy* index coded by the Polity IV project. It has a range from -10 (hereditary monarchy) to $+10$ (consolidated democracy). For studies requiring time-series, the Polity IV data is uniquely suitable as it provides annual data, starting as early as 1800. The Polity data mainly measures the limits on executive authority and the degree of political participation. Although imperfect, we consider this data most closely aligned with our notion of political institutions.⁸ Other regime type indicators focus purely on elections,

⁷See Perotti (1996), Deininger and Squire (1998), Barro (2000), and Forbes (2000).

⁸Polity IV scores countries on five indices capturing the openness of the political process and the constraints placed on individual actors. We interpret the data as measuring the degree of open institutions with narrow mandates.

political freedoms, or the power of the elite, which makes them more suitable in other contexts (also see Cheibub et al., 2010, for a discussion).

From Polity IV, we also derive two additional measures of political shocks. We code a dummy for positive regime changes (*Regchange* $+(t-1)$) as major positive changes of the political structure identified by at least a three-point improvement in the polity score. Conversely, we code negative regime changes (*Regchange* $-(t-1)$) as a negative change of at least three points in the polity score, including interregna and state failure. We expect negative regime changes to increase the probability of stagnation, while positive regime changes may have a stagnation deterring effect. Yet, regime transitions may also be a sign of political instability or consolidation of power.

We also include a dummy for the irregular exit of leaders (*Leader Exit* $_{(t-1)}$) based on *Archigos 2.9* (Goemans, Gleditsch, and Chiozza, 2009) as a proxy for internal shocks to a country's political regime. This variable codes an irregular exit whenever a country's major leader (president, chancellor, dictator, and so forth) lost power by violating established rules and conventions. Such cases include, but are not limited to, the loss of power due to the removal by foreign intervention, assassinations, ill health, and domestic popular protest with foreign support. We focus on leader exit and not entry, as our aim is to link periods of stagnation to unexpected adverse events and not to their possible resolution. Some studies of growth accelerations have accounted for the sudden death of leaders in office but usually not other types of exit (Hausmann et al., 2005; Jones and Olken, 2008).

In order to investigate the impact of large scale violence on stagnation spells, we include a dummy for the occurrence of *War/Conflict* $_{(t-1)}$ based on the UCDP/PRIO *Armed Conflict Dataset v.4-2011* (Gleditsch et al., 2002). We expect countries that are the location of an interstate war or large civil strife to be especially prone to falling into a stagnation spell. Our measure codes a war if UCDP/PRIO records a conflict intensity of two or higher, corresponding to at least 1,000 battle-related deaths in a country-year, and if the country is recorded as a location of war. In the case of multiple conflicts, we select the conflict with the highest intensity.

5 Empirical strategy

Most extant studies of growth episodes use pooled probit or non-linear panel methods to study the *onset* of specific growth episodes, such as accelerations or collapses. We know of only two papers concerned with modeling the duration of growth spells using observations within the episode (Hausmann et al., 2008; Berg et al., 2012). Other studies exploit only part of the data, either retaining only the first observation of the episodes of interest, or comparing averages of covariates before and after a regime change (e.g. Jones and Olken, 2008). However, we are unaware of any theoretical justification for assuming that determinants of stagnation explain only its onset and not its continuation. Apart from the obvious loss of efficiency, the neglect of information from within episodes also brings with it the risk of rare event bias (King and Zeng, 2001).

In fact, state dependence is likely be a crucial feature of growth episodes in general and of stagnation spells in particular. Yet, apart from Jerzmanowski's (2006) Markov-switching models, the literature does not model the incidence of a growth spell as a dynamic process. The likely explanation is that, in dynamic non-linear models, special efforts are required to address econometric issues such as *spurious* state dependence,

endogeneity of the lagged state, unobserved heterogeneity, the initial conditions problem, and non-linear interaction effects.

We propose to phase the issues and first introduce dynamics in a generic linear model of stagnation episodes. This generic model will be used for an initial screening and selection of potential explanatory variables. The linear estimates will also be used to provide initial variance estimates for more complex nonlinear models. Ultimately, linear estimates remain interesting as they are much less demanding in terms of statistical assumptions than the potentially more efficient nonlinear models and provide a robust benchmark.

For ease of exposition, we write the general model in index form, leaving the functional form unspecified.

$$y_{it} = \mathbf{1}[\alpha y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}'_{it}y_{i,t-1}\boldsymbol{\gamma} + \mu_i + \nu_{it} > 0], \quad i = 1, \dots, N, \quad t = 2, \dots, T_i \quad (1)$$

where y_{it} indicates whether or not a country i is stagnating in year t , $y_{i,t-1}$ is the lagged state, \mathbf{x}_{it} is a vector of covariates, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are parameters to be estimated, μ_i is a time-invariant unobserved country effect, and ν_{it} a residual time-varying error. The interaction term ($\mathbf{x}'_{it}y_{i,t-1}\boldsymbol{\gamma}$) allows the impact of the covariates to be different in (or just after) a stagnation spell than in expansions. For now, the unobserved effects may be of the ‘fixed’ or ‘random’ variety: nothing is assumed about their distribution or the absence of correlation with the explanatory variables. In most specifications, we will also include quinquennial dummies among the regressors though not their interactions with $y_{i,t-1}$.

In the linear case, the elements of the parameter vector $\boldsymbol{\gamma}$ can be interpreted as slope shifts in the effects of the variables in \mathbf{x}_{it} if a country is or has just been in a stagnation spell (i.e. $y_{i,t-1} = 1$), while mean shifts are captured by α . This model allows us to test the hypothesis that specific elements or subsets of $\boldsymbol{\gamma}$ are equal to zero. Macroeconomics and political economy theory provide little guidance as to whether and how some effects should differ in the initial and later stages of a stagnation episode. Our modeling strategy is to “let the data decide” which variables in \mathbf{x}_{it} require an interaction term with $y_{i,t-1}$ and which do not. We proceed in four steps. First, we estimate fully interacted linear probability models specified according to eq. (1). Second, we test whether the interaction terms that are individually insignificant at the 10% level may also be considered jointly insignificant. Third, based on these tests, we specify a ‘parsimonious’ reference model retaining only those interactions that pass our inclusion criteria. Fourth, we compare the preceding results with those found using non-linear probability models.

Linear probability models

Simplest of all is to estimate eq. (1) as a linear probability model (LPM) with country fixed effects (FE). This approach is particularly attractive, as FE-OLS requires no distributional assumptions about the unobserved effects and the OLS coefficients are usually a good approximation of the partial effects near the means of the variables. However, the variance of the dependent variable is known to be of the form $\mathbf{x}'_{it}\boldsymbol{\beta}(1 - \mathbf{x}'_{it}\boldsymbol{\beta})$, making the LPM inherently heteroskedastic. Furthermore, the LPM can predict probabilities outside the unit interval and hence non-positive variances, unless the predictions are arbitrarily trimmed.

In model (1), the LPM has two further shortcomings. First, it implies awkward restrictions on the unobserved effects. Second, since the same unobserved effects (μ_i) also

appear in the lagged state $y_{i,t-1}$, their presence means that the ‘Within’ OLS estimate of α is downward biased (Nickell, 1981). The ‘Nickell bias’ is decreasing in T . Our panel has an average length of about 18 years; therefore the bias should be moderate.

For comparison, we also estimate eq. (1) using differenced and system Generalized Methods of Moments (GMM) estimators. Differenced GMM uses lagged levels of order two and higher as instruments for the endogenous regressors in a differenced equation (Holtz-Eakin et al., 1988; Arellano and Bond, 1991). To alleviate problems of weak instrumentation, Arellano and Bover (1995) and Blundell and Bond (1998) proposed the ‘system GMM’ estimator that instruments levels with lagged differences and estimates the equation both in differences and in levels simultaneously. Under the appropriate conditions system GMM is consistent and more efficient than differenced GMM. The extra moment conditions require that the lagged differences are not correlated with the unobserved effects. In addition, system GMM requires that the initial conditions (y_{i1}) represent a stationary equilibrium, which is arguably an unnatural assumption for the analysis of stagnation spells. We apply two-step system GMM with a small sample correction due to Windmeijer (2005).

System GMM is not a universal panacea. Bun and Windmeijer (2010) recently showed that the level equation in system GMM also suffers from a weak instruments problem. Differenced or system GMM estimates are also often unstable and strongly depend on the instrument matrix used (Roodman, 2009). For this reason, we do not use GMM for model selection but only apply it to the ‘parsimonious’ specification to check whether the results remain within a reasonable range of the FE-OLS estimates.

Non-linear probability models

Simple within or first-difference transformations cannot eliminate unobserved heterogeneity in non-linear probability models like logit and probit models. Instead, the assumptions made on the structure of the unobserved effects determine which type of model can be estimated. We apply two techniques: fixed effects logit and dynamic random-effects probit. On the one hand, the fixed effects logit estimator is less restrictive in its assumptions about the unobserved heterogeneity but similarly to the LPM with FE, it does not deal with the endogeneity of the lagged state. On the other hand, the dynamic random effects probit estimator requires explicit assumptions about the unobserved heterogeneity, but has been modified to account for the endogeneity of the lagged state, including solutions for the initial conditions problem.

Fixed effects logit. The dummy variables fixed effects logit model estimated by unconditional Maximum Likelihood (ML) runs into a statistical problem. Even in a simplified version of eq. (1) without the lagged state, we need a consistent estimate over $t = 1, \dots, T_i$ for each of the unobserved effects μ_i . Any inconsistency introduced there will contaminate the estimate of β . This is the well-known *incidental parameters problem* which, for balanced panels, creates a bias in the ML estimator of β in the order of $1/T$ (Neyman and Scott, 1948). Given the lengths of our series, we do not expect the bias to be large.

Chamberlain (1980) observed that there is a computational trick that allows consistent estimation of the parameter vector but not the constants by conditioning on the sum of observed outcomes *within* groups ($\sum_{t=1}^{T_i} y_{it}$). In the conditional logit model the incidental parameters (μ_i) drop out. However, conditioning on the sum of the observed outcomes

comes at a cost. Since groups in which y_{it} does not change over T_i provide no information for the likelihood, they drop out of the log-likelihood. If there is strong persistence the number of observations used in the estimation may fall a lot. Likewise, time-invariant effects cancel out of the estimation equation. They cannot be estimated, neither can partial effects since these depend on the expected value of the unobserved effects.

Given the expectation that the unconditional ML estimator is not too strongly biased and allows estimation of partial effects, we estimate both models and compare their results.

Dynamic random effects probit. Conditioning on the sum of the observed outcomes does not work for the equivalent probit model. Furthermore, the standard random effects probit model assumes that the unobserved heterogeneity is strictly unrelated to the explanatory variables. The presence of the lagged state ($y_{i,t-1}$) together with μ_i violates this assumption even if α is actually zero (Wooldridge, 2010, p. 626). This is the problem of *true* versus *spurious* state dependence. The estimated effect of $y_{i,t-1}$ depends on three sources: (1) serial correlation in the errors, (2) correlation with the unobserved effects, and (3) true state dependence (Greene, 2011, p. 769). In such a setting, the ordinary fixed or random effects estimation techniques do not provide consistent estimates. In addition, the outcome path may be severely influenced by the initial conditions (y_{i1}), which enter the unconditional likelihood function and prohibit integrating out the unobserved effects (μ_i). We are faced with two related problems: the violation of strict exogeneity in the presence of unobserved heterogeneity, and the treatment of the initial condition in the log-likelihood.

A relaxation of the strict exogeneity assumption known as *correlated random effects* has been developed for static models following the ideas of Mundlak (1978) and Chamberlain (1984). The approach allows for correlation between \mathbf{x}_{it} and μ_i but restricts the unobserved effects to depend on means (or other values) of the explanatory variables according to $\mu_i = \eta_0 + \bar{\mathbf{x}}_i' \boldsymbol{\eta}_2 + \varepsilon_i$, where the ε_i are assumed to be i.i.d. and normally distributed. The vector $\bar{\mathbf{x}}_i$ consists of time-averages (or other values, like initial levels) of the regressors \mathbf{x}_{it} .⁹ While this approach relaxes the strict exogeneity assumption, it does not address the issue of dynamics and the related initial conditions problem.

Several solutions to the initial conditions problem have been proposed (see Heckman, 1981; Orme, 2001; Wooldridge, 2005). Wooldridge (2005) suggests to condition the density (y_{i1}, \dots, y_{iT_i}) on the observed history of the covariates *and* the initial values by specifying a distribution of the unobserved effects given the initial conditions. Rabe-Hesketh and Skrondal (2013) show that this simple method performs well in general and Akay (2012) provides Monte Carlo evidence that a constrained model (with time-averages) also works well in unbalanced panels as long the time series lengths are moderately large.

Concretely, Wooldridge (2005) proposes to condition on the entire time series of the strictly exogenous variables less the initial period plus the initial condition (y_{i1}) – an approach that easily extends to interactions. A convenient way to specify the conditional distribution of the unobserved effects is $\mu_i | y_{i1}, \bar{\mathbf{x}}_i \sim \mathcal{N}(\eta_0 + \eta_1 y_{i1} + \bar{\mathbf{x}}_i' \boldsymbol{\eta}_2 + \bar{\mathbf{x}}_i' y_{i1} \boldsymbol{\eta}_3, \sigma_\varepsilon^2)$, where $\bar{\mathbf{x}}_i$ are the time-averages as in Mundlak-Chamberlain. This implies the following parametric specification for the unobserved effects: $\mu_i = \eta_0 + \eta_1 y_{i1} + \bar{\mathbf{x}}_i' \boldsymbol{\eta}_2 + \bar{\mathbf{x}}_i' y_{i1} \boldsymbol{\eta}_3 + \varepsilon_i$, where the ε_i are normal, i.i.d., and independent of ($y_{i1}, \bar{\mathbf{x}}_i$).

⁹Further, this model implies $\text{Corr}(\varepsilon_i + \nu_{it}, \varepsilon_i + \nu_{is}) = \sigma_\varepsilon^2 / (\sigma_\varepsilon^2 + \sigma_\nu^2)$ for any $t \neq s$.

As a final model, we specify the constrained probit version of eq. (1):

$$P(y_{it} = 1 | \mathbf{x}_{i2}, \dots, \mathbf{x}_{iT_i}, y_{i,t-1}, \dots, y_{i1}, \varepsilon_i) = \Phi(\alpha y_{i,t-1} + \mathbf{x}'_{it} \boldsymbol{\beta} + \mathbf{x}'_{it} y_{i,t-1} \boldsymbol{\gamma} + \eta_0 + \eta_1 y_{i1} + \bar{\mathbf{x}}'_i \boldsymbol{\eta}_2 + \bar{\mathbf{x}}'_i y_{i1} \boldsymbol{\eta}_3 + \varepsilon_i) \quad (2)$$

where \mathbf{x}_{it} is the vector of explanatory variables, $y_{i,t-1}$ is the lagged state, $\bar{\mathbf{x}}_i$ is the vector of time-averages of the covariates, the vector $\boldsymbol{\gamma}$ allows for differential effects of the covariates depending on the state ($\mathbf{x}_{it} y_{i,t-1}$), η_1 measures the effect of the initial condition (y_{i1}), and the vector $\boldsymbol{\eta}_3$ measures the effect of the averaged covariates depending on the initial condition ($\bar{\mathbf{x}}_i y_{i1}$). Estimation still proceeds over $i = 1, \dots, N$ and $t = 2, \dots, T_i$.

Appendix B outlines why the average partial effects (APEs), which are not identified in the conditional logit model, are in fact identified in this random effects probit model. This model addresses all three fundamental issues involved in our research problem. It specifies a data-coherent functional form, relaxes the strict exogeneity assumption, and consistently estimates the APE of time-varying variables and the lagged state variable in the presence of unobserved effects.

Partial effects of interactions. It is well known that the partial effects in non-linear models are not constant and that the model coefficients only indicate the direction and approximate significance of the effects. A somewhat less well known fact is that non-linearity necessarily implies that the coefficients of the interaction terms do not represent the sign, size or significance of the underlying interaction effect (e.g. see Ai and Norton, 2003; Greene, 2010).

Let $F(\mathbf{w}_{it}; \boldsymbol{\theta}) = F(\alpha y_{i,t-1} + \mathbf{x}'_{it} \boldsymbol{\beta} + \mathbf{x}'_{it} y_{i,t-1} \boldsymbol{\gamma} + \mu_i)$ denote a generic non-linear version of our interaction model with an implicit error, and let hats indicate estimated values. The change in the partial effect of $x_{k,it} \in \mathbf{x}_{it}$ due to a regime switch into stagnation ($\Delta y_{i,t-1} = 1$) is

$$PE(\hat{\gamma}_k)_{it} = \Delta F'_k(\mathbf{w}_{it}; \hat{\boldsymbol{\theta}}) = F'_k(\mathbf{w}_{it}; \hat{\boldsymbol{\theta}} | y_{i,t-1} = 1) - F'_k(\mathbf{w}_{it}; \hat{\boldsymbol{\theta}} | y_{i,t-1} = 0) \quad (3)$$

where $F'_k(\mathbf{w}_{it}; \hat{\boldsymbol{\theta}})$ can denote either a partial derivative with respect to $x_{k,it}$ or a difference, depending on whether $x_{k,it}$ is discrete or continuous, and in either case we write $F'_k(\mathbf{w}_{it}; \hat{\boldsymbol{\theta}}) = \partial F(\mathbf{w}_{it}; \hat{\boldsymbol{\theta}}) / \partial x_{k,it}$.

Clearly, $\hat{\gamma}_k$ is not equal to the interaction effect as it would be in a linear model. Now it is straightforward to also compute the APE for either $\Delta y_{i,t-1} = 1$ or $\Delta y_{i,t-1} = -1$ as

$$APE(\hat{\gamma}_k) = \frac{1}{N(\bar{T} - 1)} \sum_{i=1}^N \sum_{t=2}^{T_i} \frac{\Delta F'(\mathbf{w}_{it}; \hat{\boldsymbol{\theta}})}{\Delta y_{i,t-1}} \quad (4)$$

where \bar{T} is the average panel size.

We compute the standard errors of the APEs using the delta method, so that the asymptotic variance of the APE of an interaction with the lagged state is

$$\widehat{AVar}[APE(\hat{\gamma}_k)] = \frac{\partial APE(\hat{\gamma}_k)}{\partial \hat{\boldsymbol{\theta}'}} \hat{\boldsymbol{\Omega}} \frac{\partial APE(\hat{\gamma}_k)}{\partial \hat{\boldsymbol{\theta}}} \quad (5)$$

and $\hat{\boldsymbol{\Omega}}$ is the ML estimate of the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}$. For the observation-specific partial effects the Jacobian vectors are not averaged.

6 Results and discussion

Linear models. Table 3 reports the results from variations of the LPM. Since we primarily use the linear models for variable selection, we only briefly discuss the results and defer the economic interpretation until the description of the preferred specification in the next subsection. Column (1) is the standard fixed effects model with standard errors clustered at the country level. Column (2) is similar but allows for arbitrary heteroskedasticity and autocorrelation within both country and time clusters (Cameron, Gelbach, and Miller, 2011). Column (3) shows the FGLS estimator with clustering at the country level via the weights. Column (4) is the parsimonious FGLS specification discarding those interaction terms that are individually and jointly insignificant. Columns (5) and (6) re-estimate (4) using system GMM.

Two joint hypothesis tests reported in Table 3 are key to our model-building approach. First, we test if the coefficients of the interactions with the lagged state and the coefficient of the lagged state are jointly zero; taken together these variables comprise Set I. In all models this hypothesis is rejected, indicating the presence of dynamics. Second, we test whether the interactions with the lagged state that are individually insignificant at the 10% level (i.e., Set II) can also be jointly omitted. In columns (1) to (4), the insignificant interactions pass this joint-exclusion test, but not in the GMM estimates.

All fully-interacted specifications provide very similar results. Inflation, financial openness, trade openness, and negative regime changes clearly affect the probability of entering a stagnation spell. Considering the interactions, the coefficients of inflation, trade openness, and negative regime changes are significant in most but not all specifications. Interestingly, the interaction terms often have the *opposite* sign of the non-interacted coefficients, indicating that the respective effects are weaker within a stagnation spell. Column (4) only retains the significant interactions and represents our parsimonious specification which we later re-estimate with non-linear techniques. We still find evidence that inflation, trade openness, and negative regime changes have a different impact within the spell than on the onset probabilities. However, the interaction effect of financial openness is, again, insignificant.

The system GMM specifications in columns (5) and (6) assess the stability of the parameter estimates once we account for the endogeneity of the lagged state and the interactions with the lagged state. While the significance of individual coefficients changes between the two GMM models, the sizes and signs of the previously highlighted coefficients remain broadly similar. In column (5) we use the second and third lag of the predetermined variables as instruments. Nevertheless, the *J*-test indicates that the number of instruments is too large relative to the group size. We address this concern in column (6) by collapsing the instrument set (see Roodman, 2009). There is no evidence of instrument endogeneity and the estimated parameters come closer to the earlier least-squares estimates. However, applying GMM in our context is not ideal. A moderate \bar{T} quickly leads to instrument proliferation and problems in identifying an instrument set that balances efficiency gains with decreasing relevance. Accordingly, we place less emphasis on the GMM results and rely more on verifying the results from the preceding linear models with non-linear techniques.

The linear models point to several preliminary conclusions. First, we find that inflation, financial openness, trade openness, and negative regime changes have a statistically significant effect on the probability of stagnation, and so do to a lesser extent exchange rate undervaluation and inequality. Second, state dependence plays a large

Table 3 – Linear Models – Probability of Stagnation

Independent Variable	(1) LPM	S.E.	(2) LPM	S.E.	(3) FGLS	S.E.	(4) FGLS	S.E.	(5) SYS-GMM	S.E.	(6) SYS-GMM	S.E.
Log GDP _C (<i>t</i> -1)	0.052	0.073	0.052	0.094	0.061	0.073	0.040	0.065	-0.018	0.035	-0.029	0.020
Inflation(<i>t</i> -1)	0.005*	0.003	0.005	0.004	0.005**	0.002	0.005**	0.002	0.003	0.004	0.013	0.010
Δ ToT(<i>t</i> -1)	-0.080	0.081	-0.080	0.082	-0.092	0.109	-0.036	0.054	-0.035	0.043	-0.056	0.049
Δ Real Exports(<i>t</i> -1)	-0.136	0.169	-0.136	0.177	-0.080	0.143	-0.014	0.047	0.011	0.067	0.055	0.062
REFR Value(<i>t</i> -1)	-0.150**	0.088	-0.150	0.113	-0.083	0.075	-0.016	0.039	-0.146***	0.046	-0.155***	0.042
Fin. Openness(<i>t</i> -1)	-0.043***	0.016	-0.043***	0.020	-0.042***	0.017	-0.034**	0.016	-0.034**	0.023	-0.020	0.026
Trade Openness(<i>t</i> -1)	-0.236***	0.081	-0.236***	0.073	-0.233***	0.062	-0.231***	0.058	-0.185	0.148	-0.286**	0.125
Inequality(<i>t</i> -1)	-0.007	0.004	-0.007	0.005	-0.003	0.004	-0.002	0.003	0.007**	0.003	0.005**	0.002
Polity2(<i>t</i> -1)	-0.011	0.007	-0.011	0.007	-0.008*	0.005	-0.003	0.003	-0.003	0.004	-0.004	0.004
Regchange +(<i>t</i> -1)	0.033	0.068	0.033	0.077	0.006	0.103	0.017	0.032	0.036	0.036	0.031	0.043
Regchange -(<i>t</i> -1)	0.338*	0.187	0.338	0.202	0.277*	0.149	0.274*	0.143	0.450	1.050	0.348*	0.210
Leader Exit(<i>t</i> -1)	-0.071	0.053	-0.071	0.055	-0.045	0.054	0.004	0.028	-0.037	0.043	-0.033	0.043
War/Conflict(<i>t</i> -1)	0.083	0.065	0.083	0.091	0.097	0.087	-0.005	0.042	-0.022	0.068	0.016	0.080
Log GDP _C (<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.002	0.048	0.002	0.050	-0.039	0.042						
Inflation(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	-0.004*	0.003	-0.004	0.004	-0.005**	0.002	-0.005**	0.002	-0.002	0.004	-0.012	0.010
Δ ToT(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.031	0.093	0.031	0.084	0.067	0.125						
Δ Real Exports(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.117	0.166	0.117	0.169	0.073	0.151						
REFR Value(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.146	0.103	0.146	0.105	0.081	0.075						
Fin. Openness(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.021	0.019	0.021	0.019	0.034*	0.018	0.023	0.016	0.064**	0.029	0.038	0.039
Trade Openness(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.186**	0.084	0.186**	0.088	0.207***	0.066	0.208***	0.059	0.132	0.169	0.292*	0.148
Inequality(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.001	0.003	0.001	0.004	0.001	0.003						
Polity2(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.007	0.007	0.007	0.007	0.007	0.005						
Regchange +(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	-0.017	0.068	-0.017	0.076	0.011	0.108	-0.277*	0.160	-0.399	1.040	-0.363*	0.211
Regchange -(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	-0.354*	0.184	-0.354*	0.193	-0.280*	0.166						
Leader Exit(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	0.101*	0.060	0.101	0.068	0.059	0.060						
War/Conflict(<i>t</i> -1) × <i>y</i> (<i>t</i> -1)	-0.171**	0.081	-0.171*	0.090	-0.126	0.096						
<i>y</i> (<i>t</i> -1)	0.158	0.473	0.158	0.505	0.486	0.423	0.222***	0.059	0.363**	0.160	0.317**	0.141
Constant	0.758	0.746	0.373	0.846	0.463	0.647	0.617	0.560	0.368	0.392	0.508*	0.261
Observations	1691		1691		1691		1691		1691		1691	
Country FE	Yes		Yes		Yes		Yes		-		-	
Clusters	country		country+year		country		country		country		country	
Joint: Country FE [p-values]	0.000		0.000		0.000		0.000		0.053		0.024	
Joint: 5-Year FE [p-values]	0.101		0.451		0.430		0.000		0.000		0.000	
Joint: Set I [p-values]	0.000		0.000		0.708		0.167		0.759		0.038	
Joint: Set II [p-values]	0.796		0.226						390		35	
Instruments												
AR1 [p-value]												
AR2 [p-value]												
Hansen's <i>J</i> [p-value]												

Notes: ***, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$.

role in determining whether a country experiences a stagnation episode or not. Third, in all specifications there is considerable evidence of unobserved heterogeneity at the country level. Fourth, all significant interaction terms with the lagged state point into the direction opposite of their linear counterparts.

Non-linear models. Turning to the non-linear models, we first discuss the heterogeneity of the interaction effects using a logit specification and then present the results from the dynamic random effects probit estimator (our preferred specification).

Table 4 shows the estimation results from two types of fixed effects logit estimators. As a reference, column (1) shows the fully interacted model estimated using conditional maximum likelihood, where the fixed effects are not estimated but drop out. Column (2) is the conditional logit equivalent of the parsimonious linear model, and column (3a) is the same model using dummy variables for the country fixed effects. The last column (3b) reports the APEs based on column (3a). The APEs of interaction terms are reported similarly to coefficients of a linear model; that is, if $y_{i,t-1} = 0$, the APE is reported in the row of the non-interacted variable and, for $y_{i,t-1} = 1$, the total APE is the sum of the former and the APE of the interaction term.

A complication in interpreting the results of the conditional logit models is that the APEs are not identified, for the simple reason that the individual country effects are not estimated. In order to approximate the APEs, we estimate the equivalent dummy variables fixed effects logit model and compute the APEs based on its results. Columns (2) and (3a) show that the parameter estimates and the corresponding standard errors remain very similar, justifying this approach.

We still find evidence that inflation, financial openness, trade liberalization, and negative regime changes significantly affect the probability of stagnation. However, for all but negative regime changes, the APEs of the interaction terms are *statistically insignificant* and, in the case of financial openness, the interaction effect no longer points in the opposite direction. Applying fixed effects logit also substantially reduces the estimation sample to 1314 observations in 62 countries, for lack of within-group variation. This loss of observations makes it impossible to identify the effect of the negative regime change interaction term; as a consequence, the interaction is dropped from the earlier parsimonious model.

The APEs are close to the linear approximation in terms of size. Notable changes are that now the effects of changes in the terms of trade, real exchange rate (RER) undervaluation, and inequality are all significant at the 10% level and have increased substantially in absolute size. The degree of state dependence identified by the logit model is also somewhat higher than in the linear approximation (a 32.1 percentage points higher probability of continuation if a country was stagnating the year before).

Table 4 may seem to provide little evidence that some variables have a different impact on the onset of a stagnation spell than on its continuation. However, in non-linear models, the partial effects of the interaction can be very heterogeneous at the observation level and may be individually significant for a substantial subset of observations. By definition, APEs ‘average out’ this type of heterogeneity which can be especially troublesome if the partial effects switch signs. In Figure 2 we examine these non-linearities more closely. The figure graphs the distributions of partial effects for the interaction terms of inflation, financial openness, and trade openness over the predicted probability of stagnation, as well as the associated p -values of a Wald-test of the null that the interaction effect at each particular observation is zero.

Table 4 – Logit Models – Probability of Stagnation

Independent Variable	(1)		(2)		(3a)		(3b)	
	Logit CML	S.E.	Logit CML	S.E.	Logit ML	S.E.	APEs	S.E.
Log GDP $C_{(t-1)}$	0.417	0.707	0.430	0.695	0.430	0.738	0.058	0.099
Inflation $_{(t-1)}$	0.037	0.025	0.043	0.026	0.046	0.028	0.006*	0.004
Δ ToT $_{(t-1)}$	-0.556	0.571	-0.828*	0.479	-0.885*	0.516	-0.119*	0.070
Δ Real Exports $_{(t-1)}$	-0.784	1.163	-0.621	0.735	-0.681	0.795	-0.091	0.107
RER Value $_{(t-1)}$	-1.375**	0.567	-0.957*	0.503	-1.020*	0.533	-0.137*	0.071
Fin. Openness $_{(t-1)}$	-0.453***	0.121	-0.396***	0.110	-0.421***	0.118	-0.058***	0.016
Trade Openness $_{(t-1)}$	-1.441***	0.486	-1.424***	0.456	-1.527***	0.489	-0.259***	0.091
Inequality $_{(t-1)}$	-0.050	0.031	-0.052*	0.029	-0.054*	0.030	-0.007*	0.004
Polity2 $_{(t-1)}$	-0.066	0.044	-0.051	0.033	-0.054	0.036	-0.007	0.005
Regchange $+_{(t-1)}$	0.193	0.561	0.129	0.376	0.136	0.397	0.018	0.054
Regchange $-_{(t-1)}$	2.102**	0.900	1.945**	0.938	2.038**	0.991	0.288**	0.135
Leader Exit $_{(t-1)}$	-0.501	0.426	-0.060	0.317	-0.058	0.334	-0.007	0.044
War/Conflict $_{(t-1)}$	0.115	0.382	-0.428	0.342	-0.454	0.357	-0.059	0.045
Log GDP $C_{(t-1)} \times y_{(t-1)}$	0.082	0.284						
Inflation $_{(t-1)} \times y_{(t-1)}$	-0.012	0.012	-0.019	0.015	-0.021	0.016	-0.002	0.002
Δ ToT $_{(t-1)} \times y_{(t-1)}$	-1.128	1.025						
Δ Real Exports $_{(t-1)} \times y_{(t-1)}$	0.290	1.244						
RER Value $_{(t-1)} \times y_{(t-1)}$	0.859	0.624						
Fin. Openness $_{(t-1)} \times y_{(t-1)}$	0.144	0.117	0.045	0.107	0.045	0.116	-0.011	0.022
Trade Openness $_{(t-1)} \times y_{(t-1)}$	0.947*	0.528	0.885*	0.487	0.939*	0.521	0.153	0.094
Inequality $_{(t-1)} \times y_{(t-1)}$	0.003	0.019						
Polity2 $_{(t-1)} \times y_{(t-1)}$	0.013	0.042						
Regchange $+_{(t-1)} \times y_{(t-1)}$	-0.164	0.620						
Leader Exit $_{(t-1)} \times y_{(t-1)}$	0.833*	0.506						
War/Conflict $_{(t-1)} \times y_{(t-1)}$	-1.051*	0.537						
$y_{(t-1)}$	-0.004	2.813	1.189***	0.456	1.282***	0.487	0.321***	0.029
Observations	1314		1314		1314			
Country FE	Yes		Yes		Yes			
5 Year FE	Yes		Yes		Yes			
Clustered Errors [Country]	Yes		Yes		Yes			
Countries	62		62		62			
Log-pseudolikelihood	-465.864		-469.311		-555.547			

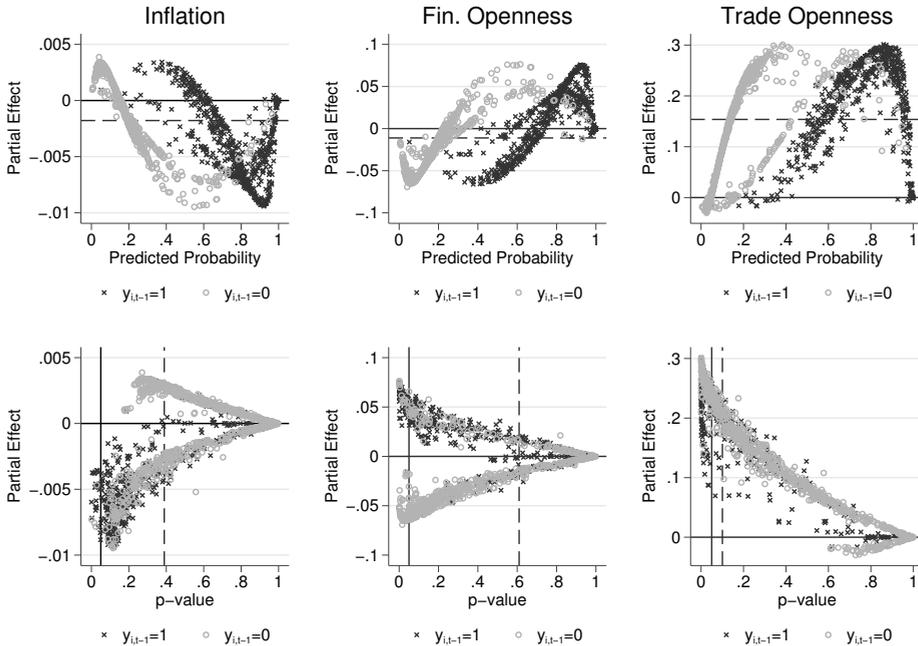
Notes: In column (1) Regchange $-_{(t-1)} \times y_{(t-1)}$ had been dropped due to a lack of within-group variance. The asymptotic standard errors of the APEs are computed via the delta method.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Clearly, the partial effects of all three interaction terms are extremely non-linear and tend to include both positive and negative values. In the case of inflation, we find that for some observations the partial effect is positive, while for most observations it is negative and insignificant. The effect is significant only at negative values for a very small fraction of the distribution (30 observations) and ranges from -0.95% to 0.39% , which is moderately large compared to an APE of 0.6% when $y_{i,t-1} = 0$. The predicted partial effects group into two families of curves with an S-shape. The curves at lower stagnation probabilities are the partial effects for observations where $y_{i,t-1} = 0$ (symbol: o) and the curves going across higher stagnation probabilities are the predicted effects for observations where $y_{i,t-1} = 1$ (symbol: x). The dashed line refers to the APE in the upper panel and the p -value of a Wald test that the APE is zero in the lower panel.

The interaction effect for financial openness is also S-shaped but sloping upwards and ranges from -6.9% to 7.6% . The partial effect is statistically significant for 14% of the predicted outcomes at both negative and positive values. However, similarly to inflation, the large range of insignificant negative and positive values supports the conclusion that the effect is not different from zero on average. In contrast, the evidence of a significant interaction is relatively strong for trade openness. For most observations the partial effect

Figure 2 – Partial Effects of Interaction Terms in Logit Model



Notes: The dashed line refers to the APE in the upper panel and the p-value of a Wald test that the APE is zero in the lower panel. “Predicted Probability” is the probability of stagnation $P(y_{it} = 1)$.

of the interaction term is positive and very large, with an overall range from -3% to 30% . For about 32% of the observations in the sample the partial effect of the interaction term is individually significant at the 5% level.

Overall, the logit specifications confirm the findings of the linear models given a few refinements. Inflation, financial openness, trade openness, and negative regime changes remain significant predictors of being in stagnation. Furthermore, these estimates suggest that exchange rate undervaluation, changes in the terms of trade, and inequality affect the probability of stagnation. However, the evidence in favor of interaction effects with the lagged state for any other variable than trade openness is weak; the estimates suggest that only trade openness has a different impact on the onset of a stagnation spell than its continuation.

We now turn to the last and final set of estimates. Table 5 presents the random effects probit estimates of the dynamic panel model. For comparison purposes, we first report the naïve pooled probit version of the fully-interacted dynamic model and then its random effects counterpart estimated according to the Wooldridge approach – columns (1) and (2), respectively. Column (3a) is parsimonious version of the dynamic random effects model in column (2) and our preferred specification. Column (3b) reports the APEs of the variables of interest and their associated standard errors.

We report several additional rows in Table 5 to display the regression results more

Table 5 – Probit Models – Probability of Stagnation

Independent Variable	(1)		(2)		(3a)		(3b)	
	Probit	S.E.	Woold	S.E.	Woold	S.E.	APEs	S.E.
Log GDPC _(t-1)	0.027	0.103	0.449	0.377	0.398	0.367	0.083	0.076
Inflation _(t-1)	0.017***	0.006	0.012	0.008	0.020***	0.008	0.005***	0.002
Δ ToT _(t-1)	-0.328	0.300	-0.435	0.476	-0.321	0.349	-0.067	0.072
Δ Real Exports _(t-1)	-0.226	0.540	-0.570	0.565	-0.362	0.358	-0.075	0.074
RER Value _(t-1)	-0.719***	0.202	-0.786**	0.331	-0.513*	0.278	-0.106*	0.058
Fin. Openness _(t-1)	-0.056	0.039	-0.205***	0.074	-0.153**	0.073	-0.038**	0.018
Trade Openness _(t-1)	-0.567***	0.217	-0.926***	0.240	-0.774***	0.228	-0.218***	0.068
Inequality _(t-1)	0.014**	0.007	-0.026	0.016	-0.022	0.015	-0.004	0.003
Polity2 _(t-1)	-0.014	0.013	-0.031	0.020	-0.027*	0.016	-0.006*	0.003
Regchange + _(t-1)	0.039	0.269	0.138	0.405	0.066	0.220	0.014	0.046
Regchange - _(t-1)	1.074*	0.557	1.238**	0.617	1.047*	0.584	0.214*	0.112
Leader Exit _(t-1)	-0.196	0.188	-0.118	0.235	-0.006	0.171	-0.001	0.035
War/Conflict _(t-1)	0.110	0.257	0.280	0.374	-0.199	0.246	-0.041	0.051
Log GDPC _(t-1) \times $y_{(t-1)}$	-0.259	0.182	-0.129	0.154				
Inflation _(t-1) \times $y_{(t-1)}$	-0.008	0.006	-0.004	0.008	-0.012	0.008	-0.003	0.002
Δ ToT _(t-1) \times $y_{(t-1)}$	-0.091	0.393	-0.097	0.750				
Δ Real Exports _(t-1) \times $y_{(t-1)}$	-0.435	0.574	0.286	0.745				
RER Value _(t-1) \times $y_{(t-1)}$	0.635	0.390	0.532*	0.308				
Fin. Openness _(t-1) \times $y_{(t-1)}$	0.096	0.060	0.086	0.078	-0.019	0.068	-0.006	0.017
Trade Openness _(t-1) \times $y_{(t-1)}$	0.474	0.321	0.816***	0.262	0.646***	0.238	0.186***	0.067
Inequality _(t-1) \times $y_{(t-1)}$	0.005	0.012	0.008	0.011				
Polity2 _(t-1) \times $y_{(t-1)}$	0.008	0.022	0.005	0.020				
Regchange + _(t-1) \times $y_{(t-1)}$	0.166	0.284	-0.137	0.482				
Leader Exit _(t-1) \times $y_{(t-1)}$	0.290	0.259	0.201	0.290				
War/Conflict _(t-1) \times $y_{(t-1)}$	-0.267	0.404	-0.803*	0.458				
y_0			0.969	2.621	1.244***	0.477		
$y_{(t-1)}$	3.407*	1.787	1.554	1.569	1.020***	0.229	0.389***	0.035
Constant	-1.287	1.105	-0.733	2.482	-2.958	1.871		
Observations	1586		1471		1596			
η_2 -estimated?	NO		YES		YES			
η_3 -estimated?	NO		YES		YES			
5-Year FE	YES		YES		YES			
Countries	90		81		90			
Pseudo-R ²	0.423		0.269		0.255			
Log-likelihood	-632.218		-574.211		-596.272			

Notes: In column (1) Regchange -_(t-1) \times $y_{(t-1)}$ had been dropped due to a lack of within-group variance. The asymptotic standard errors of the APEs are computed via the delta method.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

succinctly. The row “ η_2 -estimated?” refers to whether the Mundlak-Chamberlain time-averages are included. As these are not interesting by themselves, we do not report their coefficients and partial effects. Similarly, the row “ η_3 -estimated?” reports whether the interaction terms are included twice, once as an interaction with the lagged state and once as an interaction with the initial condition. We also do not report the parameter estimates of the latter, as they primarily account for the non-random nature of the unobserved effects.

The random effects probit model gives different (and superior) estimates compared to the pooled probit version.¹⁰ Several parameters exhibit sign changes and substantial changes in significance levels. More interestingly, even though the Wooldridge estimator of the dynamic model includes several additional terms and requires the assumption that the regressors are only correlated with the unobserved heterogeneity through their

¹⁰The coefficients are scaled differently. The comparison should be done in terms of relative sizes, signs, and significance levels – where appropriate.

averages, the results are remarkably similar to those of the linear models (in Table 3) and logit specifications (in Table 4). We do not compute the partial effects at the observation-level as only the APEs across the entire distribution of the unobserved heterogeneity are identified (see Appendix B and Wooldridge, 2005).

The APE of inflation is similar to that in the previous models. According to the Wooldridge estimator, a unit increase in the inflation index in $t - 1$ translates into 0.5 percentage points higher probability of stagnation in year t , if the country was not in stagnation in $t - 1$. The APE of the interaction effect is -0.3% and insignificant at all conventional levels, supporting the view that inflation increases the chances of stagnation, no matter if this occurs within or outside of a stagnation spell. We find no evidence in favor of the hypothesis that (higher) inflation helps exiting a stagnation spell.

Next, the effect of financial openness is also similar to that estimated by the linear probability model. A unit increase in the Chinn-Ito index towards more openness reduces the probability of stagnation by 3.8 percentage points. The APE of the interaction term is small and insignificant (-0.6%), suggesting that there is no difference between onset and continuation probabilities. This result is particularly interesting, since capital account openness is sometimes restricted during crises to avert capital flight. Yet we find no evidence that it increases the risk of stagnation when a country is already stagnating, but instead it seems to be associated with a lower probability of crisis throughout.

In the case of trade openness, a discrete change towards openness when $y_{i,t-1} = 0$ has an average partial effect of -21.8% , which is similar to the previous results. The interaction term is highly significant and has a very large effect on the predicted probability of stagnation (18.6%). Adding these two effects, we get the APE for the probability of continued stagnation if $y_{i,t-1} = 1$ and the country is open to trade which is still negative but statistically insignificant (-3.2%). Trade openness has already been shown to affect growth rates (Wacziarg and Welch, 2008), growth accelerations (Hausmann et al., 2005), and the duration of positive growth spells (Berg et al., 2012). Our results add that open economies are significantly less likely to fall into a stagnation spell – it is the single largest effect of a non-political variable in all specifications – but also that openness does not significantly improve the chances of exiting a stagnation spell.

This finding suggests that, on top of typically being linked to higher average growth rates, trade openness is ‘good for growth’ in the sense that it protects countries from stagnation. While trade benefits growth through specialization, technological diffusion, and higher investment levels, it can also stabilize and accelerate growth but it does not necessarily lead to shorter crises (especially when these are caused by international factors). Arguably, the weaker within-stagnation effect could also be driven by the fact that countries often liberalize in response to an ongoing crisis which often yields no immediate benefits and can temporarily increase volatility (Wacziarg and Welch, 2008).

Regarding the political shocks, the confidence interval of the APE of negative regime changes widens a bit (the t -statistic is 1.90), but the absolute size of the effect remains large. If a country experiences a negative regime change, then the probability of stagnation increases by 21.4 percentage points. As in the logit models, we could not include the interaction term due to a lack of time variation. Whereas negative regime changes were previously linked to the onset of growth accelerations (Hausmann et al., 2005), our analysis reveals the more plausible result that they strongly predict stagnation episodes. None of the other political variables, that is positive regime changes, irregular leader exits or the Polity score, are robustly related with the incidence of stagnation spells. This finding stands in sharp contrast to the ‘institutions trump other factors’

perspective (e.g. Acemoglu et al., 2001, 2002; Rodrik et al., 2004) but could be driven by our focus on annual transitions in and out of crises (which are likely to be caused by more proximate factors). Aidt and Gassebner (2010), for example, find that the level of trade integration is endogenous to the political regime. Autocratic countries trade less than democracies, but steps towards improving the accountability of leaders can lead to higher trade integration in the medium-run.

Undervaluation of the real exchange rate has a moderate and marginally significant effect on the probability of stagnation. A one standard deviation change in the undervaluation index reduces the probability of both the onset and continuation of stagnation by about 4%. We cannot corroborate the results from the logit model that changes in the terms of trade and inequality have any meaningful effect on the probability of stagnation. In addition, wars, and civil conflict do not predict stagnation spells, and there is no evidence of an effect of the level of GDP per capita.

Table 5 also reveals that the degree of state dependence is underestimated by both the LPM and FE logit model. If a country experienced stagnation in the previous year, it is 38.9 percentage points more likely to be in stagnation in the current year. Path dependence is thus an important feature of the stagnation process. The models also fit the data reasonably well (Pseudo- $R^2 = 0.255$) and the fit compares favorably to the onset models used elsewhere (e.g. Hausmann et al., 2005, 2008) – although these goodness of fit criteria are not strictly comparable. Hence, we find that dynamic models a) allow us to test refined hypotheses about the difference between onset and continuation of stagnation spells, and b) lead to better predictions of when stagnation occurs.

To summarize, the main results are robust to different specifications, assumptions about the unobserved effects, and correlation between the observed variables and unobserved country-effects. Most of the preferred specifications identify the lags of inflation and negative regime changes as strong predictors of stagnation episodes. Several models also identify trade openness and financial openness as relevant. The non-linear models add three important insights. First, there is less evidence of interaction effects with the lagged state than the LPM suggests. The interactions are highly non-linear and vary significantly across the sample, but only in the case of trade openness we find convincing evidence that its effect differs depending on whether the country was in stagnation before or not. The evidence of interaction effects between the lagged state and inflation, financial openness, or negative regime changes is weak. Second, the non-linear models show that the real exchange rate matters for stagnation. An overvalued real exchange rate raises the chances of stagnation, while undervaluation reduces it. Third, the level of state dependence is still moderate, but larger than estimated by the linear models with fixed effects.

We also find that lagging all included regressors by one year to assure that they are at least contemporaneously exogenous results in identifying very different factors as significant than if the issue of endogeneity is left unaddressed. For example Hausmann et al. (2008), who studied the onset of stagnation, found that exports, inflation, wars, and political transitions matter. We do not find significant effects for exports and wars, although we do confirm their results for inflation and negative regime changes.

Interestingly, we cannot confirm two central results of the previous literature. First, we find little evidence that, apart from negative regime changes, institutions generally affect the incidence of stagnation. Consistent with the institutionalist literature, we find that poor countries are more prone to stagnation and crisis, but our analysis contradicts the hypotheses that weak institutions increase the incidence of stagnation. On the

contrary, traditional macroeconomic variables predict the incidence of stagnation rather well. Second and contrary to previous research (Rodrik, 1999; Hausmann et al., 2008), our data does not support the finding that changes in real exports or terms of trade affect the probability of stagnation.

7 Concluding remarks

This chapter provides a dynamic analysis of the incidence of stagnation and examines if stagnation spells are determined by institutional factors, various shocks, or macroeconomic factors. Building on a recent contribution by Hausmann et al. (2008), we define stagnation spells as episodes in which GDP per capita is below previously achieved levels. We then use fixed effects linear models, GMM, fixed effects logit, and dynamic random effects probit to assess the role of institutions, political shocks, and economic factors. In contrast to the previous literature, we explicitly examine the hypothesis that the effects of variables on the onset of stagnation and on the continuation of stagnation may be different.

We identify several factors that explain the incidence of stagnation spells. Adverse regime changes have the single largest effect on the incidence of stagnation, while higher inflation increases the chances of being in stagnation. More surprisingly, we find that real exchange rate undervaluation, financial openness, and trade openness help reduce the chances of stagnation. Additionally, we evidence that trade openness primarily protects against falling into stagnation, but does not speed up recovery. For all other variables there is no evidence of a differential impact on onset versus continuation.

Though traditional macroeconomic factors have the upper hand in our explanations of stagnation compared to institutional factors, we do not yet know to what extent macroeconomic factors are in turn influenced by underlying institutional characteristics. The use of more differentiated measures and time series for institutions as well as additional techniques for dealing with endogeneity might add further insights. This is a promising avenue for future research.

Bibliography

- Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5), 1369–1401.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics* 117(4), 1231–1294.
- Acemoglu, D., S. Johnson, J. A. Robinson, and Y. Thaicharoen (2003). Institutional causes, macroeconomic symptoms: volatility, crises and growth. *Journal of Monetary Economics* 50(1), 49–123.
- Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters* 80(1), 123–129.
- Aidt, T. S. and M. Gassebner (2010). Do autocratic states trade less? *The World Bank Economic Review* 24(1), 38–76.
- Akay, A. (2012). Finite-sample comparison of alternative methods for estimating dynamic panel data models. *Journal of Applied Econometrics* 27(7), 1189–1204.
- Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58(2), 277–297.
- Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* 68(1), 29–51.
- Aron, J. (2000). Growth and institutions: A review of the evidence. *World Bank Research Observer* 15(1), 99–135.
- Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth* 5(1), 5–32.
- Berg, A., J. D. Ostry, and J. Zettelmeyer (2012). What makes growth sustained? *Journal of Development Economics* 98(2), 149–166.
- Bluhm, R. and A. Szirmai (2012). Institutions and long-run growth performance: An analytic literature review of the institutional determinants of economic growth. UNU-MERIT Working Paper Series 033, United Nations University, Maastricht Economic and social Research and training centre on Innovation and Technology.
- Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87(1), 115–143.
- Bun, M. J. G. and F. Windmeijer (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econometrics Journal* 13(1), 95–126.
- Bussière, M. and M. Fratzscher (2006). Towards a new early warning system of financial crises. *Journal of International Money and Finance* 25(6), 953–973.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics* 29(2), 238–249.
- Cerra, V. and S. C. Saxena (2008). Growth dynamics: The myth of economic recovery. *American Economic Review* 98(1), 439–57.
- Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies* 47(1), 225–238.
- Chamberlain, G. (1984). Panel data. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume 2 of *Handbook of Econometrics*, pp. 1247–1318. Elsevier.
- Cheibub, J. A., J. Gandhi, and J. R. Vreeland (2010). Democracy and dictatorship revisited. *Public Choice* 143(1/2), 67–101.
- Chinn, M. D. and H. Ito (2006). What matters for financial development? Capital controls, institutions, and interactions. *Journal of Development Economics* 81(1), 163–192.
- Crombrugge, D. d. and K. Farla (2012). Preliminary conclusions on institutions and economic performance. UNU-MERIT Working Paper Series 035, United Nations University, Maastricht Economic and social Research and training centre on Innovation and Technology.
- Deininger, K. and L. Squire (1998). New ways of looking at old issues: inequality and growth. *Journal of Development Economics* 57(2), 259–287.
- Diebold, F. X., G. Rudebusch, and D. Sichel (1993). Further evidence on business-cycle duration dependence. In J. H. Stock and M. W. Watson (Eds.), *Business Cycles, Indicators, and Forecasting*, Volume 28 of *Studies in Business Cycles*, Chapter 6, pp. 255–285. University of Chicago Press.

- Engerman, S. L. and K. L. Sokoloff (1997). Factor endowments, institutions, and differential growth paths among new world economies. In S. Haber (Ed.), *How Latin America Fell Behind*, pp. 260–296. Stanford, CA: Stanford University Press.
- Forbes, K. J. (2000). A Reassessment of the Relationship between Inequality and Growth. *American Economic Review* 90(4), 869–887.
- Glaeser, E. L., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004). Do institutions cause growth? *Journal of Economic Growth* 9(3), 271–303.
- Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002). Armed conflict 1946–2001: A new dataset. *Journal of Peace Research* 39(5), 615–637.
- Goemans, H. E., K. S. Gleditsch, and G. Chiozza (2009). Introducing Archigos: A dataset of political leaders. *Journal of Peace Research* 46(2), 269–283.
- Greene, W. H. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economics Letters* 107(2), 291–296.
- Greene, W. H. (2011). *Econometric Analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Greif, A. (2006). *Institutions and the path to the modern economy: Lessons from medieval trade*. Cambridge University Press.
- Hall, R. E. and C. I. Jones (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *Quarterly Journal of Economics* 114(1), 83–116.
- Hausmann, R., L. Pritchett, and D. Rodrik (2005). Growth accelerations. *Journal of Economic Growth* 10(4), 303–329.
- Hausmann, R., F. Rodriguez, and R. Wagner (2008). Growth collapses. In C. M. Reinhart, C. A. Vegh, and A. Velasco (Eds.), *Money, Crises and Transition*, pp. 376–428. Cambridge, MA: MIT Press.
- Heckman, J. J. (1981). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In C. F. Manski and D. L. McFadden (Eds.), *Structural Analysis of Discrete Data With Economic Applications*, pp. 179–195. Cambridge, MA: MIT Press.
- Heston, A., R. Summers, and B. Aten (2009). Penn World Table version 6.3. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- Holtz-Eakin, D., W. Newey, and H. S. Rosen (1988). Estimating vector autoregressions with panel data. *Econometrica* 56(6), 1371–1395.
- Jerzmanowski, M. (2006). Empirics of hills, plateaus, mountains and plains: A Markov-switching approach to growth. *Journal of Development Economics* 81(2), 357–385.
- Jones, B. F. and B. A. Olken (2008). The anatomy of start-stop growth. *Review of Economics and Statistics* 90(3), 582–587.
- King, G. and L. Zeng (2001). Logistic regression in rare events data. *Political Analysis* 9(2), 137–163.
- Knack, S. and P. Keefer (1995). Institutions and economic performance: cross-country tests using alternative institutional measures. *Economics & Politics* 7(3), 207–227.
- Kuznets, S. (1955). Economic growth and income inequality. *American Economic Review* 45(1), 1–28.
- Marshall, M. G. and K. Jagers (2010). Polity IV Dataset. Center for International Development and Conflict Management, College Park (MD), University of Maryland.
- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica* 46(1), 69–85.
- Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica* 16(1), 1–32.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica* 49(6), 1417–1426.
- North, D. C. and R. P. Thomas (1973). *The Rise of the Western World, a New Economic History*. Cambridge and New York: Cambridge University Press.
- North, D. C., J. J. Wallis, and B. R. Weingast (2009). *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge and New York: Cambridge University Press.
- Orme, C. D. (2001). Two-step inference in dynamic non-linear panel data models. School of Economic Studies, University of Manchester. mimeo.
- Perotti, R. (1996). Growth, income distribution, and democracy: What the data say. *Journal of Economic Growth* 1(2), 149–187.
- Pritchett, L. (2000). Understanding patterns of economic growth: searching for hills among plateaus, mountains, and plains. *World Bank Economic Review* 14(2), 221–250.
- Rabe-Hesketh, S. and A. Skrondal (2013). Avoiding biased versions of wooldridges simple solution to

- the initial conditions problem. *Economics Letters* 120(2), 346–349.
- Reddy, S. and C. Minoiu (2009). Real income stagnation of countries 1960–2001. *Journal of Development Studies* 45(1), 1–23.
- Rodrik, D. (1999). Where did all the growth go? External shocks, social conflict, and growth collapses. *Journal of Economic Growth* 4(4), 385–412.
- Rodrik, D. (2000). Institutions for high-quality growth: what they are and how to acquire them. *Studies in Comparative International Development (SCID)* 35(3), 3–31.
- Rodrik, D. (2008). The real exchange rate and economic growth. *Brookings Papers on Economic Activity Fall 2008*(2), 365–412.
- Rodrik, D., A. Subramanian, and F. Trebbi (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. *Journal of Economic Growth* 9(2), 131–165.
- Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics* 71(1), 135–158.
- Sachs, J. D. and A. Warner (1995). Economic reform and the process of global integration. *Brookings Papers on Economic Activity* 1995(1), 1–118.
- Solt, F. (2009). Standardizing the World Income Inequality Database. *Social Science Quarterly* 90(2), 231–242. SWIID Version 3.0, July 2010.
- Wacziarg, R. and K. H. Welch (2008). Trade liberalization and growth: New evidence. *World Bank Economic Review* 22(2), 187–231.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126(1), 25–51.
- Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics* 20(1), 39–54.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: MIT press.

Appendix A: Variables definitions

Table A-1 – Independent Variables, Description, Construction and Sources

Name	Description	Construction	Data Source
Log GDP $_{(t-1)}$	Logarithm of GDP per capita	$\ln(RGDPCH_{i,t-1})$	PWT 6.3
Inflation $_{(t-1)}$	Change in consumer prices	$100 \times \ln(1 + [gCPI_{i,t-1}])$	IFS & WDI ^a
Δ ToT $_{(t-1)}$	Change in terms of trade	$\ln(TOT_{i,t-1}) - \ln(TOT_{i,t-2})$	WDI & IFS ^b
Δ Real Exports $_{(t-1)}$	Change in exports volumes	$\ln(EXP_{i,t-1}) - \ln(EXP_{i,t-2})$	WDI & IFS ^c
RER Value $_{(t-1)}$	Real exchange rate valuation	see note ^d	PWT 6.3
Fin. Openness $_{(t-1)}$	Capital account openness	KAOPEN $_{i,t-1}$	Chinn-Ito '09
Trade Openness $_{(t-1)}$	Trade liberalization measure	1 if open in $t - 1$	W-W '08
Inequality $_{(t-1)}$	Gini coefficient for income	GINI $_{i,t-1}$	Solt '09
Polity2 $_{(t-1)}$	Revised combined polity score	DEMOC $_{i,t-1}$ – AUTOC $_{i,t-1}$	Polity IV
Regchange + $_{(t-1)}$	Positive regime change	based on REGTRANS $_{i,t-1}$ ^e	Polity IV
Regchange – $_{(t-1)}$	Negative regime change	based on REGTRANS $_{i,t-1}$ ^e	Polity IV
Leader Exit $_{(t-1)}$	Irregular exit of leader	1 if EXIT $_{i,t-1} \neq 1$ ^g	Archigos 2.9
War/Conflict $_{(t-1)}$	Conflicts (≥ 1000 deaths)	see note ^f	UCDP/PRIO

^a We use the IFS series (CPI y-o-y %-change based on line 64) as a benchmark and append it with the WDI series in 59 cases where the former has missing data.

^b We use the WDI series as a benchmark (which comprises of UNCTAD and IFS data) and append it with the export volume index from the IFS for missing years/countries.

^c From the WDI 2011, we use the series ‘net barter terms of trade’, and from the IFS, we derive the equivalent net barter terms of trade by dividing the unit value of exports (line 74) by the unit value of imports (line 75) and multiplying the result by 100. We then append the WDI series of growth rates with the growth rates from the IFS series whenever the former has missing information.

^d Rodrik (2008) proposes a simple way to calculate an index of “real” exchange rate (RER) overvaluation based only on the Penn World Tables. The method involves three steps. (1) compute the PPP-adjusted exchange rate: $\ln(RER_{it}) = \ln(XRAT_{it}/PPP_{it})$. (2) Estimate the Balassa-Samuelson effect: $\ln(RER_{it}) = \alpha + \beta \ln(RGDPCH_{it}) + \gamma_t + u_{it}$. (3) Take the difference between the actual RER and the predicted RER from (2), hence: $RER\ Value_{(i,t-1)} = \ln(RER_{(i,t-1)}) - \ln(\widehat{RER}_{i,t-1})$.

^e We use the Polity IV variable REGTRANS to identify regime changes in either direction based on a minimum 3-point change in a country’s democracy or autocracy score. We exclude the code 0 for “minor changes”, which denotes *any* change in the democracy or autocracy scores. Further, we do not code -77 for “interregna”, -66 for (foreign) “interruptions” and -88 for regime “transitions” as negative regime changes to avoid collinearity with the leader exit and war/conflict dummies.

^f This dummy is constructed based on the UCDP/PRIO Armed Conflict Dataset v.4-2011, 1946–2010. We first converted the conflict-year database into country-year format and then coded the intensity levels for the highest intensity conflict in any given country-year. The dummy is unity if the intensity level of the conflict was coded as 2 in $t - 1$ and the country was listed as a location of the conflict.

^g The Archigos 2.9 time-series database records entries, tenure and exits of country leaders and the conditions on which they entered and exited. In some instances there are multiple observations per country-year, in such an event we code an irregular exit if any one observation within that year is identified as “irregular”. Irregular exit refers to leaders that died in office, committed suicide, or left office due to ill health, other irregular means or the deposition by another state.

Appendix B: Average partial effects

Wooldridge (2005) shows that in the dynamic random effects probit model, a consistent and \sqrt{N} -asymptotically normal estimator of the APEs of time-varying variables is available. However, the APEs of time-invariant covariates are not identified. Using the same assumptions as in Section 5, we can write the conditional expectation over the distribution of μ_i as:

$$E[\Phi(\alpha y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta} + \mu_i)] = E[\Phi(\alpha y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta} + \eta_0 + \eta_1 y_{i1} + \bar{\mathbf{x}}'_i \boldsymbol{\eta}_2 + \varepsilon_i)] \quad (\text{B-1})$$

where \mathbf{x}_{it} denotes time-varying regressors, $\bar{\mathbf{x}}_i$ their time-averages, and $\varepsilon_i = \mu_i - \eta_0 - \eta_1 y_{i1} - \bar{\mathbf{x}}'_i \boldsymbol{\eta}_2$. The expectation runs over the distribution of $(y_{i1}, \bar{\mathbf{x}}_i, \varepsilon_i)$.

Following Wooldridge (2005), we can get rid of the unobserved effects by applying the law of iterated expectations and defining the average structural function (ASF):

$$\text{ASF}(\mathbf{x}_{it}, y_{i,t-1}) = E[E[\Phi(\alpha y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta} + \eta_0 + \eta_1 y_{i1} + \bar{\mathbf{x}}'_i \boldsymbol{\eta}_2 + \varepsilon_i) | y_{i1}, \bar{\mathbf{x}}_i]] \quad (\text{B-2})$$

$$= E[\Phi[(\alpha y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta} + \eta_0 + \eta_1 y_{i1} + \bar{\mathbf{x}}'_i \boldsymbol{\eta}_2) \times (1 + \sigma_\varepsilon^2)^{-1/2}]] \quad (\text{B-3})$$

$$= E[\Phi(\alpha_\varepsilon y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta}_\varepsilon + \eta_{\varepsilon,0} + \eta_{\varepsilon,1} y_{i1} + \bar{\mathbf{x}}'_i \boldsymbol{\eta}_{\varepsilon,2})] \quad (\text{B-4})$$

Assuming that $\varepsilon_i \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2)$ and $\varepsilon_i \perp (y_{i1}, \bar{\mathbf{x}}_i)$, the scaled parameters in (B-4) are what standard random effects probit estimates. The panel-level likelihood based on the density of $(y_{i1}, \dots, y_{iT_i})$ given $(y_{i1}, \bar{\mathbf{x}}_i)$ can be written as

$$\mathcal{L}_i = \int_{\mathbb{R}} \left[\prod_{t=2}^{T_i} \Phi(q_{it}(\alpha y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta} + \eta_0 + \eta_1 y_{i1} + \bar{\mathbf{x}}'_i \boldsymbol{\eta}_2 + \varepsilon_i)) \right] \sigma_\varepsilon^{-1} \phi(\varepsilon_i / \sigma_\varepsilon) d\varepsilon_i \quad (\text{B-5})$$

where $q_{it} = (2y_{it} - 1)$ and $\phi(\cdot)$ is the standard normal pdf. The model log-likelihood is $\ln \mathcal{L} = \sum_{i=1}^N \ln \mathcal{L}_i$.

The scaled parameters also show up directly in the average structural function (ASF). A consistent estimator of the ASF is the simple average across all observations. The derivative of the ASF with respect to a continuous time-varying regressor, the finite difference for a binary regressor, or a mix thereof, is equivalent to the APE of that variable. This approach can be easily extended to include interactions with the lagged state and other non-linearities. Wooldridge's device to get to the APEs is to always average across the distribution of $(y_{i1}, \bar{\mathbf{x}}_i)$ first and then to specify the derivatives/differences. E.g., the APE of a continuous time-varying variable (without an interaction) in \mathbf{x}_{it} is:

$$\text{APE}(\beta_k) = \beta_{\varepsilon,k} \left[\frac{1}{N(\bar{T} - 1)} \sum_{i=1}^N \sum_{t=2}^{T_i} \phi(\alpha_\varepsilon y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta}_\varepsilon + \eta_{\varepsilon,0} + \eta_{\varepsilon,1} y_{i1} + \bar{\mathbf{x}}'_i \boldsymbol{\eta}_{\varepsilon,2}) \right] \quad (\text{B-6})$$

where we also average over time to obtain a single scale factor. It is straightforward to apply the results of Section 5 to the APEs of the interaction terms.

Chapter 3

Do weak institutions prolong crises?

*On the identification, characteristics, and duration of declines during economic slumps**

Abstract

This chapter characterizes economic slumps by a trend break or shift in the growth regime with a restricted pattern. We identify 58 episodes between 1950 and 2008 among 138 countries, and then examine the duration of the decline phase. Some declines last very long and we put several contributing factors to the test. We find evidence that weak political institutions precede crises and reforms occur thereafter. Strong political institutions shorten the duration of declines, while ethnic cleavages prolong them. Furthermore, we observe an interaction effect between institutions and ethnic cleavages, suggesting that strong institutions can overcome the adverse effects of fractionalization.

Keywords: economic slumps, crises, institutions, structural breaks, duration analysis

JEL Classification: O43, O11, C41, F43

*This chapter is based on Bluhm, R., de Crombrughe, D. and A. Szirmai. 2014. Do weak institutions prolong crises? On the identification, characteristics, and duration of declines during economic slumps, CESifo Working Paper Series 4594, CESifo Group Munich.

1 Introduction

The last sixty years of growth have been far from steady. For every “growth miracle” we can easily find a counterpart in the form of a collapse. For example, the East Asian miracle was interrupted by the Asian financial crisis, China’s take-off in 1978 was preceded by decades of adverse economic policies, Latin America was frequently rocked by political turmoil and economic volatility, and several African nations went from “up and coming” in the 1950s to stagnation after 1973. Moreover, during the post-war period, there is a long list of relatively short-lived advanced economy crises. What can we learn from such abrupt changes in growth? Why do some countries deal better with negative growth shocks than others?

The instability of growth is of great concern in economics because it affects short-run and medium-run growth performance (e.g. Ramey and Ramey, 1995). A growing literature on trend breaks has established that growth is often not steady but instead characterized by switching among growth regimes. Growth does not follow a constant trend but consists of qualitatively different episodes, such as crises, recoveries, periods of stagnation, and accelerations. This perspective provides new stylized facts. For example, growth is relatively easy to ignite (Hausmann et al., 2005) but much harder to sustain (Berg et al., 2012). However, the negative implications of volatile growth are just beginning to be explored. Long-lasting slumps can nullify decades of positive growth and there is no guarantee that lost potential output is ever fully recouped (Cerra and Saxena, 2008). It thus becomes important to ask, why do some declines last so much longer than others?

A potential explanation is that the duration of declines during slumps is driven by the prevailing structure and quality of institutions. Institutions create specific political and economic incentives, solve or worsen coordination failures and define the set of feasible policies. Seminal contributions to the institutions and growth literature link stronger institutions to higher *levels* of GDP per capita (Acemoglu et al., 2001, 2002). Others have shown that strong institutions, democracy and political stability bring about reduced output volatility (e.g. Acemoglu et al., 2003; Mobarak, 2005). However, there is still a lack of evidence convincingly linking institutions to short and medium-run growth dynamics.

Each type of growth episode has distinct characteristics. We can analyze the switching among growth regimes, the rate of growth within a regime, the duration of a regime, or even the typical sequence of regime switches that makes up a growth path. Out of this plethora of possibilities, this chapter focuses on three points. First, how can we identify large economic slumps empirically? Second, is there any evidence of institutional change when slumps occur? Third, conditionally on the occurrence of a slump, do weak institutions prolong the duration of the decline phase?

We specifically focus on the duration of declines, as the onset of economic crises can be triggered by a variety of external or internal factors which are not (always) linked to weak institutions. However, the way a country deals with a negative shock, and whether the decline phase takes longer than in other countries, depends on the political system’s ability to react with coordinated policies and avert outright social conflict. This notion derives from a large body of political economy theory putting social tension and the ability of resilient political institutions to manage such conflict at the center of development theory (e.g. Acemoglu and Robinson, 2006; North et al., 2009; Besley and Persson, 2011). Some of these theories argue that weakly institutionalized societies, or ‘limited access orders’, are prone to collapses, and that the declining rents associated with

a crisis undermine the institutional set-up and the prevailing political arrangements (e.g. North et al., 2009). Weak institutions thus bring with them an increased vulnerability to crises and potentially much longer declines once slumps occur. Similar mechanisms are suggested in the literature on institutions and macroeconomic volatility (Acemoglu et al., 2003). Even if sound policy responses are available, a combination of coordination failures, rent seeking and power struggles combined with ethnic cleavages may lead to longer declines in weakly institutionalized environments.

Ethnic heterogeneity itself has been linked to a variety of coordination failures leading to inadequate policies, low provision of public goods and conflict.¹ However, diversity may also be beneficial. In fact, a particular level of heterogeneity may be optimal (Ashraf and Galor, 2013) and necessary to reap the advantages of skill complementarities in highly diversified economies (Alesina and Ferrara, 2005). Even if there are no benefits to diversity, its negative effects may become muted as “richer societies have developed institutional features that allow them to better cope with the conflict element intrinsic in diversity” (Alesina and Ferrara, 2005, p. 763). We show that in the context of economic declines the (negative) effects of heterogeneity depend on political institutions.

The ‘delayed stabilizations’ literature provides another perspective on how conflict over the costs of adjustment always leads to delayed reform (Alesina and Drazen, 1991) and how political institutions affect the timing of the adjustment (Spolaore, 2004). Similarly, the model of policy non-adoption by Fernandez and Rodrik (1991) outlines a status-quo bias due to uncertainty about the benefits of reform when the losers cannot be compensated *ex post*. These ideas naturally carry over to the process of formulating policy responses during slumps where more group heterogeneity can increase delay.

Our findings are consistent with these theoretical perspectives. First, there is evidence of weak political institutions preceding slumps and signs of political reforms in the following years. Second, longer decline phases are robustly linked to weak institutions and particularly strongly to a measure of ethnic cleavages (ethno-linguistic fractionalization). Ethnic cleavages are especially important for understanding declines in Sub-Saharan Africa. Third, we show that political institutions and ethnic fractionalization interact. In weakly institutionalized and highly fragmented societies declines last considerably longer than in more homogeneous countries or countries with stronger constraints on the executive. Current political economy theory does not sufficiently capture these findings. We also provide evidence suggesting that these factors determine the depth of a slump, but exclusively through the duration of declines and not the average rate of contraction.

The chapter is structured as follows. Section 2 outlines the identification of slumps and defines the decline phase. Section 3 briefly discusses the data and characteristics of the estimated slumps, and provides evidence of endogenous institutional change. Section 4 analyzes the duration of the decline phase and discusses the results. Section 5 concludes.

2 Identifying slumps

Restricted structural breaks

Beginning with Pritchett’s (2000) classification of post-World War II growth experiences into “Hills, Plateaus, Mountains, and Plains”, a growing literature analyzes the characteristics of different types of growth episodes. These papers usually employ either

¹For a review see Alesina and Ferrara (2005).

simple or more complex tests of structural stability to define and identify the episode of interest.² For example, Hausmann et al. (2005) use economic criteria to isolate growth accelerations and then date their beginning with a simple breakpoint test. Other authors, such as Jones and Olken (2008) and Berg et al. (2012), use versions of the Bai and Perron (1998, 2003) test for multiple structural breaks to classify growth episodes. A third set of papers solely relies on economic criteria to identify and date the episode of interest (e.g. Hausmann et al., 2008).

Not every change in growth rates amounts to a regime switch. The main advantage of econometric tests for multiple structural breaks over any set of predefined economic criteria is that they allow for an inferential approach to identifying growth regimes and measuring their duration. Methods based on deterministic economic criteria cannot discriminate among multiple plausible starting points or assess whether an episode truly constitutes a departure from the previous growth regime. However, since standard structural break methods leave the particular type of structural change unspecified, these tests may not identify the theoretically desired type of regime switch but rather any form of significant change which must then be classified *ex post*. Furthermore, while break estimators work well for identifying growth spurts, they perform poorly when it comes to identifying recessions or growth collapses.³

To improve the identification of what we interchangeably refer to as deep recessions, slumps, or growth collapses, Papell and Prodan (2012) propose a *two-break model with parameter restrictions*. They demonstrate that this modified structural change approach consistently identifies well-known slumps, such as the Great Depression in the United States. The key innovation is to impose features of the desired pattern directly instead of searching for unrestricted structural changes. Their two-break model accounts for three growth regimes (a pre-slump regime, a contraction/ recovery regime, and a post-slump regime) and places sign restrictions onto the estimated coefficients to ensure the breaks occur in the desired direction. Since this approach is a version of Bai's (1999) sequential likelihood ratio test, the number of slumps – which is not known in advance – can then be estimated by recursively applying the model on ever smaller sub-samples until all breaks in the GDP per capita series have been found.

The restricted structural change approach can easily be modified in principle to allow for other plausible structures, such as three-break models (e.g., to estimate a pre-slump regime, a decline, a recovery and a post-slump regime). However, estimating three or more breaks for each slump quickly becomes computationally expensive and does not necessarily provide better results than a simpler two-break model.⁴ While Papell and Prodan (2012) focus on the question whether growth in a few developed countries eventually returns to its pre-slump trend path, we apply a variant of this method to identify slumps in a large sample of countries over the period from 1950 to 2008.

We define slumps according to three criteria. First, a slump is a *departure from a previously positive trend*. Second, a slump must begin with *negative growth in the first year*. Third, all slumps should be *pronounced regime switches* and not just minor business cycle fluctuations. The precise meaning of 'pronounced' will vary and depend

²With the exception of the Markov-switching models used in Jerzmanowski (2006).

³As Papell and Prodan (2012) point out, this applies to the entire class of "generic" tests for single and multiple breaks.

⁴Let $q = T - 2\tau T - h$, where τ is the trimming fraction and h is the distance between breaks, then the two-break model estimates $(q^2 + q)/2$ regressions for the first iteration, while a three-break model already requires $\sum_{i=1}^q (i^2 + i)/2 = (1/12)q(q+1)(2q+4) - 1$, with $q = T - 2\tau T - 2h$ to now allow for three breaks. Additional results are available on request.

on the country's idiosyncratic growth process.

We capture these criteria in the following partial structural change model:

$$y_t = \alpha + \beta t + \gamma_0 \mathbf{1}(t > tb_1) + \gamma_1 (t - tb_1) \mathbf{1}(t > tb_1) + \gamma_2 (t - tb_2) \mathbf{1}(t > tb_2) + \sum_{i=1}^p \delta_i y_{t-i} + \epsilon_t \quad (1)$$

where y_t is the log of GDP per capita, β is a time trend, γ_0 is the coefficient on an intercept break occurring together with a trend change (γ_1) after the first break at time tb_1 , γ_2 is the coefficient for a second trend change occurring after the second break at time tb_2 , $\mathbf{1}(\cdot)$ is an indicator function selecting the regime, p is the optimal lag order determined by the Bayesian information criterion (BIC) to parametrically adjust for the presence of serial correlation, and $\{\epsilon_t\}$ is a martingale difference sequence satisfying $\mathbb{E}[|\epsilon_t|] < \infty$ and $\mathbb{E}[\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots] = 0$.

Eq. (1) formalizes the notion that the evolution of GDP per capita around a slump is a simple function of time split into three different growth regimes: (1) a *pre-slump regime* from the beginning of the time series of a country until time tb_1 , (2) a *slump/recovery regime* lasting from time $tb_1 + 1$ to time tb_2 , and (3) a *post-slump regime* from time $tb_2 + 1$ onwards. The true location of the breakpoints is not assumed known but estimated within the model. We impose *two restrictions* to make sure we only select breaks meeting our definition of slumps. First, we require $\beta > 0$, so that growth must be positive in the years before a slump begins. Second, we also impose the condition that $\gamma_0 < 0$, so that a slump always starts with a drop in the intercept.⁵ Slope shifts are left unrestricted, so that the model can catch unfinished slumps (e.g., declines from tb_1 onwards, possibly lasting until the end of a country's time series).

We implement the sequential break search algorithm as follows. First, we fit the structural change model specified in eq. (1) for all possible combinations of tb_1 and tb_2 . We always exclude 5% of the observations at the beginning and end of the sample to avoid registering spurious breaks. Second, we compute the sup- W test statistic, that is, the supremum of a Wald test of the null hypothesis of no structural change ($\mathbb{H}_0 : \gamma_0 = \gamma_1 = \gamma_2 = 0$) over all pairs of break dates implying estimates satisfying both restrictions. Third, we bootstrap the empirical distribution of the sup- W statistic. If the bootstrap test rejects at the 10% significance level, we record the break pair (\hat{tb}_1, \hat{tb}_2) and split the sample into a series running until the first break and a series starting just after the second break. The process starts again on each sub-sample until the bootstrap test fails to reject the null hypothesis of no breaks or the sample gets too small. A key issue in evaluating the statistical significance of breakpoints is that the individual Wald tests over which the sup- W statistic is computed are not independent. Asymptotic tests have been derived but they tend to underreject in finite samples (Prodan, 2008) and an asymptotic distribution for our particular version of restricted structural change is not available. We construct a bootstrap Monte Carlo test in order to circumvent both issues. Appendix A gives a formal description of the break search algorithm and the bootstrap.

The structural break methods applied in this chapter assume that GDP per capita is regime-wise trend stationary. This is not a trivial requirement. Ever since the issue was first raised by Nelson and Plosser (1982), a vibrant literature has been debating the question whether most GDP series are unit-root processes or can be considered trend stationary. Originally, the conflicting views evolved around a clear divide. If an output

⁵The intercept shift implies that we assume that there is an instantaneous drop. However, by not restricting the coinciding trend break we also allow for longer lasting declines.

series is non-stationary, i.e., it has a unit root, then any stochastic shock to the economy is permanent. If the series is trend stationary, then stochastic shocks are temporary; after a while GDP is back on track as if the shock never occurred. Given the available data, this issue cannot be fully resolved. It is generally difficult – if not impossible – to convincingly differentiate between non-stationary and stationary time series when T is only moderately large.

More recently the debate has shifted. A process that is subject to structural breaks is an intermediate case. Broken-trend stationarity only implies that, within each regime, growth can be approximated by a deterministic trend, but from one regime to the next the trend path may change due to (semi-)permanent shocks. This allows for a flexible description of the growth process as several different types of trend breaks can occur. In fact, there is mounting evidence that once trend breaks are incorporated, many of the GDP series previously thought to have unit roots may in fact be broken-trend stationary (e.g. Zivot and Andrews, 1992; Ben-David and Papell, 1995). Broken trends blur the conceptual distinction. A unit root process can be thought of as a trend-stationary process with a trend that changes every year.⁶

We do not attempt to characterize all types of breaks an economy can experience, or to formally test for unit roots. Our approach is flexible and allows for different growth regimes occurring before, during and after an unknown number of slumps. We assume that there is some structure in the growth process, but do not assume that it is necessarily generated by neoclassical steady-state growth, endogenous growth or any other standard growth model. In fact, Aguiar and Gopinath (2007) recently highlighted that growth in emerging markets can be characterized by shocks to trend growth rather than transitory fluctuations around a stable trend. Thus, under certain conditions, broken trends are compatible with various models of aggregate output.

The duration of declines

Within a slump, we separate the decline from the recovery phase and focus solely on the decline phase. This is based on the conjecture that these two processes are potentially subject to very different dynamics and explained by different covariates. This “to the bottom” approach stands in contrast to the earlier literature which typically focuses on the entire duration of the slump (until the pre-slump level of GDP is regained). Instead, we argue that political institutions and ethnic cleavages matter particularly for the duration of the decline phase where policy solutions need to be agreed upon to achieve a turnaround, as opposed to the length of the recovery process which depends on the success of the chosen policy and, more generally, the type of crisis.

We define the end of a slump to have occurred with certainty in the first year $a > \hat{t}b_1$ where $y_a \geq y_{\hat{t}b_1}$. In other words, a slump is over as soon as the level of GDP per capita preceding the slump has been recovered; until then, the slump is continuing.⁷ It is important to note that the end of the slump does not in general coincide with the second break and is used only as a device to identify the trough. Once the endpoint of a slump is known, *the trough is simply the year with the lowest level of GDP per capita during the slump*. The duration of the slump is censored if GDP per capita does not reach the

⁶Every year, or at any other observation frequency.

⁷This also implies that we exclude a small number of episodes found by the sequential algorithm if these begin before the previous slump is certain to have ended. Avoiding this issue would imply estimating a third breakpoint (see footnote 4).

pre-slump level again by the end of the sample. In such a case, even if GDP per capita seems to be recovering, we do not know how long the slump may last. The censoring indicator is defined as $c = \mathbf{1}(\max_{j \in (\hat{t}b_1, T]} y_j < y_{\hat{t}b_1})$.

Given the set of possible end years $A = \{a \mid a \in (\hat{t}b_1, T] \text{ and } y_a \geq y_{\hat{t}b_1}\}$, define $a_0 = \min A$, corresponding to the (certain) end of the slump. If the set A is empty, then the slump is unfinished, and the length of the episode is censored. We estimate the trough to occur at time:

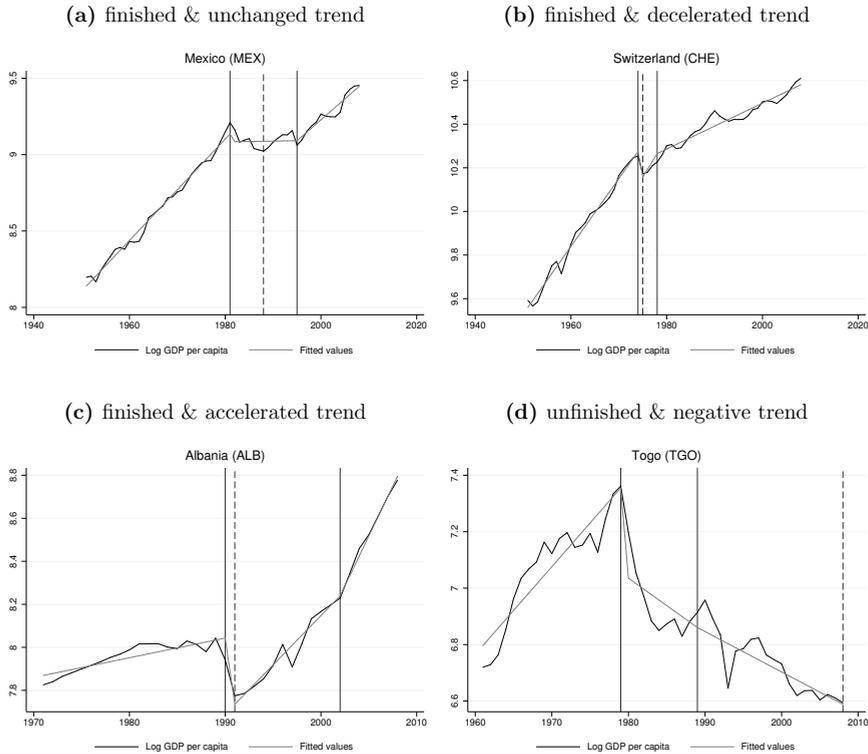
$$\hat{t}_{min} = \begin{cases} \operatorname{argmin}_{j \in (\hat{t}b_1, a_0]} y_j, & \text{if } c = 0 \\ \operatorname{argmin}_{j \in (\hat{t}b_1, T]} y_j, & \text{if } c = 1. \end{cases} \quad (2)$$

A provisional trough occurs when y_t attains a minimum after $\hat{t}b_1$. The duration of the decline phase lasting from the beginning of the slump to the estimated trough will be denoted $\tilde{t}_D = \hat{t}_{min} - \hat{t}b_1$.

These definitions also imply that in some cases we date the trough after the estimated second break, which is purely a consequence of allowing for unfinished episodes. If the slump is still ongoing, the second break may have been placed at an arbitrary point maximizing the Wald statistic but not corresponding to the start of a new growth regime.⁸ The true trough may lie in the future, that is, beyond the end of the sample. Treating such spells as censored implies that in the later analysis we only incorporate the information that (certain) exit from the slump has not yet occurred after a duration \tilde{t}_D .

Figure 1 illustrates the diversity of slumps identified by this method. Panel (a) shows a finished slump in Mexico where the trend growth rate is nearly unchanged after the second break. The slump begins in 1982 and encompasses more than a decade of political volatility, hyperinflation, high debt and low growth. The trough is found in 1988. Another short downturn occurs during the Tequila crisis in 1994 after which the Mexican economy returns to its pre-1982 growth path. Panel (b) shows a finished slump in Switzerland where the trend growth rate decelerated after the second break. In 1975, the Swiss economy was strongly affected by the Oil crisis of the mid-1970s, leading to a 7.87% drop in GDP per capita within one year. After the second break, Switzerland enters a low growth regime typical for the high income economies in Western Europe of the 1980s and 1990s. Panel (c) shows a finished slump in Albania occurring at the time of the post-communist transition with an accelerated trend after the second break. The estimated first break occurs in 1990, the trough is located in 1991, and the second break occurs in 2002, a few years after the end of the slump. While the duration of the decline phase is only one year, output contracted considerably. GDP per capita in 1991 was 15.32% lower than in 1990. Last but not least, panel (d) shows an unfinished slump with a continuing decline in Togo. Togo grew rapidly for over a decade following independence from France in 1960 but then experienced a dramatic collapse. The first break occurs in 1979, but the second break is placed at an (economically) arbitrary point to accommodate the lasting decline. Togo's GDP per capita did not recover to its pre-slump level for the next 29 years. At the end of the observed period, the decline is ongoing and the provisional trough coincides with the censoring cutoff in 2008.

⁸A solution avoiding this problem would be to test if a restricted one-break model fits better than a restricted two-break model for those cases.

Figure 1 – Four types of slumps

Note(s): Models refitted using the estimated breaks \hat{tb}_1 and \hat{tb}_2 but without the optimal $AR(p)$ terms to emphasize the trend breaks. The bold vertical lines are at \hat{tb}_1 and \hat{tb}_2 , respectively. The dashed vertical line indicates \hat{t}_{min} .

3 Data and characteristics of slumps

We apply the sequential algorithm to the entire Penn World Table (v7.0) yielding a total of 58 slumps between 1950 and 2008.⁹ We deliberately stop in 2008 to avoid the global recession of 2009 which is too close to the end of the sample for reliable break estimation. The mean duration from the first break to the trough is about 7.7 years and the median duration is three years. Ten out of the 58 slumps are censored. For these spells the location of the trough is not yet definitive. Table B-2 in the Appendix lists all episodes and provides some summary statistics.

Many covariates used in the sequel are from well-known sources and will be

⁹We only run the algorithm on countries with a population of at least one million (excluding very small countries and island economies), and at least 20 years of data. In addition, we discard some episodes that are driven by positive breaks in the slope coefficient but fail the negative growth criterion due to the presence of the $AR(p)$ terms. A simple rule is applied to these cases. We define a valid episode as an interval of two break dates $\hat{tb}_1, \hat{tb}_2 \in [\tau T, (1 - \tau)T]$ satisfying: $\exists j \in (\hat{tb}_1, \hat{tb}_2]$ such that $\min y_j < y_{\hat{tb}_1}$, where τ is the trimming fraction and T is the length of the estimation sample. This rule *only* requires that a contraction occurs within the range of the two estimated breaks, otherwise there is no slump.

discussed only summarily here. We include four major categories of variables: 1) a variety of measures for different aspects of *institutions, politics and social conflict*, 2) macroeconomic indicators of *prices, trade and exports*, 3) a set of variables for domestic and international *finance*, and 4) several *other growth determinants* (such as life expectancy or years of schooling). Table C-1 in the Appendix provides an exhaustive list of all variable names, data sources and a basic set of summary statistics for the data used throughout the chapter. Some data may not be entirely satisfactory but are simply the best available. For example, we use the Polity IV database as our primary proxy for institutional development because of a lack of other time series data capturing the characteristics of political and economic institutions.

We observe several well-known growth collapses and deep recessions. Most slumps begin between the 1970s and the early 1990s. Seven downturns occur following the oil shock in 1973–1974, eleven declines begin between 1979 and 1981 during the debt crisis of the early 1980s, and nine slumps follow the post-communist transitions of 1989–1990. Due to trimming and the cut-off in 2008, we find no beginnings of slumps in the period of the early 2000s and tranquil mid-2000s. Generally, the period between the 1970s and early 1980s is marked by heightened volatility, as has been documented in a number of studies (Easterly et al., 1993; Rodrik, 1999; Pritchett, 2000; Jones and Olken, 2008).

Table 1 – Depth and Duration by Income Level and Geographical Region

	Mean Depth	Mean Duration	Median Duration	Number of Spells	Censored Spells	Number of Countries
<i>Income Level</i>						
High Income (OECD)	-7.12	2.00	1	12	0	29
High Income (Other)	-20.84	5.38	2	8	1	12
Upper Middle Income	-21.20	5.39	2	16	2	30
Lower Middle Income	-27.40	6.00	3	11	3	34
Low Income	-34.17	15.75	16	11	4	33
<i>Geographical Region</i>						
East Asia & Pacific	-13.63	2.30	2	10	0	17
Eastern Europe & Central Asia	-19.70	3.40	2	5	0	10
Europe (excl. Eastern Europe)	-8.37	1.50	1	6	0	22
Latin America & Caribbean	-17.34	5.27	3	15	1	23
Middle East & North Africa	-33.24	8.66	9	7	3	17
North America	-2.51	1.00	1	1	0	2
South Asia	-5.33	1.00	1	1	0	6
Sub-Saharan Africa	-37.14	17.74	16	13	6	41
Total	-21.87	7.69	3	58	10	138

Note(s): Depth is defined as the percent decrease in GDP per capita at the trough relative to GDP per capita before the slump (not log difference). Mean and median duration are expressed in years. As a result of some spells being censored, both mean duration and depth are underestimated. The number of countries refers to countries with more than one million inhabitants and more than 20 observations of GDP per capita in a particular income group or region.

Table 1 summarizes the distributions of depth, duration, and number of spells across income groups and geographical regions. For this purpose, we define the depth of a decline as the percent decrease of GDP per capita at the trough relative to its pre-slump level. We detect considerably deeper slumps in low-income and middle-income countries than in high-income (OECD) countries. The spread of depth and duration is very large. High-income (OECD) countries experience relatively short declines with a comparatively soft landing. The median duration is only one year with a mean depth of about -7.1%. In

the middle, there is little substantial variation between non-OECD high-income countries and upper/lower-middle-income countries. In all of these three groups, the mean depth is in the range of -20.8% to -27.4% and the mean (median) duration varies between about 5.4 to 6 (2 to 3) years. Low-income countries experience the most dramatic declines. Both mean and median duration are about 16 years, with an associated average depth of -34.2%. Interestingly, the number of spells itself is distributed relatively evenly across the different income groups, suggesting that developed countries, too, experience their fair share of (milder) volatility.

The geographical distribution reveals three interesting patterns. First, Sub-Saharan Africa and the Middle East & North Africa are the two regions experiencing both the deepest and longest declines. The depth in these regions is particularly striking in comparison to the other regions. The mean decline is -37.1% in Sub-Saharan Africa and -33.2% in the Middle East & North Africa, which is about double of the average decline in Latin America & the Caribbean. The duration is longest in Sub-Saharan Africa, with the median spell lasting 16 years and the mean spell lasting over 17 years. Declines are shorter in the Middle East & North Africa, where the mean and median do not exceed nine years. Both regions also have the most censored/unfinished spells due to their long duration. Second, countries in Latin America & the Caribbean experienced slumps most frequently, but the average decline was only moderately deep and lasted for about five years. Third, we find similar mean and median durations when comparing Eastern Europe & Central Asia to East Asia & Pacific, but much deeper slumps in the former. As expected, there are comparatively few, short and mild declines in North America, Europe (excluding Eastern Europe), and – more surprisingly – South Asia.

Table 1 suggests a relatively strong association of both the mean duration and mean depth of the decline phase with different income levels. This is particularly interesting, since we subsequently model these observed differences in duration between high and low income economies with more fundamental factors such as institutions and ethnic fractionalization. Furthermore, there is substantial regional heterogeneity which will have to be taken into account in the analysis.

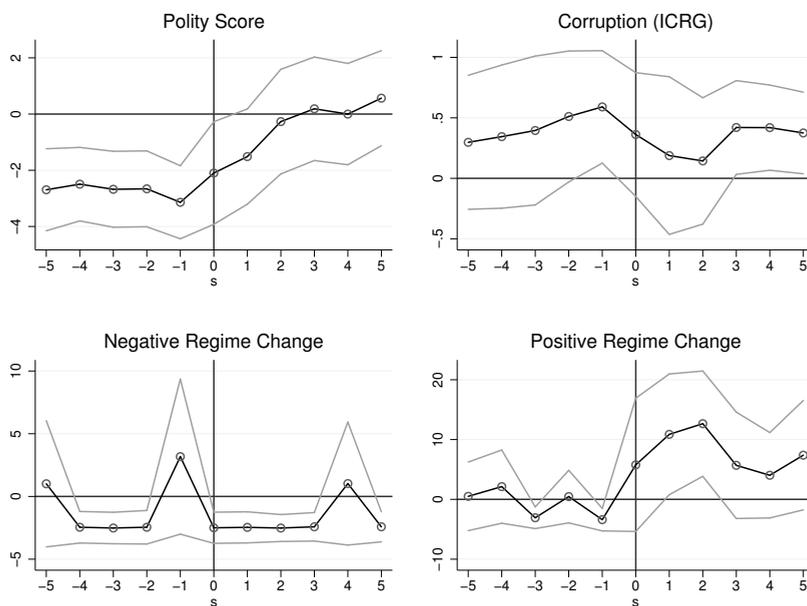
Is there evidence of institutional change occurring before, during or after a slump? To study this question descriptively, we employ an event methodology often used in the literature on currency and banking crises (e.g. Eichengreen et al., 1995; Gourinchas and Obstfeld, 2012). The basic idea is to use dummy variables indicating the imminence or recentness of the start of the slump as a means of detecting changes in the relative mean of each time-varying covariate. The coefficients of the dummies measure if and how the covariate changes around the time the slump hits, and their standard errors quantify the associated uncertainty.

We run the following regression for each measure of institutions: $x_{it} = \sum_{s=-5}^5 \delta_{t, \hat{t}b_1+s} \beta_s + \mu_i + \epsilon_{it}$ where $\delta_{t, \hat{t}b_1+s}$ is the Kronecker delta which is equal to one if $t = \hat{t}b_1 + s$ and zero otherwise, β_s are coefficients, μ_i is an unobserved country effect and ϵ_{it} is an idiosyncratic error term. We set $s \in \{-5, \dots, 0, \dots, 5\}$, so that the result is an 11-year window around the break date $\hat{t}b_1$. The first year of the slump is $s = 1$ corresponding to $t = \hat{t}b_1 + 1$. The standard errors are robust to heteroskedasticity and also autocorrelation within both country *and* time clusters (Cameron, Gelbach, and Miller, 2011). We plot the estimates of the coefficients (including 95% confidence bands) as they represent *the conditional expectation of the covariate x_{it} at time s relative to “normal”*

times.¹⁰

Figure 2 shows how certain institutional and political dynamics evolve around the downbreak. The Polity score is much lower before a slump occurs, but increases towards normal levels thereafter. In the five years before a slump, the conditional expectation is between 2.5 and 3.1 points lower than in normal times and until the break date these differences are significant at the 5%-level. This suggests that prior deficiencies in institutions increase vulnerability to slumps *and* institutions start to improve when slumps occur. All the subcomponents of the combined Polity score, including *constraints on the executive*, exhibit very similar trends (not shown, available on request). Conversely, the ICRG's 6-point corruption indicator shows a moderate, yet insignificant, decrease in corruption in the first two years of a slump. The ICRG series suffers from low coverage; it begins only in 1984 while a majority of the slumps in our sample start earlier.

Figure 2 – Institutions & Politics



The association of reforms and slumps is confirmed by the time profile of the probabilities of negative or positive regime changes, measured as a minimum three-point downward or upward change in the Polity score. There is little evidence that negative regime changes precede downbreaks or systematically occur thereafter. Interestingly, there is an upward trend in the probability of positive regime changes from the eve of the slump onwards. In the first and second year of a slump, the probability is 10-12% higher than in normal times. While these point estimates are individually significant, the sequence of positive estimates is even more convincing.

This pattern suggests a new stylized fact: slumps are often preceded by weak

¹⁰In this case, “normal” refers to all observations other than the 11 years around the downbreak. The working paper version of this chapter also reports this analysis for a host of other growth determinants.

institutions. However, it seems that sharply negative growth creates room for political reforms. A bold interpretation would be that given prior institutional deficiencies, slumps bring about a form of *creative political destruction* by altering power relations and increasing the pressure on governments to pursue institutional change (North et al., 2009). In other words, deep slumps offer a window of opportunity for institutional improvements, illustrating the *endogenous nature of reforms*.

4 The duration of declines

Methodology

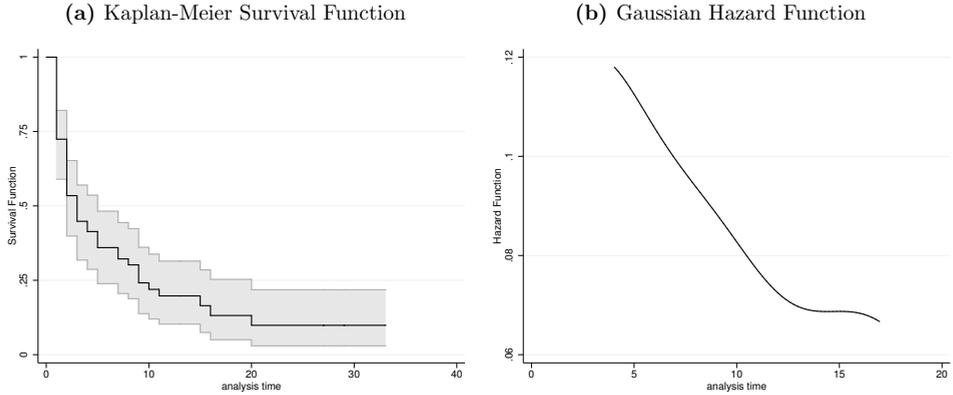
There are two major approaches to dealing with duration data: semi-parametric Cox models and parametric models. The main advantage of Cox regression is that the so-called baseline hazard can have any shape (as long as the conditional hazard is proportional). Parametric models, on the contrary, require more explicit assumptions about the shape of the baseline hazard but in turn also allow more detailed predictions.

We take a parametric approach. Parametric models can either be cast as proportional hazard (PH) or as accelerated failure time (AFT) models. PH models directly begin with a log-linear specification of the hazard function. In other words, they model the instantaneous probability of an event occurring at a particular time conditional on the event not having occurred before. Proportionality then implies that the hazard function for each subject in the sample is a multiple of the baseline; that is, the baseline hazard is scaled up and down by the different realizations of the covariates. From the hazard function, we can derive the survival function which expresses the cumulative probability of the event not having occurred until the observed time. The event of interest is usually defined as the ‘exit’ from an ‘initial state’. In our case, the ‘initial state’ corresponds to the decline phase and ‘exit’ corresponds to entering the recovery phase. If a country is observed to exit the decline phase at some time, its contribution to the likelihood is the probability of the recovery starting at that particular time (conditional on the decline phase lasting until that time). If there is no observed exit from the decline phase, then the observation is censored and only the survival probability enters the likelihood. AFT models describe the same relationships but begin with a log-linear model of time itself. Hence, AFT models closely resemble linear regression. The hazard function and survival function are then characterized indirectly (and without the restriction of proportionality) by the distribution of the error terms in the log-linear model.

All parametric models assume a certain shape of the baseline hazard. The exponential model assumes that the hazard is constant over the entire duration process. Models with a Weibull or Gompertz distribution allow for flat, monotonically increasing or monotonically decreasing hazard rates. Log-normal and log-logistic models provide a non-monotonic function that is first increasing and then decreasing. Alternatively, the generalized gamma distribution is very flexible and encompasses the exponential, Weibull and log-normal distributions but is more demanding to estimate.

We have no strong theoretical prior that the hazard function must follow a particular shape. We may expect some countries to exit rather quickly and others to take longer, but it is difficult to determine *ex ante* if remaining in the decline phase for very long leads to a deterioration of fundamentals and thus a decreasing hazard, or if the probability of exit is actually increasing because countries are bound to enter the recovery phase eventually.

Figure 3 shows the non-parametric Kaplan-Meier estimate of the unconditional

Figure 3 – Unconditional Survival and Hazard Functions

Note(s): The Kaplan-Meier survival curve is a non-parametric estimate of the probability of remaining in the decline state at each unit of analysis time. 95% confidence intervals are shown in grey. The corresponding hazard function has been smoothed using a Gaussian kernel with boundary adjustment and bandwidth 3.

survival function and the (smoothed) Gaussian hazard function. About 47% of the spells in our sample end after only two years of decline, but the unconditional probability of exiting the decline phase is monotonically decreasing. Nevertheless, the shape of the conditional hazard may be very different. We take a flexible approach by first relying on a log-normal parameterization and then testing the robustness of our preferred specification under different distributional assumptions. It is important to note that the log-normal model does not imply proportionality of the conditional hazard function. Hence, it can not be represented in PH form. We provide a more detailed description of how log-normal AFT models are estimated in Section 5.

Let analysis time be \tilde{t} , where $\tilde{t} \equiv t - t_0$ and $t_0 = \hat{tb}_1$, so that we can refer to the calendar times t and t_0 when necessary. We specify the following regression equation for crisis durations in AFT form:

$$\ln \tilde{t} \equiv \ln(t - t_0) = \alpha + \beta INS_0 + \gamma ELF + \mathbf{x}'_0 \boldsymbol{\xi} + \mathbf{z}'_t \boldsymbol{\zeta} + \epsilon_t \quad (3)$$

where INS_0 is a measure of institutions fixed at t_0 , ELF is a time-invariant measure of ethnic fractionalization, $\mathbf{x}_0 = (x_{0,1}, x_{0,2}, \dots, x_{0,k})$ is a $k \times 1$ vector of controls fixed at t_0 , $\mathbf{z}_t = (z_{t,1}, z_{t,2}, \dots, z_{t,m})$ is an $m \times 1$ vector of strictly exogenous time-varying controls, and – for the log-normal model – ϵ_t is distributed $\mathcal{N}(0, \sigma_\epsilon^2)$. All parameters, including σ_ϵ^2 , are estimated with Maximum Likelihood (ML). Unfinished spells are accounted for but censoring is assumed independent of the duration. Our coefficients of interest are β and γ . We suppress the country-spell index to simplify the exposition.

The estimated coefficients are semi-elasticities of the expected duration with respect to the covariates, or elasticities if the covariate is in logs. The term ‘accelerated failure time’ derives from the interpretation of the implied effects. If the coefficient of the covariate is positive, then the expected duration until the event is prolonged by larger realizations of the covariate. In our case, this is equivalent to delayed exit from the decline phase (later start of the recovery). If the coefficient is negative, then the expected duration is

shortened and the recovery will start earlier.

A complication of using time-varying covariates is possible feedback running from the duration to the covariates. If such feedback occurs, the estimated coefficients are biased and the usual test statistics are invalid (Lancaster, 1990; Kalbfleisch and Prentice, 2002). In order to avoid this problem, we simply take the last pre-slump value of all potentially endogenous covariates at t_0 , including our measure of political institutions, so that no feedback from slumps to the covariates is possible. Hence, we can rule out simultaneous causality. This is particularly important given that the previous section highlights that political institutions may endogenously respond to crises. On the other hand, ethnic fractionalization is assumed to be strictly exogenous; we do not expect the ethnic configuration of a country to change as a short-run response to a crisis.

The fact that countries can have several recurrent slumps is a minor concern in our application; only eight of the 58 spells in our data are not the first spell for a given country. To account for this dependence, we allow the variances of the parameter estimates to be correlated across spells in the same country. This procedure assumes that the sequence of repeated spells does not matter. We show in the robustness section that our results hold when this assumption is relaxed.

Dealing with at most 48 exits in 58 decline spells over the entire period of 1950 to 2008 requires a careful approach to model selection, since we have to match these episodes with data over the almost six decades spanned by them. Including many control variables with different patterns of missing data then easily results in small samples, so that including a large set of controls is not feasible. Even at more moderate sample sizes, care needs to be taken to guard against overfitting. We employ a two-step approach to arrive at a parsimonious specification. First, we fit variable-by-variable regressions and reduce the set of controls based on statistical significance (p -value $< .1$). We select only those variables clearly exhibiting a correlation with the duration of declines. Results of this step are relegated to Appendix E. Second, using the smaller set of controls, we then extend our base specification in several ways and examine its robustness.

Results

We model the duration of declines as a function of executive constraints, ethno-linguistic fractionalization, initial GDP per capita, and the real US interest rate. Constraints on the executive is our preferred proxy of institutional quality for two reasons. First, it is widely used in the empirical literature as a measure of institutional constraints placed on political actors and has already been linked to macroeconomic volatility (e.g. Acemoglu et al., 2003; Acemoglu and Johnson, 2005). Second, it is conceptually rooted in the economic theory of institutions, more so than any of the broader measures capturing wider aspects of the political regime (e.g. democracy or autocracy). Controlling for initial GDP matters, as executive constraints are correlated with the level of development (the correlation coefficient is 0.43), and both potentially determine the duration of declines.

For fractionalization, we use a measure from Desmet et al. (2012), who recently developed a very detailed set of estimates of linguistic diversity. They compute the probability that two randomly chosen individuals in a country belong to different ethno-linguistic groups at 15 levels of ‘the language tree’. These new measures of fractionalization capture the historical nature of ethnic and linguistic differentiation into increasingly narrower groups over time. We use two variables at both extremes of the spectrum. ELF1 is the most aggregate level, measuring only crude distinctions such as

Indo-European versus non-Indo-European languages. ELF15 is the most disaggregate level, differentiating among the language groups known today. Desmet et al. (2012) show that aggregate fractionalization matters more for civil conflict and redistribution while the disaggregate level strongly predicts growth differentials. Hence, we use the latter as our primary measure.

The variable selection results reported in Section 5 show that the basic correlations are mostly as expected. One notable exception is the lack of correlation between conflict and the duration of declines. Stronger institutions are associated with shorter declines, regardless of the measure. Higher initial GDP predicts shorter declines but the effect is not significant. Conversely, higher fractionalization and a higher US interest rate predict longer declines. Yet these findings could be driven by omitted variables.

Table 2 – Additive effects of institutions and fractionalization

VARIABLES	(1) ln \bar{t}	(2) ln \bar{t}	(3) ln \bar{t}	(4) ln \bar{t}	(5) ln \bar{t}	(6) ln \bar{t}
Executive Constraints (INS_0)	-0.181*** (0.066)	-0.149 (0.102)	-0.007 (0.075)	-0.172*** (0.063)	-0.111* (0.064)	-0.178*** (0.058)
Fractionalization (ELF_{15})	0.015*** (0.004)	0.016*** (0.006)	0.019*** (0.005)	0.016*** (0.004)	0.014*** (0.005)	0.016*** (0.004)
Initial ln GDP per capita	0.282** (0.125)	1.014*** (0.290)	0.293** (0.132)	0.559*** (0.171)	0.459*** (0.164)	0.197* (0.106)
Real US Interest Rate	0.063 (0.044)	0.065 (0.073)	0.098* (0.054)	0.081* (0.044)	0.074* (0.042)	0.087* (0.048)
Trade Openness (de jure)	-0.772** (0.312)					
Trade Openness (de facto)	0.010** (0.005)					
Manufactures (% Exports)		-0.007 (0.010)				
Export Diversification		-0.003 (0.012)				
Export Sophistication		-1.559** (0.777)				
Financial Depth			-0.007 (0.013)			
Private Credit			-0.011 (0.008)			
Infant Mortality				0.003 (0.009)		
Life Expectancy				-0.020 (0.038)		
Education (All)				-0.026 (0.074)		
Constant	-1.411 (0.989)	6.637 (4.802)	-1.346 (1.035)	-2.426 (2.633)	-3.346** (1.401)	-0.553 (0.868)
Region FE	NO	NO	NO	NO	YES	NO
Exits	42	18	25	45	47	47
Spells	51	22	32	55	57	57
Years of Decline	314	166	193	325	346	346
Log- \mathcal{L}	-59.259	-22.668	-36.422	-65.321	-63.827	-72.090
Pseudo-R ²	0.248	0.277	0.218	0.211	0.263	0.168

Note(s): The standard errors are cluster-robust at the country level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2 addresses the issue of omitted variables by adding groups of variables measuring similar though different aspects of a certain theme, such as trade or finance. All variables, except *de facto* trade openness, enter with the expected sign. The broad patterns are very interesting. Above all, the effect of fractionalization is extremely robust in all specifications and varies only within a narrow band. A one percentage point increase

in fractionalization is estimated to prolong the decline phase by about 1-2%. Further, the coefficient of executive constraints has a negative sign throughout and is often significant. Most specifications imply that a one point improvement in executive constraints leads to 11-18% reduction in the duration of the decline phase. The coefficient becomes small when we control for financial variables, but this is the result of losing roughly half the sample. The standard error blows up when controlling for export characteristics, though the coefficient remains similar. We find that sustaining either a sophisticated export mix or a complex financial system requires well-developed institutions (the correlations are 0.56 and 0.62, respectively). These complex features of modern economies can be characterized as outcomes of institutional development (Acemoglu et al., 2003), so we do not interpret them as a lack of robustness.

Column (1) shows that *de jure* trade openness is negatively associated with the expected duration, while *de facto* openness has a significant positive effect. The opposing signs may seem puzzling at first, but *de facto* openness could be capturing international contagion or excessive import (export) reliance, especially once “institutionalized” openness is controlled for. The coefficients and standard errors for institutions and fractionalization are unaffected. Column (2) illustrates that the effects of executive constraints and fractionalization are robust to the inclusion of export sophistication, the share of manufactured exports, and export diversification. The standard error for executive constraints increases due to the diminished sample size but the coefficient remains in the usual range. Turning to the financial variables, column (3) reveals that including private credit and financial depth reduces the coefficient of executive constraints but also the sample size substantially. The partial effects of these variables are statistically very weak. Column (4) illustrates that neither life expectancy, nor infant mortality, nor schooling have clear effects on the expected duration. We take a different approach to addressing the issue of omitted variables in column (5) by including region dummies to account for regional heterogeneity that is otherwise not captured by the observed covariates. Both the coefficients and their standard errors on institutions and fractionalization are within the usual range, providing further support for the assertion that the effect of institutions is reasonably robust. Column (6) is our preferred specification. This model captures most of the effects we are interested in and uses nearly all available spells. Taken together, the regressions in Table 2 show that the effect of fractionalization is very robust and the effect of institutions is only strongly affected by financial depth and private credit – two measures that we consider to be observed institutional outcomes, and that come at the cost of much data loss.

Intriguingly, the coefficient of initial GDP per capita has a positive sign, implying the counterintuitive result that higher initial GDP prolongs declines. Yet, as the single-variable specifications in Table E-1 of the Appendix show, the simple correlation between initial GDP and the duration of declines is negative, though insignificantly so (p -value = 0.39). The coefficient of initial GDP and its standard error are very unstable. Across all specifications, the implied elasticities of the duration with respect to initial GDP range from near zero to about one, depending on the sample size and the controls. While we keep initial GDP in the models so as to not spuriously assign its effect to institutions, we only treat it as a control variable and do not attempt to give its effect a causal interpretation. The robustness section also shows that the statistical significance of initial GDP strongly depends on the choice of baseline hazard.

Table E-3 in the Appendix assesses the effects of adding each control separately to our preferred specification and reveals some additional findings. Complementing the results

of Desmet et al. (2012), we find that when controlling for both the disaggregate and aggregate measures of fractionalization, the aggregate measure (ELF1) has no effect on the duration of declines. Several of the variables that passed the selection step have effects that are not robust in a multivariate setting. The coefficients of the share of manufactured exports, export diversification, and inequality point in the expected direction but are insignificant by a large margin. Life expectancy, infant mortality and years of schooling are insignificant and do not affect the partial effects of institutions or fractionalization. Otherwise, the same pattern as in Table 2 emerges. The coefficient of institutions only becomes insignificant when the sample size is diminished substantially.

The estimated effects of political institutions are economically meaningful. In the log-normal model, mean and median duration are both estimated by the exponentiated linear prediction allowing straightforward predictions of the expected duration. Conditionally on having entered a slump, a country with the lowest score on the executive constraints measure is expected to decline for about 9.1 years, while a country with the highest score is expected to decline only for about 3.1 years. The mean of executive constraints in the estimation sample is about 2.4, implying a duration of 7.1 years.

Next we examine if there is an interaction effect between institutions and fractionalization. The rationale for this hypothesis is simple. Given a political economy in which ethnic tension challenges the ability of political actors to take coordinated action, more cohesive institutions may help to overcome this vulnerability. Countries with a high degree of ethnic fractionalization may require strong institutions just to compensate. Conversely, countries with a greater degree of ethnic homogeneity may make do with less developed institutions to achieve a similar degree of social coordination. This hypothesis is a less restrictive variant of the idea that there is a multiplicative effect between social conflict and institutions in response to external shocks (Rodrik, 1999).

Table 3 reports the corresponding results. In order to ease the interpretation, we subtract the sample average of the institutions and fractionalization variables from their observed values before estimating each model. We denote the demeaned variables by \overline{INS}_0 and $\overline{ELF15}$. This has the following effect. If either one of the two variables is at its mean, then the interaction term is zero and the only relevant coefficient is the non-interacted variant. As a result, the coefficient of the executive constraints variable directly measures the effect of institutions at the average level of fractionalization, and *vice versa*. For values other than the mean, the coefficient on the interaction term needs to be taken into account.

There is considerable evidence of an interaction effect. In the same specifications where we find a robust effect of institutions, we also find a significant interaction effect between executive constraints and ethno-linguistic fractionalization. The partial effect of one variable at the mean of the other is at least as significant as in the corresponding specification without an interaction effect. The interaction term is negative and significant at the 5%-level in all perturbations but columns (2) and (3). In column (2) there is simply not enough data to estimate this effect, whereas in column (3) the sign and size of the interaction coefficient is actually very similar to other estimations. Since our earlier preferred specification is nested in column (6), testing the null that the interaction term is zero is equivalent to a test that this model fits the data better. A likelihood ratio test also prefers the interaction model and the pseudo- R^2 improves from 0.168 to 0.197.

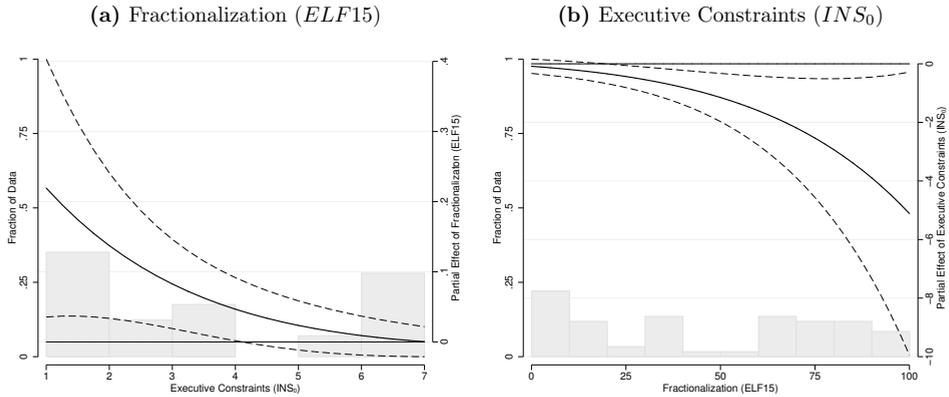
Figure 4 illustrates that the effects estimated in the interaction model are both economically and statistically significant. It plots the predicted semi-elasticities of the expected duration with respect to one variable of the interaction term over representative

Table 3 – Interaction effects of institutions and fractionalization

VARIABLES	(1) $\ln \bar{t}$	(2) $\ln \bar{t}$	(3) $\ln \bar{t}$	(4) $\ln \bar{t}$	(5) $\ln \bar{t}$	(6) $\ln \bar{t}$
Executive Constraints (\widehat{INS}_0)	-0.366*** (0.098)	-0.137 (0.126)	-0.091 (0.101)	-0.280*** (0.090)	-0.217** (0.087)	-0.288*** (0.080)
Fractionalization ($\widehat{ELF15}$)	0.016*** (0.004)	0.016*** (0.006)	0.019*** (0.004)	0.017*** (0.004)	0.017*** (0.004)	0.018*** (0.004)
$\widehat{INS}_0 \times \widehat{ELF15}$	-0.005*** (0.002)	0.000 (0.002)	-0.003 (0.002)	-0.004** (0.002)	-0.003** (0.001)	-0.004*** (0.001)
Initial ln GDP per capita	0.275** (0.123)	1.007*** (0.293)	0.238* (0.135)	0.533*** (0.175)	0.439*** (0.163)	0.198* (0.106)
Real US Interest Rate	0.068 (0.042)	0.064 (0.072)	0.100* (0.054)	0.084* (0.043)	0.076* (0.042)	0.098** (0.047)
Trade Openness (de jure)	-0.690** (0.298)					
Trade Openness (de facto)	0.015*** (0.005)					
Manufactures (% Exports)		-0.008 (0.012)				
Export Diversification		-0.004 (0.011)				
Export Sophistication		-1.503* (0.784)				
Private Credit			-0.008 (0.009)			
Financial Depth			-0.008 (0.011)			
Infant Mortality				0.001 (0.007)		
Life Expectancy				-0.020 (0.030)		
Education (All)				-0.045 (0.070)		
Constant	-1.122 (0.910)	6.940 (4.896)	0.213 (1.247)	-1.425 (2.066)	-2.552* (1.362)	0.025 (0.872)
Region FE	NO	NO	NO	NO	YES	NO
Exits	42	18	25	45	47	47
Spells	51	22	32	55	57	57
Years of Decline	314	166	193	325	346	346
Log- \mathcal{L}	-54.001	-22.648	-35.514	-62.746	-61.681	-69.540
Pseudo-R ²	0.315	0.278	0.238	0.243	0.288	0.197

Note(s): The standard errors are cluster-robust at the country level. *** p<0.01, ** p<0.05, * p<0.1

values of the other, including a 95% confidence interval. Three points stand out: 1) the effect of executive constraints clearly depends on fractionalization (and *vice versa*), 2) both partial effects are significant over most of the distribution, and 3) both partial effects consistently have the expected sign. In the background, Figure 4 also shows histograms of the sample data. Executive constraints scores cover the entire range from 1 to 7, and ethno-linguistic fractionalization ranges from near-zero (0.07%) to near-total heterogeneity (96%). The predictions cover a wide range of the durations. At the average score of executive constraints, a country with the highest (lowest) degree of ethnic heterogeneity is expected to decline for about 12.6 years (2.1 years). Hence, it would be difficult to understand the effects of institutions without considering fractionalization. Stronger institutions also have the potential to overcome the adverse effects of high levels of ethnic fractionalization. At the 75th percentile of ethnic heterogeneity ($ELF15 = 89.7$), a country with the highest (lowest) score of executive constraints is expected to decline for about 1.8 years (18.3 years).

Figure 4 – Partial Effects in Interaction Model

Note(s): The partial effects are based on column (6) in Table 3 and are computed over the entire range of the variable on the horizontal axis while keeping all other regressors at their means. The dashed lines are upper and lower 95% confidence limits.

One way to interpret the effects of fractionalization and institutions is through the lens of the delayed stabilizations literature (Alesina and Drazen, 1991). When (ethnic) groups engage in a ‘war of attrition’ over the burden of reform and are uncertain about how the reform benefits all other groups (hence their willingness to bear the costs), then policy reform is delayed until the weakest group concedes. The expected time until stabilization occurs is expected to increase with the number of groups involved in the decision-making process and the veto points they possess, so that the adjustment speed depends on the political system (Spolaore, 2004). However, this interpretation does not explain the strong interaction effect between executive constraints and ethnic heterogeneity very well.

We propose a theoretical mechanism whereby groups facing a crisis have to decide on a policy response under uncertainty about post-crisis outcomes. When the executive is unconstrained, then some groups may have an incentive to delay cooperation by fear of boosting the strength of the independent executive and its power to expropriate them in the aftermath of the crisis. If groups can fortify their position through blocking agreement on policy, then such a mechanism could generate the observed interaction.¹¹

Robustness

We briefly illustrate that our main conclusions are unaffected by the choice of the baseline hazard, extending the sample and adding unobserved heterogeneity, the exclusion of influential regions, and different ways of accounting for recurrent spells. In addition, we also analyze the depth of the decline and the average rate of contraction.

Table 4 tackles the issues of functional form and model selection. We report our preferred specification in the first column and then show estimates based on five alternative forms of the hazard function. Column (2) uses a log-logistic hazard instead of the log-normal shape. The estimated shape parameter ($\ln \gamma$) is negative, implying that the hazard is first increasing then decreasing as in the log-normal model. This

¹¹In Chapter 4 we formalize this mechanism and provide a theory of delayed responses to crises.

Table 4 – Robustness: functional form

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Log-normal $\ln \hat{t}$	Log-logistic $\ln \hat{t}$	Exponential $\ln \hat{t}$	Weibull $\ln \hat{t}$	Gompertz $\ln \hat{t}$	Cox PH $\ln \hat{t}$
	Coefficients			Hazard Ratios ($H_0 : HR = 1$)		
Executive Constraints (INS_0)	-0.178*** (0.058)	-0.185*** (0.067)	1.229*** (0.074)	1.263*** (0.089)	1.222*** (0.071)	1.221*** (0.082)
Fractionalization ($ELF15$)	0.016*** (0.004)	0.016*** (0.005)	0.978*** (0.005)	0.974*** (0.007)	0.979*** (0.005)	0.979*** (0.006)
Initial ln GDP per capita	0.197* (0.106)	0.235** (0.112)	0.787 (0.119)	0.765 (0.146)	0.786* (0.113)	0.768 (0.137)
Real US Interest Rate	0.087* (0.048)	0.084* (0.051)	0.947 (0.058)	0.928 (0.061)	0.949 (0.057)	0.951 (0.064)
$\ln \sigma$ (Log-normal)	-0.065 (0.093)					
$\ln \gamma$ (Log-logistic)		-0.580*** (0.105)				
$\ln p$ (Weibull)				1.219** (0.107)		
γ (Gompertz)					0.985 (0.030)	
Constant	-0.553 (0.868)	-0.856 (0.901)	1.830 (2.432)	1.723 (2.884)	1.928 (2.448)	
Exits	47	47	47	47	47	47
Spells	57	57	57	57	57	57
Years of Decline	346	346	346	346	346	346
Log- \mathcal{L}	-72.090	-73.286	-75.295	-73.940	-75.192	-143.142
AIC	156.180	158.571	160.590	159.879	162.384	294.285
Pseudo-R ²	0.168	0.164	0.208	0.210	0.160	0.088

Note(s): The standard errors are cluster-robust at the country level. *** p<0.01, ** p<0.05, * p<0.1.

lends itself to the following interpretation. In the first few years of a decline, some countries are able to recover quickly. However, the longer the decline lasts, the more economic fundamentals deteriorate making it increasingly difficult to enter a recovery. Columns (3) to (6) have a different interpretation. We no longer report coefficients but instead *hazard ratios*, since these models are proportional hazards (PH) models by nature. They are interpreted as follows. A hazard ratio greater than one implies a higher instantaneous probability of exiting the decline. A hazard ratio smaller than one implies a lower instantaneous probability of exiting the decline. Column (3) is the exponential or constant hazard model. The results remain very similar (given the altered interpretation), but the log-likelihood decreases somewhat. Column (4) uses a Weibull parameterization which allows for monotonically increasing or decreasing hazard rates. This model also has a shape parameter (p) which allows testing for whether the rate increases, decreases or is constant. The estimate suggests that the baseline hazard is increasing over time. In contrast, the Gompertz model in column (5) also suggests a shape that is monotonically decreasing ($\gamma < 0$). Among these parametric models, the AIC is lowest for the log-normal distribution; that is, our preferred model fits the data best. In column (6), we specify a semi-parametric Cox model which does not restrict the shape of the baseline hazard. The Cox model also suggests that the probability of exiting a spell first increases very briefly and then decreases monotonically. However, the imposed proportional hazard restriction comes at a great cost in terms of fit.

The previous section has already shown that the findings are robust to the inclusion of regional fixed effects. Column (1) in Table 5 goes two steps further: it extends the

Table 5 – Robustness: sample, heterogeneity, dropping regions, and multiple failures

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Extended ln \hat{t}	No SSA ln \hat{t}	No MNA ln \hat{t}	No LAC ln \hat{t}	No multiple ln \hat{t}	PWP multiple ln \hat{t}
Executive Constraints (<i>INS₀</i>)	-0.178*** (0.049)	-0.159*** (0.055)	-0.199*** (0.071)	-0.189** (0.074)	-0.199*** (0.064)	1.263*** (0.096)
Fractionalization (<i>ELF15</i>)	0.011*** (0.004)	0.005 (0.004)	0.018*** (0.005)	0.016*** (0.005)	0.015*** (0.004)	0.981*** (0.006)
Initial ln GDP per capita	0.301*** (0.105)	0.358*** (0.101)	0.263* (0.139)	0.179 (0.115)	0.196* (0.111)	0.759 (0.137)
Real US Interest Rate	0.091** (0.036)	0.103** (0.042)	0.066 (0.048)	0.090 (0.065)	0.086 (0.060)	0.940 (0.067)
Constant	-1.415 (0.904)	-1.843** (0.936)	-1.055 (1.063)	-0.393 (0.870)	-0.405 (0.900)	
VCE	–	cluster	cluster	cluster	cluster	cluster
Frailties	shared	–	–	–	–	–
Strata	–	–	–	–	–	spell #
Exits	70	40	43	34	40	47
Spells	82	44	50	43	50	57
Years of Decline	466	178	294	271	312	346
Log- \mathcal{L}	-109.168	-50.584	-64.255	-54.643	-63.715	-123.435
Pseudo-R ²	0.094	0.151	0.163	0.172	0.162	0.095

Note(s): *** p<0.01, ** p<0.05, * p<0.1.

sample by running the break search algorithm with a significance level of 20% to detect more episodes, and it includes country-level effects. We now identify 82 spells in total, out of which 70 are completed. Each country now also has a so-called gamma distributed frailty; the duration analysis equivalent of random effects in linear models. As usual, they are assumed to be uncorrelated with any of the covariates (which is unlikely to hold in practice). Our results are robust to these two modifications; the coefficients of interest hardly change. Interestingly, there is only some evidence in favor of unobserved effects altogether. A Likelihood Ratio test for the presence of shared frailties fails to reject the null at 5% significance ($p = 0.054$).

To examine if results are driven by specific regions, we re-estimate our preferred model several times, each time removing one of the three regions with the longest spells. Column (2) drops all episodes in Sub-Saharan Africa (SSA) and reveals an interesting additional finding. While the coefficient of fractionalization (ELF15) is very robust in the previous models, its size and significance is clearly driven by African observations. Without those, the coefficient keeps the same sign but shrinks substantially and becomes insignificant at conventional levels. The coefficient of institutions remains large and significant. Since Sub-Saharan Africa has the greatest ethno-linguistic heterogeneity of all regions, this result comes as no surprise and does imply we are only estimating an “Africa effect”.¹² Nevertheless, the interaction model proposed earlier may be particularly relevant to understanding the effects of institutions and fractionalization in Africa. The parameter estimates are not sensitive to excluding either the entire Middle East and North Africa (MNA) or all of Latin America and the Caribbean (LAC).

Until now, we assumed that multiple spells of the same type are interchangeable. The last two columns of Table 5 investigate if this relatively strong form of conditional independence is a reasonable assumption. Column (5) shows that our findings are robust to excluding all spells other than the first, which rules out any dependency across

¹²Also see Chapter 4, where we explore this issue further. This is a common issue in the literature, e.g. see Sturm and de Haan (2015) who link various fractionalization measures to income inequality.

recurrent spells. The coefficient of executive constraints becomes even larger and the effect of fractionalization is virtually unchanged. Column (6) takes a different approach and specifies a conditional risk set model or stratified Cox model due to Prentice et al. (1981, PWP). The model accounts for ordering of the events but assumes that a subject cannot experience another event until the previous event has occurred. This is a natural assumption, as – by definition – a country cannot exit a second decline phase before having left the first. The results (hazard ratios) are qualitatively similar.

Table 6 – Robustness: Average rate of decline and total depth

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	\bar{g}_i	\bar{g}_i	$\bar{t}_D \times \bar{g}_i$	$\bar{t}_D \times \bar{g}_i$	$\bar{t}_D \times \bar{g}_i$	$\bar{t}_D \times \bar{g}_i$
Executive Constraints (INS_0)	0.003 (0.002)	0.001 (0.002)	0.034*** (0.010)	0.005 (0.010)	0.027** (0.010)	0.002 (0.012)
Fractionalization ($ELF15$)	-0.000 (0.000)	-0.000* (0.000)	-0.003*** (0.001)	-0.005*** (0.001)	-0.003*** (0.001)	-0.005*** (0.001)
Interaction ($INS_0 \times ELF15$)		0.000* (0.000)		0.001*** (0.000)		0.001*** (0.000)
Initial ln GDP per capita	-0.001 (0.006)	-0.001 (0.006)	-0.022 (0.024)	-0.022 (0.024)	-0.052 (0.033)	-0.045 (0.035)
Real US Interest Rate	-0.002 (0.002)	-0.002 (0.002)	-0.001 (0.009)	-0.006 (0.008)	0.001 (0.009)	-0.003 (0.008)
Constant	-0.030 (0.049)	-0.022 (0.049)	0.007 (0.197)	0.098 (0.186)	0.308 (0.310)	0.352 (0.310)
Region FE	NO	NO	NO	NO	YES	YES
Spells	57	57	57	57	57	57
Log- \mathcal{L}	106.795	107.408	12.267	14.939	15.987	18.209
R ²	0.009	0.011	0.289	0.340	0.295	0.333

Note(s): Standard errors are cluster-robust at the country level. *** p<0.01, ** p<0.05, * p<0.1.

Last but not least, we construct a simple test to examine whether the effects of institutions and ethnic cleavages on the depth of a slump run solely through the duration process or if they also affect the rate of contraction. Since the depth of the decline is the product of the estimated duration and the average rate of contraction, we can define $\bar{g}_i \equiv (y_{i,\hat{t}_{min}} - y_{i0}) / (\hat{t}_{min} - t_0) \equiv (y_{i,\hat{t}_{min}} - y_{i0}) / \bar{t}_D$ as the average rate of decline and $\bar{t}_D \times \bar{g}_i \equiv y_{i,\hat{t}_{min}} - y_{i0}$ as the overall depth of the decline. Here y_{it} is still the log of GDP per capita in country i at time t , and \bar{g}_i is by construction negative. We have already analyzed the duration of declines. Now we run OLS regressions explaining the rate of contraction (\bar{g}_i) and depth ($\bar{t}_D \times \bar{g}_i$) to isolate the channel through which the previously estimated effects run.

Column (1) in Table 6 shows that we find no evidence in favor of an effect of either executive constraints or fractionalization on the average rate of decline. Column (2) confirms the absence of robust effects for the specification with an interaction term. In fact, in both columns most coefficients are insignificant and estimated to be virtually zero. The variables are not mean-centered in this case. The remaining columns of Table 6 examine the depth of the slump. Columns (3) and (4) illustrate that we now recover the previously estimated effects (even with similar significance levels and, naturally, with reversed signs since \bar{g}_i is negative). These results are robust to the inclusion of region fixed effects as shown in columns (5) and (6). The interaction models explain about 33–34% of the variation in depth, highlighting the relevance of the estimated effects. Together with the duration analysis and the analysis of the average rate of decline, this leads us to conclude that a) political institutions and ethnic heterogeneity have robust

effects on the overall depth of slumps, and b) these effects run primarily through the duration of the decline phase and not the rate of contraction.

5 Concluding remarks

This chapter makes several contributions to a burgeoning literature on growth episodes and structural breaks in growth performances. We show that a restricted structural change approach, as in Papell and Prodan (2012), works well as an inferential method for identifying slumps, big recessions or growth collapses in a large sample of countries. We find a substantial number of slumps in developing and developed countries alike, suggesting that severe downward volatility is an ubiquitous phenomenon in the post-war period. There is systematic evidence of weak political institutions before slumps hit and positive institutional change during and in the immediate aftermath of slumps. Our interpretation of this stylized fact is that, while institutions may cause growth, volatility can also contribute to endogenous institutional change. Severe economic crises bring about what we may call creative political destruction and raise the pressure for institutional reform in a very broad sense. Our key finding is that the duration of declines depends on institutions and also particularly strongly on ethno-linguistic fractionalization. An important extension is that the effect of executive constraints is non-linear and depends on the level of fractionalization, an interaction which is not fully captured by current theories of policy reform and delay. We also show that effects of political institutions and ethnic cleavages on the depth of declines run through the duration.

As a whole, our results lend broad support to political economy theories emphasizing the respective roles of institutions and ethnic cleavages. Effective coordination and responses to slumps are hampered by a high degree of social tension as captured by ethno-linguistic fractionalization. However, particularly strong political institutions can put in place coordination mechanisms that are able to contain or resolve these conflicts within the institutional framework. At the same time, our findings imply that in less ethnically fragmented societies political institutions are less important as a determinant of the length of declines.

These results do not suggest that sound macroeconomic policies as such do not matter, though they provide some indication that these policies may be secondary to more fundamental factors. We need more research on which specific institutional features allow better policies to emerge. While the previous literature has stressed the role of positive growth spurts, we show that slumps matter a lot and that the decline phase can last very long in some cases. We think a comparison of the relative effects of slumps versus accelerations on long-run GDP levels would be an interesting extension of our findings.

Bibliography

- Acemoglu, D. and S. Johnson (2005). Unbundling institutions. *Journal of Political Economy* 113(5), 949–95.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5), 1369–1401.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics* 117(4), 1231–1294.
- Acemoglu, D., S. Johnson, J. A. Robinson, and Y. Thaicharoen (2003). Institutional causes, macroeconomic symptoms: volatility, crises and growth. *Journal of Monetary Economics* 50(1), 49–123.
- Acemoglu, D. and J. A. Robinson (2006). *Economic Origins of Democracy and Dictatorship*. Cambridge University Press.
- Aguiar, M. and G. Gopinath (2007). Emerging market business cycles: The cycle is the trend. *Journal of Political Economy* 115(1), 69–102.
- Alesina, A. and A. Drazen (1991). Why are stabilizations delayed? *American Economic Review* 81(5), 1170–1188.
- Alesina, A. and E. L. Ferrara (2005). Ethnic diversity and economic performance. *Journal of Economic Literature* 43(3), 762–800.
- Ashraf, Q. and O. Galor (2013). The “Out of Africa” hypothesis, human genetic diversity, and comparative economic development. *American Economic Review* 103(1), 1–46.
- Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory* 13, 315–352.
- Bai, J. (1999). Likelihood ratio tests for multiple structural changes. *Journal of Econometrics* 91(2), 299–323.
- Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.
- Bai, J. and P. Perron (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18(1), 1–22.
- Barro, R. J. and J.-W. Lee (2010, April). A new data set of educational attainment in the world, 1950–2010. Working Paper 15902, National Bureau of Economic Research.
- Beck, T., A. Demirgüç-Kunt, and R. Levine (2010). Financial institutions and markets across countries and over time: The updated financial development and structure database. *World Bank Economic Review* 24(1), 77–92.
- Ben-David, D. and D. H. Papell (1995). The great wars, the great crash, and steady state growth: Some new evidence about an old stylized fact. *Journal of Monetary Economics* 36(3), 453–475.
- Berg, A., J. D. Ostry, and J. Zettelmeyer (2012). What makes growth sustained? *Journal of Development Economics* 98(2), 149–166.
- Besley, T. and T. Persson (2011). The logic of political violence. *Quarterly Journal of Economics* 126(3), 1411–1445.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics* 29(2), 238–249.
- Cerra, V. and S. C. Saxena (2008). Growth dynamics: The myth of economic recovery. *American Economic Review* 98(1), 439–57.
- Chinn, M. D. and H. Ito (2006). What matters for financial development? capital controls, institutions, and interactions. *Journal of Development Economics* 81(1), 163–192.
- Desmet, K., I. Ortuno-Ortín, and R. Wacziarg (2012). The political economy of linguistic cleavages. *Journal of Development Economics* 97(2), 322–338.
- Diebold, F. X. and C. Chen (1996). Testing structural stability with endogenous breakpoint a size comparison of analytic and bootstrap procedures. *Journal of Econometrics* 70(1), 221–241.
- Easterly, W., M. Kremer, L. Pritchett, and L. Summers (1993). Good policy or good luck? *Journal of Monetary Economics* 32(3), 459–483.
- Eichengreen, B., A. K. Rose, and C. Wyplosz (1995). Exchange market mayhem: the antecedents and aftermath of speculative attacks. *Economic Policy* 10(21), 249–312.
- Fernandez, R. and D. Rodrik (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *American Economic Review* 81(5), 1146–1155.

- Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002). Armed conflict 1946-2001: A new dataset. *Journal of Peace Research* 39(5), 615–637.
- Goemans, H. E., K. S. Gleditsch, and G. Chiozza (2009). Introducing Archigos: A dataset of political leaders. *Journal of Peace Research* 46(2), 269–283.
- Gourinchas, P.-O. and M. Obstfeld (2012). Stories of the twentieth century for the twenty-first. *American Economic Journal: Macroeconomics* 4(1), 226–265.
- Hansen, B. E. (2000). Testing for structural change in conditional models. *Journal of Econometrics* 97(1), 93–115.
- Hausmann, R., J. Hwang, and D. Rodrik (2007). What you export matters. *Journal of Economic Growth* 12(1), 1–25.
- Hausmann, R., L. Pritchett, and D. Rodrik (2005). Growth accelerations. *Journal of Economic Growth* 10(4), 303–329.
- Hausmann, R., F. Rodriguez, and R. Wagner (2008). Growth collapses. In C. Reinhart, C. Vegh, and A. Velasco (Eds.), *Money, Crises and Transition*, pp. 376–428. Cambridge, Mass.: MIT Press.
- Jerzmanowski, M. (2006). Empirics of hills, plateaus, mountains and plains: A markov-switching approach to growth. *Journal of Development Economics* 81(2), 357–385.
- Jones, B. F. and B. A. Olken (2008). The anatomy of start-stop growth. *Review of Economics and Statistics* 90(3), 582–587.
- Kalbfleisch, J. D. and R. L. Prentice (2002). *The statistical analysis of failure time data* (2nd ed.). New York: John Wiley.
- Lancaster, T. (1990). *The econometric analysis of transition data*. Cambridge University Press.
- Lane, P. R. and G. M. Milesi-Ferretti (2007). The external wealth of nations mark ii: Revised and extended estimates of foreign assets and liabilities, 1970-2004. *Journal of International Economics* 73(2), 223–250.
- MacKinnon, J. (2009). Bootstrap hypothesis testing. In D. Belsley and E. Kontoghiorghes (Eds.), *Handbook of Computational Econometrics*, Chapter 6, pp. 183–213. John Wiley & Sons, Ltd.
- Mobarak, A. (2005). Democracy, volatility, and economic development. *Review of Economics and Statistics* 87(2), 348–361.
- Nelson, C. R. and C. R. Plosser (1982). Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics* 10(2), 139–162.
- North, D. C., J. Wallis, and B. Weingast (2009). *Violence and Social Orders A Conceptual Framework for Interpreting Recorded Human History*. Cambridge University Press.
- Papell, D. H. and R. Prodan (2012). The Statistical Behavior of GDP after Financial Crises and Severe Recessions. *The B.E. Journal of Macroeconomics* 12(3), 1–31.
- Prentice, R. L., B. J. Williams, and A. V. Peterson (1981). On the regression analysis of multivariate failure time data. *Biometrika* 68(2), 373–379.
- Pritchett, L. (2000). Understanding patterns of economic growth: searching for hills among plateaus, mountains, and plains. *World Bank Economic Review* 14(2), 221–250.
- Prodan, R. (2008). Potential pitfalls in determining multiple structural changes with an application to purchasing power parity. *Journal of Business & Economic Statistics* 26(1), 50–65.
- Ramey, G. and V. A. Ramey (1995). Cross-country evidence on the link between volatility and growth. *American Economic Review* 85(5), 1138–1151.
- Rodrik, D. (1999). Where did all the growth go? External shocks, social conflict, and growth collapses. *Journal of Economic Growth* 4(4), 385–412.
- Solt, F. (2009). Standardizing the World Income Inequality Database. *Social Science Quarterly* 90(2), 231–242. SWIID Version 3.0, July 2010.
- Spolaore, E. (2004). Adjustments in different government systems. *Economics & Politics* 16(2), 117–146.
- Sturm, J.-E. and J. de Haan (2015). Income inequality, capitalism and ethno-linguistic fractionalization. *American Economic Review: Papers and Proceedings*. forthcoming.
- Wacziarg, R. and K. H. Welch (2008). Trade liberalization and growth: New evidence. *World Bank Economic Review* 22(2), 187–231.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). The MIT press.
- Zivot, E. and D. W. K. Andrews (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics* 10(3), 251–270.

Appendix A: Estimation of structural breaks

Sequential procedure for testing and dating breaks

The procedure described here is a modification of Bai's (1997) sequential likelihood ratio tests for structural change – see also the extensions in Bai and Perron (1998) and in Bai (1999). We make an important simplifying assumption, namely, that all output series are regime-wise trend-stationary. Verifying this assumption is beyond the scope of this chapter, as testing for unit roots in the presence of structural breaks (with sufficient power and size) is still contested territory and our output series have only a moderate time dimension ($T < 60$ years). We implement the sequential procedure in six steps.

1. Determine the optimal $AR(p)$ trend model using the Bayesian information criterion to adjust for serial correlation up to a maximum lag count (p_{max}). We set $p_{max} = 4$.
2. Specify the partial structural change model:

$$y_t = \alpha + \beta t + \gamma_0 \mathbf{1}(t > tb_1) + \gamma_1 (t - tb_1) \mathbf{1}(t > tb_1) + \gamma_2 (t - tb_2) \mathbf{1}(t > tb_2) + \sum_{i=1}^p \delta_i y_{t-i} + \epsilon_t$$

where y_t is the log of GDP per capita in year t , tb_i are the possible break dates, $\mathbf{1}(\cdot)$ is an indicator function, and p is the lag order as determined by the optimal $AR(p)$ model. We require that $tb_2 \geq tb_1 + h$ for $h = 4$. In other words, the period between two successive breaks making up the same episode is at minimum 4 years.

3. Define trimming parameter τ , where typically $\tau \in [0.05, 0.25]$. The breaks are in the ranges $tb_1 \in [\tau T, (1 - \tau)T - h]$ and $tb_2 \in [\tau T + h, (1 - \tau)T]$. We set $\tau = 0.05$. Let Λ_τ denote the set of all possible episodes $[tb_1, tb_2] \subset [\tau T, (1 - \tau)T]$.¹³
4. Compute the sup- W test statistic of the null of no break versus at least one break ($\mathbb{H}_0 : \gamma_0 = \gamma_1 = \gamma_2 = 0$). The supremum is taken over all episodes in Λ_τ with a positive estimate of β and a non-positive estimate of γ_0 :

$$\sup_{[tb_1, tb_2] \in \Lambda_\tau} W(tb_1, tb_2) = \sup_{[tb_1, tb_2] \in \Lambda_\tau} \left(\frac{T - K}{3} \right) \frac{SSR^r - SSR^u}{SSR^u}$$

where K is the number of parameters, SSR^r denotes the sum of squared residuals from a regression imposing \mathbb{H}_0 , and SSR^u the sum of squared residuals from a regression imposing only $\beta > 0$ and $\gamma_0 \leq 0$.

5. The critical value and empirical p -value of sup- W statistic is bootstrapped.¹⁴
6. If the sup- W statistic is significant at the desired level, the remaining sample is split into two new sub-samples from the beginning to the first break and from the second break to the end, then the procedure restarts at (4) using the estimated AR-order from before. If the bootstrapped sup- W^* test fails to reject in each sub-sample, or the sub-samples are too small ($T \leq 20$), then the procedure stops.

¹³For simplicity of exposition, we suppress an additional index running over the sub-samples (defined in Step 6). T refers to the number of observations of the currently active sample. The notation neglects the discontinuity of actual observation times.

¹⁴In finite samples comparable asymptotic tests often have poor size and power (see Prodan, 2008).

Bootstrapping the sup-Wald statistic

There have been several suggestions on how to best bootstrap structural change tests in particular or other popular time series tests in general. For example, Hansen (2000) suggests employing a fixed-design bootstrap allowing for non-stationarity, lagged dependent variables and conditional heteroskedasticity. MacKinnon (2009), on the contrary, shows that the recursive bootstrap of Diebold and Chen (1996) gives results superior to most other bootstrap types (fixed-parameter, sieve, pairs, block, double block) as well as the asymptotic test in a simple application of an AR(1) model with an endogenous break. Papell and Prodan (2012) also favor a recursive bootstrap though they do not compare it to other methods. We use a recursive bootstrap similar to Diebold and Chen (1996). Comparing methods systematically is beyond the scope of this chapter.¹⁵ In line with usual notation, we denote all bootstrap quantities with the superscript ** . The bootstrap procedure is as follows.

1. Specify the optimal break model under the \mathbb{H}_0 of no structural breaks in the specified sample using the BIC as before and obtain the residuals:

$$\hat{e}_t = y_t - \hat{\alpha} - \hat{\beta}t - \sum_{i=1}^p \hat{\delta}_i y_{t-i}.$$

2. Draw new residuals: $\hat{e}_t^* = u_t$, with $u_t \sim \text{i.i.d. } \mathcal{N}(0, \hat{\sigma}_e^2)$.
3. Construct a bootstrap sample of equal size as the original sample:

$$y_t^* = \hat{\alpha} + \hat{\beta}t + \sum_{i=1}^p \hat{\delta}_i y_{t-i}^* + \hat{e}_t^*, \quad \forall t = 1 + p, \dots, T$$

where y_{t-i}^* is the observed y_{t-i} only in the case of a *fixed-design* bootstrap, otherwise y_t^* must be constructed *recursively* (conditional on p observed initial values).

4. Rerun the break search algorithm on the bootstrap series $\{y_t^*\}$, including determination of the optimal AR(p) model, and compute bootstrapped test statistics $\sup_{[tb_1^*, tb_2^*] \in \Lambda_\tau} W_j^*$, where j indexes the current bootstrap iteration.
5. Repeat from Step (2) until $j = B$, where B is the total number of bootstrap replications. We set $B = 1000$.
6. The bootstrap p -value (\hat{p}^*) is obtained by counting the proportion of the estimated bootstrap test statistics that are greater than the originally calculated test statistic:

$$\hat{p}^* = \frac{1}{B} \sum_{j=1}^B \mathbf{1} \left(\sup_{[tb_1^*, tb_2^*] \in \Lambda_\tau} W_j^* > \sup_{[tb_1, tb_2] \in \Lambda_\tau} W(tb_1, tb_2) \right).$$

The critical value is the $(1 - \alpha^s)B^{th}$ largest bootstrapped sup- W^* statistic, where α^s is the desired significance level (10% throughout the text, unless otherwise noted).

¹⁵We use a parametric recursive bootstrap, but informally compared the results to other techniques. Hansen's fixed-design bootstrap generates (too) many questionable slumps and the Wild bootstrap rejects (too) often. Residual and parametric bootstraps give similar results.

Appendix B: List of episodes

Table B-1 – Global Parameters

Data:	PWT	Max AR (p_{max}):	4
Sample start:	1950	Bootstrap replications:	1000
Sample end:	2008	Bootstrap errors:	parametric
Trimming (τ):	0.05	Bootstrap type:	recursive
Min. tb_i distance (h):	4	Bootstrap significance (α^s):	0.1

Table B-2 – Estimated and Filtered Breaks with Troughs: 58 Episodes*

Code	T_0	\hat{tb}_1	\hat{t}_{min}	\hat{tb}_2	T	Sup- W	Critical W	p-value	Drop (%)	Duration	c
ALB	1970	1990	1991	2002	2008	18.5	13.6	0.007	-15.32	1	0
ARE	1986	1990	1999	2002	2008	29.1	14.5	0.003	-10.90	9	0
AUS	1950	1954	1957	1966	2008	8.3	8.7	0.064	-0.72	3	0
AUS	1967	1989	1991	1998	2008	10.1	10.7	0.059	-2.29	2	0
BDI	1960	1971	1972	1988	2008	9.9	11.3	0.089	-3.23	1	0
BEL	1950	1957	1958	1973	2008	12.8	12.1	0.029	-2.24	1	0
BGR	1970	1988	1997	1997	2008	16.3	12.8	0.010	-23.79	9	0
BHR	1970	1980	1987	1986	2008	14.4	11.0	0.010	-44.12	7	1
BRA	1950	1980	1983	2003	2008	12.5	12.3	0.043	-14.60	3	0
CAF	1960	1978	2005	2005	2008	8.3	8.7	0.060	-46.38	27	1
CHE	1950	1974	1975	1978	2008	10.7	10.6	0.047	-7.87	1	0
CHL	1951	1953	1954	1972	1973	12.0	8.5	0.017	-9.06	1	0
CHL	1951	1974	1975	1979	1980	13.3	10.8	0.021	-16.50	1	0
CHL	1951	1981	1983	1995	2008	12.6	11.4	0.025	-21.22	2	0
CHN	1952	1960	1962	1977	2008	13.9	12.9	0.029	-23.71	2	0
CMR	1960	1986	1995	1990	2008	12.0	12.3	0.055	-40.46	9	1
COG	1960	1974	1977	1982	2008	11.9	12.5	0.069	-21.35	3	0
CRI	1950	1955	1956	1963	1979	11.4	11.3	0.048	-4.39	1	0
CRI	1950	1980	1982	2002	2008	17.2	10.6	0.002	-17.47	2	0
CUB	1970	1988	1993	1995	2008	11.4	12.5	0.072	-34.70	5	0
CYP	1950	1973	1975	1977	2008	15.5	9.7	0.001	-31.40	2	0
CYP	1978	1990	1991	1995	2008	11.6	14.6	0.098	-10.19	1	0
DNK	1950	1954	1955	1965	2008	12.9	11.7	0.022	-1.56	1	0
DZA	1960	1984	1994	1996	2008	10.9	8.2	0.013	-14.09	10	0
ETH	1950	1972	1992	1993	2008	11.5	10.2	0.020	-30.68	20	0
FIN	1950	1989	1993	2006	2008	10.6	10.8	0.057	-16.34	4	0
GAB	1960	1976	1987	1997	2008	10.6	11.2	0.062	-50.56	11	1
GMB	1960	1982	1998	2002	2008	16.4	11.2	0.006	-25.33	16	0
GRC	1951	1973	1974	1994	2008	17.9	11.6	0.003	-6.92	1	0
GTM	1950	1980	1988	1984	2008	15.1	12.3	0.015	-19.14	8	0
HUN	1970	1990	1992	2004	2008	15.6	13.5	0.018	-10.56	2	0
IDN	1960	1997	1999	2001	2008	13.5	10.6	0.013	-17.49	2	0
IRN	1955	1976	1981	1980	2008	15.9	11.6	0.004	-56.78	5	1
IRQ	1970	1990	2003	1994	2008	9.1	8.9	0.046	-66.43	13	1
JPN	1950	1973	1974	1990	2008	13.5	13.4	0.050	-2.85	1	0
MEX	1950	1981	1988	1995	2008	11.9	11.0	0.038	-17.03	7	0
MNG	1970	1990	1993	2003	2008	46.5	11.7	0.000	-41.81	3	0
MOZ	1960	1981	1986	1995	2008	12.6	12.0	0.037	-24.99	5	0
MYS	1955	1984	1986	1993	2008	9.1	10.5	0.093	-7.47	2	0
NPL	1960	1979	1980	2000	2008	10.6	8.9	0.025	-5.33	1	0
NZL	1950	1974	1978	1992	2008	9.9	10.5	0.070	-9.03	4	0

Continued on next page

Table B-2 – *Continued from previous page*

Code	T_0	\hat{t}_{b_1}	\hat{t}_{min}	\hat{t}_{b_2}	T	Sup- W	Critical W	p-value	Drop (%)	Duration	c
OMN	1970	1979	1980	1985	2008	12.4	9.0	0.007	-21.61	1	0
PER	1950	1958	1959	1966	1976	11.9	9.3	0.022	-6.91	1	0
PER	1950	1977	1992	1992	2008	11.0	10.3	0.037	-29.30	15	0
PHL	1950	1983	1985	2003	2008	12.8	10.2	0.007	-16.78	2	0
POL	1970	1979	1982	1993	2008	13.8	12.1	0.027	-22.55	3	0
PRY	1980	1989	2002	2002	2008	8.8	8.8	0.049	-14.24	13	1
RWA	1960	1993	1994	1997	2008	18.0	7.9	0.001	-45.38	1	0
SAU	1986	1992	1999	2002	2008	14.6	13.3	0.039	-18.75	7	0
SLE	1961	1995	1999	2006	2008	14.2	11.1	0.011	-41.65	4	1
SLV	1950	1978	1983	1987	2008	18.2	10.2	0.002	-25.82	5	0
TGO	1960	1979	2008	1989	2008	9.6	10.1	0.065	-53.60	29	1
THA	1950	1996	1998	2003	2008	10.7	7.8	0.003	-14.17	2	0
TTO	1950	1961	1963	1969	1981	16.8	14.9	0.020	-0.78	2	0
TTO	1950	1982	1993	2006	2008	12.4	12.6	0.054	-28.96	11	0
UGA	1950	1977	1986	1987	2008	11.6	10.5	0.029	-30.27	9	0
USA	1950	1957	1958	1966	2008	8.7	9.3	0.075	-2.51	1	0
ZMB	1955	1968	2001	2000	2008	15.0	10.9	0.007	-68.99	33	1

* Out of a total of 70 episodes identified by the sequential algorithm, 12 are invalid slumps. The invalid episodes are [country code (spell number)]: AUT (1), AUT (2), CHN (1), FIN (1), HKG (1), IRN (1), MRT (1), PRY (1), TZA (1).

Appendix C: Data sources and summary statistics

Table C-1 – Summary Statistics: break date to trough

VARIABLE	Mean	Std. Dev.	$N \times T$	Source
<i>Institutions, Politics & Conflict</i>				
Polity Score	-1.90	6.99	346	Polity IV
Democracy	2.73	3.61	330	Polity IV
Autocracy	4.69	3.74	330	Polity IV
Executive Recruitment	4.92	2.27	330	Polity IV
Executive Constraints	3.18	2.28	330	Polity IV
Political Competition	4.11	3.38	330	Polity IV
Regime Duration	18.14	22.70	347	Polity IV
Negative Regime Change	0.01	0.12	347	Polity IV
Positive Regime Change	0.10	0.29	347	Polity IV
Corruption (ICRG)	2.63	1.10	193	ICRG
Fractionalization (ELF1)	18.36	18.69	348	Desmet et al. (2012)
Fractionalization (ELF15)	63.68	30.71	348	Desmet et al. (2012)
Inequality (Gini)	45.83	11.65	192	Solt (2009)
Leader Exit	0.39	0.49	344	Goemans et al. (2009)
War/Conflict (major)	0.12	0.33	348	Gleditsch et al. (2002)
War/Conflict (any)	0.24	0.43	348	Gleditsch et al. (2002)
<i>Macro I: Prices, Trade & Exports</i>				
Inflation ($\ln(1 + \delta)$)	22.89	43.97	292	WDI/IFS
RER Undervalue	0.07	0.54	348	PWT 7.0
Current Account Balance (% of GDP)	-3.98	6.70	254	WDI
Δ Terms of Trade	-4.11	17.72	224	WDI/IFS
Manufactures (% of Exports)	22.65	24.27	264	WITS/ COMTRADE
Trade Openness (de facto)	67.85	37.43	348	PWT 7.0
Trade Openness (de jure)	0.23	0.42	306	Wacziarg and Welch (2008)
Export Sophistication	8.43	0.42	234	Hausmann et al. (2007)
Δ Export Sophistication	1.48	7.44	233	Hausmann et al. (2007)
Export Diversification	65.91	24.58	264	WITS/ COMTRADE
<i>Macro II: Finance</i>				
Capital Account Openness	-0.49	1.28	304	Chinn and Ito (2006)
Financial Integration	115.30	88.18	309	Lane and Milesi-Ferretti (2007)
Financial Depth	32.35	18.68	245	Beck et al. (2010)
Financial Development	68.40	22.18	271	Beck et al. (2010)
Private Credit (% of GDP)	26.25	23.53	248	Beck et al. (2010)
FDI Liabilities (% of GDP)	15.11	15.66	309	Lane and Milesi-Ferretti (2007)
External Debt Liabilities (% of GDP)	65.22	59.18	309	Lane and Milesi-Ferretti (2007)
External Leverage ^a	165.29	327.09	307	Lane and Milesi-Ferretti (2007)
<i>Other Determinants</i>				
Initial ln GDP per capita ^b	8.20	1.21	348	PWT 7.0
Real US Interest Rate ^c	1.90	2.44	348	FRED
Infant Mortality ^d	73.37	40.23	348	World Population Prospects
Life Expectancy ^d	58.63	10.55	348	World Population Prospects
Telephones (per 100 people)	5.24	9.78	312	WDI
Education (primary)	3.14	1.71	327	Barro and Lee (2010)
Education (secondary)	1.12	0.83	327	Barro and Lee (2010)
Education (all)	4.44	2.47	327	Barro and Lee (2010)

^a Following Gourinchas and Obstfeld (2012), external leverage is $l_i = (\tau + A_i/Y_i)(\tau + NA_i/Y_i + E_{ij}/Y_i)^{-1}$, where τ is the market value of assets to output (set to 3) and j is the rest of the world, A_i/Y_i is assets over GDP, NA_i/Y_i is net foreign assets over GDP and E_{ij}/Y_i equity over GDP. The ratio is always > 0 if $NA_i > -300$, this condition is not satisfied in very few cases; we set these missing.

^b Initial refers to ln GDP per capita at t_0 , that is, the last year before the slump.

^c Deflated three months treasury bill rate.

^d Converted into annual data by interpolation. If the average is for the years 1950-55, we assume it is reached in the 1952 and linearly interpolate to the middle of the next group (1957), and so on. The data is from the 2010 edition of the World Population Prospects (medium-fertility variant).

Appendix D: Duration Method

Log-normal Accelerated Failure Time (AFT) models

Given the model $\ln(\tilde{t}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \epsilon$, log-normality implies the following relationships. Setting all covariates zero, the expected survival time is $E[\ln \tilde{t} | \mathbf{x} = \mathbf{0}] = \beta_0$. Hence, the baseline survival and hazard functions are

$$S_0(\tilde{t}) = 1 - \Phi((\ln \tilde{t} - \beta_0)\sigma^{-1}) \quad \text{and} \quad \lambda_0(\tilde{t}) = \frac{\phi((\ln \tilde{t} - \beta_0)\sigma^{-1})}{(1 - \Phi((\ln \tilde{t} - \beta_0)\sigma^{-1}))\sigma\tilde{t}}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal pdf and cdf, respectively.

Including (time-invariant) covariates is equivalent to scaling the baseline survival functions. The conditional survival curve is defined as $S(\tilde{t}|\mathbf{x}) = S_0(\tilde{t}) (\exp(-\mathbf{x}'\boldsymbol{\beta})\tilde{t})$. This implies $S(\tilde{t}|\mathbf{x}) = 1 - \Phi((\ln \tilde{t} - (\beta_0 + \mathbf{x}'\boldsymbol{\beta}))\sigma^{-1})$; that is, the intercept can be absorbed into $\boldsymbol{\beta}$. The density and cumulative probability functions are defined implicitly.¹⁶

Time-varying covariates introduce two complications. First, the hazard rate at each unit of analysis time \tilde{t} is not independent from previous realizations of the time-varying covariates. Second, the covariates must be *strictly exogenous*, as otherwise feedback may occur from the duration to future realizations of the covariates. Following Lancaster (1990) and Kalbfleisch and Prentice (2002) these issues can be formalized as follows. For time-varying covariates $\mathbf{x}(\tilde{t})$, let $\mathbf{x}^H(\tilde{t})$ denote the covariate path up until time \tilde{t} , so that $\mathbf{x}^H(\tilde{t}) \equiv \{\mathbf{x}(u), 0 \leq u \leq \tilde{t}\}$ for all $\tilde{t} \geq 0$, then the conditional hazard function is:

$$\lambda(\tilde{t}|\mathbf{x}^H) = \lim_{d\tilde{t} \rightarrow 0} \frac{\Pr(\tilde{t} \leq \tilde{T} < \tilde{t} + d\tilde{t} \mid \tilde{T} \geq \tilde{t}, \mathbf{x}^H(\tilde{t} + d\tilde{t}))}{d\tilde{t}}$$

Lancaster (1990, pp. 26–30) and Kalbfleisch and Prentice (2002, p. 196) define strict exogeneity as $\Pr(\mathbf{x}^H(\tilde{t}) \mid \mathbf{x}^H(u), \tilde{T} \geq u) = \Pr(\mathbf{x}^H(\tilde{t}) \mid \mathbf{x}^H(u), \tilde{T} = u)$ for all $0 < u \leq \tilde{t}$. The condition states that the future path of the time-varying covariate is not affected by the event occurring at present.

We can now derive the partial likelihood.¹⁷ Suppose we know the event occurs at \tilde{t}_i , the likelihood contribution of an observation i at time $j = \tilde{t}_i$ then is $\mathcal{L}_i = S(j)\lambda(j)$. The likelihood contribution of an observation that has not failed at time j , so that $j < \tilde{t}_i$, is just the probability of survival until j : $\mathcal{L}_i = S(j)$. Hence, right-censoring is essentially nothing else than an observation at analysis time j that is still in the sample but has not yet failed and thus extends easily to (exogenous) time-varying covariates.

Using the notation for grouped data from Wooldridge (2010, p. 1016), the log-likelihood of the log-normal model with time-varying covariates can be expressed as:

$$\ln \mathcal{L}(\boldsymbol{\beta}, \sigma) = \sum_{i=1}^N \left[\sum_{j=1}^{\tilde{t}_i-1} \ln \alpha_j(\mathbf{x}'_{ij}\boldsymbol{\beta}, \sigma) + (1 - c_i) \ln \left(1 - \alpha_{\tilde{t}_i}(\mathbf{x}'_{i\tilde{t}_i}\boldsymbol{\beta}, \sigma) \right) \right]$$

where $\alpha_j(\cdot) = \exp[-\int_{\alpha_j-1}^{\alpha_j} \lambda(s, \cdot) ds]$ measures survival over the given interval and c_i indicates if observation i is censored. The inner sum (first term) is the probability of survival until $\tilde{t}_i - 1$ and the second term is the conditional probability of failure at \tilde{t}_i .

¹⁶It follows that an expression for the hazard function conditional on the covariates is $\lambda(\tilde{t}|\mathbf{x}) = \lambda_0(\tilde{t} \exp(-\mathbf{x}'\boldsymbol{\beta})) \exp(-\mathbf{x}'\boldsymbol{\beta})$; these hazards are not proportional.

¹⁷This does not apply to frailty models where the likelihoods are more involved.

Appendix E: Variable selection and robustness

Table E-1 – Base Models

	Coefficient	SE	p-value	Exits	Spells	Years	$\log \mathcal{L}$
Constant Only	1.346	0.180	0.00	48	58	348	-87.86
Initial ln GDP per capita	-0.124	0.144	0.39	48	58	348	-87.44
Real US Interest Rate	0.096	0.047	0.04	48	58	348	-86.55

Note(s): All models include a constant. The standard errors are cluster-robust at the country level.

Table E-2 – Variable Selection

	Coefficient	SE	p-value	Exits	Spells	Years	$\log \mathcal{L}$
Inflation ($\ln(1 + \delta)$)	-0.002	0.004	0.68	38	45	234	-64.62
RER Underval	-0.144	0.333	0.67	48	58	348	-86.13
Trade Openness (de jure)	-1.019	0.304	0.00	43	52	316	-74.89
Trade Openness (de facto)	0.014	0.005	0.00	48	58	348	-80.82
Current Account Balance	-0.000	0.027	1.00	27	34	222	-47.79
Manufactures (% Exports)	-0.015	0.007	0.04	24	31	236	-42.26
Δ Terms of Trade	-0.007	0.017	0.68	24	27	164	-36.63
Export Diversification	-0.015	0.008	0.07	24	31	236	-42.29
Export Sophistication	-2.131	0.574	0.00	28	34	241	-45.86
Capital Account Openness	-0.016	0.125	0.90	32	41	275	-59.63
Financial Integration	0.000	0.003	0.88	35	43	271	-61.67
Financial Depth	-0.015	0.005	0.00	26	33	195	-44.81
Financial Development	0.006	0.009	0.55	31	39	266	-57.87
External Debt Liabilities	0.000	0.007	0.98	35	43	271	-61.69
External Leverage	0.003	0.014	0.82	35	43	271	-61.64
FDI Liabilities	-0.004	0.018	0.83	35	43	271	-61.67
Private Credit	-0.013	0.004	0.00	28	35	198	-47.09
Polity IV Score	-0.064	0.018	0.00	47	57	346	-80.27
Democracy Score	-0.118	0.032	0.00	47	57	346	-80.43
Autocracy Score	0.127	0.038	0.00	47	57	346	-80.57
Executive Recruitment	-0.163	0.057	0.00	47	57	346	-81.90
Executive Constraints (INS_0)	-0.218	0.065	0.00	47	57	346	-79.70
Political Competition	-0.122	0.038	0.00	47	57	346	-80.89
Regime Durability	-0.002	0.005	0.72	47	57	346	-84.96
Corruption (ICRG)	-0.486	0.141	0.00	14	18	98	-19.29
Fractionalization ($ELF1$)	0.018	0.007	0.01	48	58	348	-83.82
Fractionalization ($ELF15$)	0.018	0.004	0.00	48	58	348	-78.07
Inequality (Gini)	0.045	0.023	0.05	22	27	137	-34.73
Leader Exit	0.424	0.360	0.24	47	57	346	-84.18
War/Conflict (major)	0.179	0.875	0.84	48	58	348	-86.19
War/Conflict (any)	0.469	0.553	0.40	48	58	348	-85.73
Infant Mortality	0.014	0.006	0.02	48	58	348	-83.18
Life Expectancy	-0.060	0.030	0.05	48	58	348	-83.03
Education (Primary)	-0.356	0.096	0.00	46	56	327	-76.39
Education (Secondary)	-0.448	0.165	0.01	46	56	327	-79.76
Education (All)	-0.254	0.063	0.00	46	56	327	-76.17
Telephones per capita	-0.021	0.014	0.13	30	38	257	-52.57

Note(s): All models also include initial GDP per capita, the real US interest rate, and a constant. The standard errors are cluster-robust at the country level.

Table E-3 – Preferred Specification: Variable-by-Variable Models

VARIABLES	(1) $\ln \bar{t}$	(2) $\ln \bar{t}$	(3) $\ln \bar{t}$	(4) $\ln \bar{t}$	(5) $\ln \bar{t}$	(6) $\ln \bar{t}$	(7) $\ln \bar{t}$	(8) $\ln \bar{t}$	(9) $\ln \bar{t}$	(10) $\ln \bar{t}$	(11) $\ln \bar{t}$	(12) $\ln \bar{t}$
Executive Constraints (INS_0)	-0.176** (0.070)	-0.170*** (0.058)	-0.137 (0.085)	-0.178** (0.085)	-0.058 (0.078)	-0.024 (0.078)	-0.012 (0.072)	-0.134 (0.114)	-0.178*** (0.058)	-0.155*** (0.060)	-0.156*** (0.060)	-0.179*** (0.064)
Fractionalization ($ELF15$)	0.019*** (0.004)	0.012*** (0.006)	0.010* (0.006)	0.008 (0.006)	0.018*** (0.004)	0.019*** (0.005)	0.019*** (0.005)	0.014*** (0.005)	0.016*** (0.006)	0.016*** (0.004)	0.016*** (0.004)	0.015*** (0.004)
Initial \ln GDP per capita	0.402*** (0.118)	0.061 (0.121)	0.135 (0.156)	0.120 (0.155)	0.887*** (0.234)	0.286** (0.131)	0.282** (0.132)	0.520** (0.244)	0.188 (0.114)	0.374** (0.170)	0.399** (0.184)	0.401*** (0.124)
Real US Interest Rate	0.066 (0.046)	0.085* (0.047)	0.122 (0.083)	0.130 (0.082)	0.090* (0.049)	0.086 (0.054)	0.106** (0.051)	-0.019 (0.077)	0.088* (0.048)	0.099** (0.046)	0.097** (0.046)	0.070 (0.045)
Trade Openness (de jure)	-0.807** (0.329)											
Trade Openness (de facto)		0.009** (0.005)										
Manufactures (% Exports)			-0.010 (0.009)									
Export Diversification				-0.005 (0.010)								
Export Sophistication					-2.151*** (0.479)							
Financial Depth						-0.018** (0.007)						
Private Credit							-0.017*** (0.005)					
Inequality (Gini)								0.022 (0.022)				
Fractionalization ($ELF1$)									0.002 (0.009)			
Infant Mortality										0.007 (0.005)		
Life Expectancy											-0.029 (0.023)	
Education (All)												-0.088 (0.069)
Constant	-2.080** (0.966)	0.232 (0.944)	0.440 (1.444)	0.902 (1.595)	11.805*** (2.948)	-1.176 (1.030)	-1.324 (1.039)	-4.164** (2.004)	-0.479 (0.930)	-2.538 (1.724)	-0.542 (0.936)	-1.786** (0.884)
Exits	42	47	24	24	28	25	27	22	47	47	47	45
Spells	51	57	31	31	34	32	34	27	57	57	57	55
Years of Decline	314	346	236	236	241	193	196	137	346	346	346	325
Log- \mathcal{L}	-61.318	-70.200	-39.236	-39.655	-39.083	-36.857	-38.267	-31.316	-72.051	-71.252	-71.192	-66.142
Pseudo- R^2	0.222	0.189	0.120	0.111	0.245	0.209	0.230	0.138	0.168	0.177	0.178	0.202

Note(s): The standard errors are cluster-robust at the country level. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 4

Ethnic divisions, political institutions and the duration of declines

*A political economy theory of delayed recovery**

Abstract

This chapter analyzes the duration of large economic declines and provides a theory of delayed recovery. First, we develop a formal political economy model that illustrates a simple mechanism of how weak constraints on the political executive can lead to longer declines in ethnically heterogeneous countries. The model shows how uncertain post-recovery incomes and a ‘winner-take-all’ threshold effect create a commitment problem rendering a cooperative equilibrium inaccessible. Holding out can benefit groups by reducing the threshold effects in subsequent periods, thus limiting the remaining uncertainty. Placing strong constraints on the executive solves this commitment problem by reducing the uncertainty from the threshold effects, which brings about cooperation earlier on. Second, we then test several empirical predictions from the model using standard data on linguistic heterogeneity and more detailed data on ethnic power configurations. We find that the partial correlations are consistent with the proposed theory. The effect of executive constraints on the length of declines is very large in heterogeneous countries, but practically disappears in ethnically homogeneous societies. The adverse effect of heterogeneity is driven by the number of groups; increasing political concentration works in the opposite direction.

Keywords: economic crises, delayed recovery, political economy

JEL Classification: E6, O43, J15

*This chapter is based on Bluhm, R. and K. Thomsson. 2015. Ethnic cleavages, institutions and the duration of economic slumps, MERIT Working Papers 003, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

1 Introduction

Why do economic declines in Sub-Saharan Africa and some parts of the globe last so much longer than in others, say, Western Europe and North America? We propose a novel answer to this question which links ethnic heterogeneity and the powers of the political executive to the failure to agree on a policy response to the shock, even when the policy is economically effective and socially desirable.

The main contribution of this chapter is to provide a theory of how ethnic diversity and political institutions interact during economic declines and to test the implications of this theory. We focus on the process of policy formulation during the decline phase of a slump and illustrate a simple mechanism of how weak constraints on the political executive can lead to longer declines in ethnically heterogeneous countries. The key issue we highlight is a commitment problem among winners and losers of the recovery process. Uncertainty about post-recovery incomes and a ‘winner-take-all’ threshold effect caused by imperfect political institutions can lead to delayed cooperation. Next, we assess the model empirically. We examine the central predictions of the model using both standard data on linguistic heterogeneity and a more detailed data set on the level of access to executive power enjoyed by the prevailing ethno-political groups.

The main theoretical results are threefold. First, delayed cooperation can occur in equilibrium. Weak constraints on the executive act as a political friction in ethnically diverse countries that can lead to large social inefficiencies. Second, stronger political institutions can entirely resolve this issue and bring about cooperation early on. Third, the commitment problem is getting worse when the number of groups increases. We also derive an additional result which takes the relative size (strength) of the groups into account. We then show empirically that the partial correlations are consistent with the theory. The effect of executive constraints on the length of declines is very large in heterogeneous countries, but muted in ethnically homogeneous societies. This finding is robust to the different data sets, as well as region and decade dummies. We also show that greater political concentration shortens declines and, *vice versa*, that a more even distribution of political power across groups increases delay. An important policy implication is that political institutions can contain the adversarial element of ethnic diversity and thus play a critical role in heterogeneous countries.

The rest of the chapter is organized as follows. Section 2 briefly outlines the empirical motivation: an interaction effect of ethnic diversity and political institutions on the duration of economic declines. In Section 3, we argue that the rich literatures on ethnicity and political institutions in economics and political science have not yet offered a thorough explanation for this interaction. In Section 4, we outline our model of how ethnic fractionalization and weak constraints on the executive can lead to delayed cooperation. In Section 5, we discuss the data, the empirical strategy and the main empirical results. Section 6 concludes.

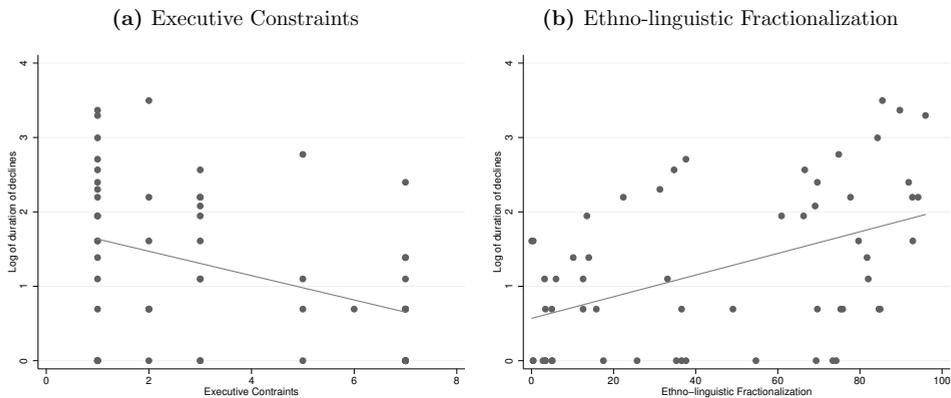
2 Empirical motivation

This chapter is motivated by a growing empirical literature which has established that economic growth is often not steady but instead characterized by different growth regimes. It is long known that the correlation of growth rates across decades is low (Easterly et al., 1993). A key finding of the newer growth episodes literature is that growth accelerations are triggered by a variety of factors but are difficult to sustain (Hausmann et al., 2005;

Berg et al., 2012). Several years of positive growth can easily be followed by long and deep slumps. Such negative shocks can wipe out previous welfare gains and are often characterized by persistent output loss (Cerra and Saxena, 2008). In light of these findings, it becomes important to understand why some declines last so much longer than others and what factors are associated with longer (or shorter) durations.

In Chapter 3, we discuss the econometric identification of the decline phase of economic slumps and then analyze its duration. We empirically examine if the duration of the decline phase of large economic slumps is, among other factors, shaped by political institutions and ethnic cleavages. In a departure from the previous literature, we specifically focus on the *duration of declines* for three reasons. First, the onset of a slump may be brought about by many factors which are not necessarily related to a country's political institutions or level of social cohesion, but the duration of declines depends on socioeconomic groups agreeing on coordinated responses. Second, the dynamics of recoveries differ a lot from the dynamics of declines (both empirically and theoretically). Third, most of the variation in the overall depth of slumps is due to the duration of the decline phase and not due the rate of contraction.

Figure 1 – Unconditional correlations with the duration of declines



Note(s): The durations are based on the 58 slumps estimated in Chapter 3 for the panel (b), and 57 slumps for panel (a) as we lack scores on the index of executive constraints for Trinidad and Tobago in 1961. No adjustment has been made for censored observations (unfinished declines).

We first show that the duration (in years) until a recovery starts increases with greater degrees of ethnic cleavages, and that it decreases with stronger constraints on the executive. A slump is defined by a trend break or shift in the growth regime with a restricted pattern. The duration of the decline phase is simply the time from the downbreak until the trough. Institutional strength is measured by the constraints placed on the political executive using an index scaled 1 to 7 (least to most constrained). Ethnic heterogeneity is proxied for by an index of linguistic fractionalization (scaled 0 to 100). Figure 1 illustrates the unconditional correlation of the (log) duration of declines with executive constraints (-0.38) and ethno-linguistic fractionalization (0.47).

We also provide evidence of a more subtle pattern: the adverse effect of high ethnic heterogeneity is conditional on the quality of political institutions. A log-normal survival regression of the duration of the decline phase on executive constraints, ethno-linguistic

fractionalization and an interaction term yields the following (the standard errors are in parentheses below the coefficients):¹

$$\widehat{\ln \tilde{t}} = \frac{1.808}{(0.201)} - \frac{0.254}{(0.080)} XCONST_0 + \frac{0.018}{(0.004)} ELF - \frac{0.004}{(0.002)} XCONST_0 \times ELF$$

The estimated effects are statistically significant and qualitatively meaningful. Even though this is a very naïve reduced form model of the duration process, the basic findings are robust to more demanding specifications. We centered the two explanatory variables on their sample mean, so that the interaction term only needs to be considered when both regressors change. *Ceteris paribus*, a unit increase in executive constraints away from the average leads to a 22.4% shortening of the duration until the trough and thus predicts a substantially faster exit from the decline phase. Conversely, a one percentage point increase in ethnic heterogeneity prolongs the duration by about 1.8%. Political institutions seem to moderate the effects of ethnic heterogeneity. At perfect executive constraints the negative effect of diversity is virtually zero, whereas it is very large when the political executive has unlimited powers. We also show that the effect of these two variables on the overall depth of slumps runs through the duration and not the average rate of declines.

In Chapter 3 we are primarily concerned with the econometrics of identifying declines and establishing this stylized fact. The main objective of this chapter is twofold: first, to propose a theory that can generate such an interaction effect, and, second, to then empirically examine additional theoretical predictions using much more detailed data on ethnic groups and their political power.

3 Related literature

Ethnic heterogeneity is a fundamental determinant of prosperity. It is typically associated with low growth (Easterly and Levine, 1997), the undersupply of public goods (Alesina et al., 1999), and civil conflict (Fearon and Laitin, 2003; Esteban and Ray, 2011; Esteban et al., 2012). Ethnicity plays a rampant role in Sub-Saharan Africa where political organization is mostly ethnic (Francois et al., 2012), but diversity has also been linked to inadequate public good provision in US states (Alesina et al., 1999), school quality in Kenya (Miguel and Gugerty, 2005), or excessive deforestation in Indonesia (Alesina et al., 2014). Furthermore, high economic inequality among ethnic groups is associated with regional underdevelopment and political inequality (Alesina et al., 2012).

Heterogeneity is not necessarily a problem and viewed favorably in many literatures. In lower income economies, organizing along ethnic lines may resolve a contracting problem and help to enforce social sanctions within family or kin groups (Bates, 2000). In highly developed economies, the negative effects of heterogeneity are muted, as skill complementarities matter more, or political institutions tame the conflict element inherent in diversity (Alesina and Ferrara, 2005). Ethnic diversity and political institutions also have an impact on one another. On the one hand, the adverse effects of ethnic heterogeneity may only be relevant in weakly institutionalized societies where political leaders often use (or abuse) ethnic and other divisions in their favor (Eifert

¹The estimation is based on 57 episodes using the data from Chapter 3. A subscripted zero means that the variable is fixed at the last year before the slump to avoid endogenous feedback. The standard errors are clustered on the country level to account for repeated spells.

et al., 2010). On the other hand, diversity affects the (endogenous) choice of institutions governing the executive power of such leaders (Aghion et al., 2004).

There is some empirical evidence consistent with the view that ethnicity and political institutions interact. Collier (2000), for example, argues that ethnicity plays no role in democracies but is growth reducing in autocracies and provides evidence along these lines. Easterly (2001) empirically investigates an interaction effect between institutions and ethnicity in determining growth and conflict. However, the precise mechanisms behind how these two jointly determine the length of crises have not been investigated and may explain substantial parts of the robust negative correlation between ethnicity and growth. While there is plenty of anecdotal evidence, we are only aware of a paper by Rodrik (1999) which explicitly considers ethnicity and *negative* growth empirically (and more formally in the working paper version).

The theoretical literature on delayed reform and policy non-adoption offers important insights. Ethnic groups may be engaged in “wars of attrition” over the burden of reform, so that groups are trying to shift the costs of, say, a debt consolidation onto competing groups (Alesina and Drazen, 1991). In these models, agreement on a particular policy is required for stabilization and veto power lies either with groups represented within the executive or an effective parliamentary or non-parliamentary opposition. Stabilization occurs only once one of the groups concedes. Drazen and Grilli (1993) use this set-up to show that crises can be welfare improving by reducing delay and Spolaore (2004) examines the impact of different government systems on the expected time until a stabilization occurs. Alternatively, a socially optimal reform may not be undertaken at all because it is *ex ante* not known to which (ethnic or other political) groups the benefits will accrue (Fernandez and Rodrik, 1991). Labán and Sturzenegger (1994a,b) show that such a model can also generate delay and an endogenous economic deterioration. Both approaches have two key elements in common: 1) *uncertainty* about the expected outcomes, and 2) an *ex ante commitment problem* between (*ex post*) beneficiaries and losers of the reform. However, while instructive, this literature does not explicitly focus on ethnic diversity and constraints on the executive. As a result, it does not capture an interaction between these two factors in determining the length of declines.

This chapter also relates to the veto player literature in political science. These contributions generally find that policy stability is greater, the more numerous the players in the political system that are required to agree (Tsebelis, 1995, 2002). Veto player arguments have been used to explain why governments may not reform during an economic shock (Cox and McCubbins, 1997; Haggard, 2000), but recently Gehlbach and Malesky (2010) turn the argument on its head by demonstrating that (more) veto players weaken the power of special interest groups which encourages wholesale reform. Using a different setup based on the selectorate framework by Bueno de Mesquita et al. (2005), Hicken et al. (2005) stress an alternative mechanism which suggests that accountability of the executive matters in response to exchange-rate devaluations. They conclude that greater checks on the executive do not aid the recovery which stands in sharp contrast to the results developed in this chapter.

The degree of ethnic diversity is endogenous in the (very) long run. Heterogeneity is related to migratory distance from Africa (Ashraf and Galor, 2013), the duration of settlements and the history of the state (Ahlerup and Olsson, 2012), and variation in terrain and land endowments (Michalopoulos, 2012). At the micro-level, people may choose their group affiliation and switch groups depending on how discernible the individual features are which identify group membership (Caselli and Coleman, 2013).

Ethnic markers are usually more salient than other group identifiers which may explain why interest groups organize among ethnic lines to limit “infiltration” by outsiders.

We do not expect ethnic compositions to change fundamentally in the short run (especially in the post-colonial period). However, ethnicity is not always the most prominent political fault line in a society and the degree of access to political power of a particular group varies over time (Posner, 2004). Early empirical studies of the effects of ethnic heterogeneity (e.g. Easterly and Levine, 1997) use data Soviet ethnographers published in 1964, incorporate possibly irrelevant cleavages and do not account for differences in political power. Several later studies use up-to-date data on *linguistic* fragmentation (e.g. Fearon, 2003; Desmet et al., 2012), but still remain confined to the cross-section and disregard political power. Wimmer et al. (2009), as well as Cederman et al. (2010), present a new data set which explicitly aims to remedy this situation. The *Ethnic Power Relations* (EPR) data codes the degree of access to executive power by different groups, focuses on politically relevant groups, and employs a more flexible notion of political division capturing the main fault line in a particular country (such as ethnicity, language, race or religion).² In our empirical part, we use the latest EPR data and contrast the results to more traditional measures of linguistic diversity.

4 Theory

We model group interactions during a slump as a cooperation game where groups decide on whether to formulate a policy response to a crisis that will initiate a recovery. We first focus on the symmetric two group case and then extend the model to allow for unequal sizes and multiple groups.

4.1 Basic setup

We consider a population normalized to unity and split into J equal-sized (ethnic) groups. These J groups constitute the players of the game. Time is discrete and there is an infinite number of periods, indexed by $t = 1, 2, \dots, \infty$. The per-period discount rate is δ . With slight abuse of notation, groups are indexed by $j = 1, 2, \dots, J$, where $J = 2$ for the baseline model considered in this subsection. Each group acts as a single agent and we do not analyze internal coordination issues among individual members within the same group.

Preferences. Group j receives a net income of y_j in period zero. Total initial income in the economy is normalized to unity ($\sum_j y_j = 1$). Utility in each period is $g(y_j)$, where $g(\cdot)$ is increasing in y_j , concave and identical for all groups.

Slumps: decline and recovery. When a slump occurs, output declines by a fixed amount (Δ) in the first period, which affects both groups proportionally, and then remains at that level until both groups cooperate. Total income is now $(1 - \Delta)$ as long as the slumps lasts. Once a decision to cooperate has been reached, we assume that the economy recovers within one period. Groups decide to cooperate or not based on their expected future returns to cooperation.

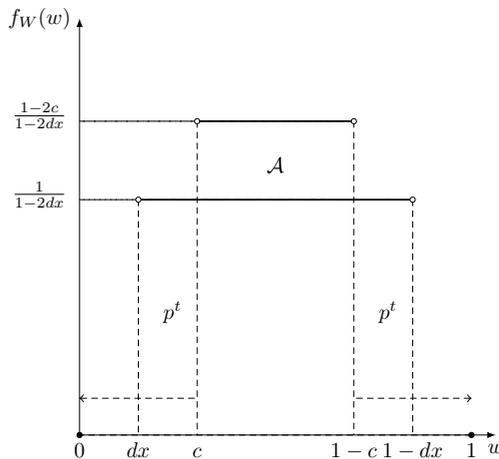
²Rainer and Trebbi (2012) and Francois et al. (2012) extend this approach further and code an ethnic group’s access to power at the ministerial level for 15 African countries.

Slumps: uncertainty. We assume that groups are uncertain about their post-recovery incomes, and their relative positions may change after the slump is over. In the baseline model where $y_1 = y_2 = 1/2$, both groups experience a uniformly distributed shock to their incomes in the first recovery period. For the first group, the shock is $\nu \sim \mathcal{U}[-y_1, 1 - y_1]$, so y_1 is now a counterfactual, and w_1 denotes actual income after the shock; and similarly for y_2 , w_2 . The other group's income undergoes a shock opposite to that of the first group. This implies that a slump hits groups unequally after the recovery, even though neither group assumes they will be hit harder *ex ante*.

Political institutions. We interpret executive constraints as limits on how much one can group gain or lose relative to the other through expropriation, as is common practice in the literature (Besley and Persson, 2011a,b). The intuition is as follows. If a particular group has been sufficiently weakened by the slump, the now stronger group may be in a position to expropriate part or all of the weaker group's income and exclude it from the political process. If the executive is unconstrained, there are no checks on this type of predatory behavior. A complete constraint on how much one group can extract means that no expropriation can occur. A partial constraint implies that expropriation occurs only when one group becomes too weak. The dominant ethnic group controls the executive and shares the spoils from expropriating the weaker group with its members.

To be more precise, we model political institutions in a simple way by including thresholds in the random shock. Boundary outcomes (expropriation) are realized only beyond these thresholds. Let the parameter $c \in [0, 1/2]$ represent the weakness of executive constraints, and let the set $\mathcal{A} = [c, 1 - c]$ be the political safe zone in which there is no expropriation. Once a group falls below c , its income is expropriated (pushed to zero) and the other group gains the remainder. Thus, $1 - 2c$ can be interpreted as the ability of one group to commit to not expropriating the other group; alternatively, $c = 1/2$ can be thought of as the total lack of constraint.

Figure 2 – Threshold effects as constraints on the executive



To fix ideas, we interpret the ‘winner-take-all’ event as political extinction of the weaker group, though it can be understood in a variety of ways. In non-democratic politics, the threshold mechanism symbolizes the potential of some ethnic groups to

exclude other groups from the political process and capture the rents of those that have been excessively weakened by the slump. Alternatively, it may even represent physical extinction due to ethnic conflict. In democratic politics, assuming that ethnic or other identity groups are represented by parties reflecting their interests, it captures the existence of thresholds that allow minorities to block political change (e.g. the filibuster rule used in the U.S. Senate as well as several state legislatures, or the 5% minority threshold used in the German *Bundestag*).

Delay. We assume that groups are able to fortify their position through non-cooperation. This implies that a group can (in part) counterbalance the uncertainty introduced by weak institutions through not cooperating, and thus potentially avoid falling below the threshold. Both groups would cooperate and the recovery would be immediate if there were no gains from delay.

In terms of the model, delay limits how likely it is that a particular group will be expropriated. We assume that delay shrinks the support of the distribution of w_j , such that after a particular number of periods of delay, say d , the support is $w_j \in [dx, 1 - dx]$. The associated parameter x is a measure of how much a group can reduce the risk of expropriation by holding out in each period. This describes a linear process for the probability of landing on either side outside the safe zone ($p^t = c - (t - 1)x$), at each t when the groups can chose to cooperate or delay.³

Figure 2 shows the distribution of w_j and illustrates the relevant regions.

Timing. The following timing summarizes the structure of the game. At $t = 0$, the economy is in its initial state. Output $\sum_j y_j = 1$ is produced and shared equally.

1. At $t = 1$, the slump occurs, and incomes decline to $(1 - \Delta)y_j$. Both groups simultaneously choose to cooperate C or delay D .
2. For all $t > 1$, incomes remain at $(1 - \Delta)y_j$ if both groups did not cooperate in the previous period. They once again simultaneously choose whether to cooperate C or delay D . If, instead, there was cooperation in the previous period, incomes recover within one period, but are subject to a random shock and groups can land outside the political safe zone with twice the probability p^t . After a recovery, each group receives the same payoff as in the first post-recovery period forever.

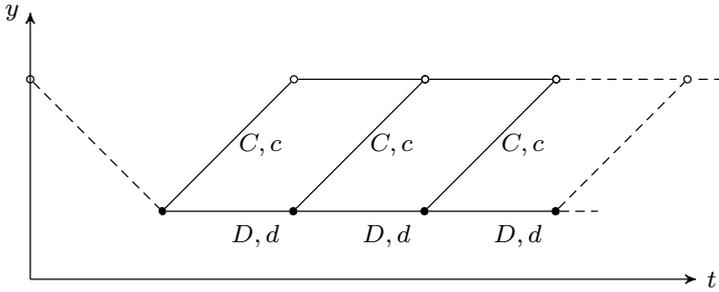
The present discounted value of the lifetime utility for each group is

$$v_j = \sum_{\tau=1}^{\infty} \delta^{\tau-1} \mathbb{E}g(\cdot) \tag{1}$$

where $g(\cdot)$ is $g((1 - \Delta)y_j)$ if the recovery has not yet occurred and $(1 - 2c)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + c(g(0) + g(1))$ otherwise. The discounted utility has two components: 1) if the economy has not recovered, groups are on a delay path, and 2) once the slump is over, they remain on a post-recovery path.

Figure 3 sketches how the economy evolves over time given different choices and presents a stylized view of the process we envision. Note that the action pair (D, d) has the same implication as (C, d) and (D, c) ; that is, cooperation of both groups is required for a recovery to occur.

³Our key results hold for a larger class of p^t -processes.

Figure 3 – A sketch of decisions and timing

The game has a symmetric structure. At each choice node (solid nodes), the comparison between any two adjacent periods always looks alike. The utility from cooperating in a particular period t when the other group cooperates in period t is

$$v_j^t(C, c) = \frac{1}{1-\delta} \left\{ (1-2p^t) \mathbb{E}[g(w_j) | w_j \in \mathcal{A}] + p^t (g(0) + g(1)) \right\} \quad (2)$$

and the utility from cooperating in the next period when the other group cooperates in period t is

$$v_j^t(D, c) = g((1-\Delta)y_j) + \frac{\delta}{1-\delta} \left\{ (1-2p^{t+1}) \mathbb{E}[g(w_j) | w_j \in \mathcal{A}] + p^{t+1} (g(0) + g(1)) \right\}. \quad (3)$$

It is useful to establish the social optimum before we characterize the non-cooperative equilibrium. Our first comment summarizes two key aspects of the planner's solution.

Comment 0. *i) The utilitarian welfare-maximizing outcome involves no delay.*

To see this, note that due to the concavity of the utility function the sum of the group's utilities is maximized when their share is equal. At equal shares, the total welfare from any non-delay path dominates any delay path.

ii) Any outcome with delay is Pareto dominated by some outcome without delay.

To see why this is the case, take any path with delay, give the groups the same shares in every period, but let the recovery happen immediately. In this case, all groups receive more in the period before the recovery than they did with delay, and the same in every period after the recovery.

The intuition behind this comment is straightforward. Given that there are two groups in the economy, a social planner would give both the same shares and avoid delay; only then is their combined utility maximized. Even if these two groups have unequal shares, an immediate recovery is beneficial to both. The social planner is unconstrained, in the sense that the solution involves no uncertainty towards the post-recovery utilities or political boundary effects. This benchmark is particularly interesting when contrasted to the non-cooperative equilibrium of the game, where groups face a trade-off between immediately recovering and falling below the threshold c , or recovering later and reducing future uncertainty.

By comparing the utilities from cooperating in the first period and in the second period it is relatively straightforward to show that delay can occur in equilibrium. Our first result establishes this.

Proposition 1. *There exist parameter values, such that all equilibria involve delay.*

Proof. See Appendix. ■

The proof to the proposition shows that all components that make the immediate cooperation scenario less attractive are conducive to delay. The key issue rendering the cooperative equilibrium inaccessible is the *ex ante* commitment problem among potential winners and losers. Hence, worse institutions, or less ability to commit to not expropriate the loser (larger c), larger gains from holding out (larger x) and a larger value placed on the future (higher δ) make immediate cooperation less likely. Conversely, a larger shock (Δ) makes cooperation more attractive since a (potentially sizable) one period loss is avoided. The concavity of $g(\cdot)$ matters in the sense that it implicitly captures how averse groups are to negative events (falling below c) or how much they value expropriating other groups (landing above c).

Note that the proposition is formally true only in a weak sense; it does not rule out that equilibria with immediate recovery could exist for some parameter values.⁴ Rather, the result should be viewed in light of Comment 0. What Proposition 1 establishes is that for some parameter values *all* equilibria are inefficient and welfare-suboptimal.

While still in the two-group case, we can already highlight an interesting comparison to the homogeneous (one group) case.

Comment 1. *Without heterogeneity, there always exists an equilibrium with immediate recovery.*

Note that if the groups were to pool their resources as one, then all the elements inducing delay – except pure miscoordination – are absent. In other words, we need antagonistic political (ethnic) groups for the proposed mechanism to work, i.e. for the model developed here to provide a theory of why there is delay. A more careful analysis of group asymmetries and multiple groups follows in the model extensions.

To better understand when we are likely to see delay, we now characterize the subgame perfect equilibrium with (the earliest possible) recovery, if such an equilibrium exists. Given the symmetric structure of the game an interior solution exists and the optimal time to recovery can be derived using equations eq. (2) and eq. (3). Our second result summarizes a central insight of the model.

Proposition 2. *Stronger constraints on the executive shorten the time to recovery.*

Proof. See Appendix. ■

⁴There are many “coordination failure” equilibria where neither group cooperates simply because they believe the other group will not. Such equilibria always exist, including an equilibrium with infinite delay. Our analysis, however, is focused on the more interesting scenarios (equilibria) where delay does not happen *only* as a result of this type coordination failure.

The proof shows that the optimal time to recovery is

$$t^* = \frac{g((1 - \Delta)y_j) - \mathbb{E}[g(w_j)|w_j \in \mathcal{A}]}{x\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - (g(0) + g(1))\}} + \frac{c}{x} + \frac{1}{1 - \delta} \quad (4)$$

where the key comparative static result is $\partial t^*/\partial c > 0$.

This proposition says that if institutions are imperfect ($c > 0$), delay is going to be longer than if the groups are able to perfectly commit to not expropriating the losers.⁵ In fact, the weaker the constraints on the executive (larger c), the longer is the expected time to cooperation. Intuitively, either group will find it optimal to delay until a point is reached when the benefit of holding out for an additional period is equal to the benefit of cooperating in this period, where the former may be the period in which all uncertainty regarding the political threshold is resolved. At this point, or the next discrete period, it is optimal to cooperate. Where exactly this point in time occurs depends on the trade off between recovering and potentially falling outside the political safe zone, or recovering later and reducing the remaining uncertainty.

For the remainder, we do not explicitly derive this equilibrium solution. Instead, we focus on the case where all uncertainty is resolved in the next period and compare different scenarios (e.g. perfect and imperfect institutions). We outline such an argument in the next comment.

Comment 2. *The existence of imperfect (weak) institutions makes delay more likely.*⁶

If institutions are perfect ($c = 0$), we have

$$\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] \geq g((1 - \Delta)y_j) \quad (5)$$

and if institutions are imperfect (and $p^t = c$ for all t), we have

$$(1 - 2c)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + c(g(0) + g(1)) \geq g((1 - \Delta)y_j). \quad (6)$$

Inequality (6) is harder to satisfy than inequality (5) under the concavity assumptions imposed on the utility function. Note that this is entirely due to the presence of weak institutions ($c > 0$).

Discussion of the model. The baseline model focuses on several key aspects of the political economy of declines. First, we have modeled group interactions during crises under uncertain post-recovery incomes in a way that highlights that groups are not able to commit to compensating the losers. There are no enforceable contracts where the winners return the (additional) post-recovery gains, which is precisely the role played by strong constraints on the executive. Second, outcomes with delay can occur in equilibrium, and they do *not* coincide with the social optimum or with efficiency. Weak institutions act as a political friction creating potentially large economic inefficiencies. Third, heterogeneity

⁵Again, this is only holds if we rule out equilibria involving immediate cooperation or infinite delay.

⁶Strictly speaking, a probabilistic statement (delay becomes “more likely”) should not be used in this comment, as, for any given set of parameters, there either exists an equilibrium with immediate recovery or not. However, we follow Acemoglu and Robinson (2006) and say that a change in the setup of the model makes a particular outcome “more likely” if it becomes an equilibrium outcome for a larger parameter set.

matters and political groups are assumed to be willing to cooperate once it is optimal to do so. Entrenched distrust would only increase delay.

We abstract from several other features, such as modeling slumps in a more realistic manner (both in terms of the decline and the recovery phase) or the precise nature of the policy response. Clearly, some policies can prolong the decline phase and make recovery more difficult. We also do not differentiate between democratic and autocratic regimes, or examine the impact of particular political constitutions (presidential or parliamentary). The exact form of the boundary events is also left open and could, for example, also represent the exclusion from public goods. Nor do we differentiate between political and economic power. Such specificities are not essential to the main argument, but clearly our contribution is particularly relevant for understanding declines in Africa where political divisions are often ethnic and executive power is shared (Francois et al., 2012).

The mechanism we propose is different than those suggested in the policy reform literature, which has previously focused on shifting the burden of reform (Alesina and Drazen, 1991) and status-quo bias (Fernandez and Rodrik, 1991). While both literatures highlight the importance of *ex ante* uncertainty (either about the costs or benefits of reform), their core focus is not on the role of political institutions in general or executive constraints in particular. The empirical content also differs substantially. For example, Drazen and Grilli (1993) stress that crises help stabilizations and Spolaore (2004) shows that political systems with a strong government (less constrained executive) reform more quickly, whereas we propose that crises coinciding with an unconstrained executive are at the heart of the problem.

4.2 Extensions: asymmetric and multigroup settings

We now briefly sketch two extensions. To extend the model to the asymmetric and J -group cases, we make the following simplifying assumptions. First, we restrict attention to the uncertainty associated with falling below the political threshold. Specifically, we assume that if a group falls within the political safe zone its share of total economic activity will be equal to its pre-recovery share. Second, we use a piecewise linear utility function, in particular:

$$g(y_j) = \begin{cases} y_j & \text{for } y_j > 0 \\ z & \text{otherwise} \end{cases} \quad (7)$$

where $z < 0$. Furthermore, for the case when there are more than two groups, we assume that at most one group can fall outside the political safe zone. We now work with a probability function $p_t(y_j)$, where we only assume $\frac{dp_t(y_j)}{dy_j} < 0$. Finally, our comparative statics will be done for the case where all uncertainty is resolved after one period of delay.

How do changes in political concentration affect the political equilibrium? Intuition may suggest that smaller groups are more afraid of falling out of the political safe zone, implying that greater asymmetry between groups increases the likelihood of delay. However, our theoretical result suggests that the effect of changes in political concentration can go either way. Several things change in the two-group case if the share of an initially weaker group moves closer to an equal allocation, so that the size of the previously more powerful group decreases in return. On the one hand, the emboldened group finds itself further away from the political threshold (c) and hence faces a lower probability of being expropriated ($p_t(y_j)$). In addition, the group also has to forgo more

utility in the delay scenario. Both work in favor of cooperation. On the other hand, the group now has more to lose if it gets expropriated and is thus less likely to cooperate. Without imposing further restrictions, the overall direction of the effect is undetermined and depends on the parameter values. We consider this an empirical issue and return to it in the next section. The following result gives the condition that has to hold for greater symmetry to lead to more delay.

Proposition 3. *A decrease in (political) concentration makes delay more likely, if the following condition holds*

$$\Delta + \frac{1}{1-\delta} \left\{ \frac{dp_1(y_j)}{dy_j} (z - y_j) - p_1(y_j) \right\} < 0. \quad (8)$$

Proof. See Appendix. ■

Using this condition, we can summarize the circumstances that determine the direction of this effect.

Comment 3. *A decrease in concentration is more likely to work in favor of delay, if the shock is smaller, the future is less heavily discounted, the negative consequence of falling outside the political safe zone is greater and the probability of that event is not very responsive to the weaker group's share.*

Up until this point, we assumed that there are only two groups deciding on whether to cooperate or not. The final proposition relaxes this constraint and highlights another key insight of the model with respect to group heterogeneity (assuming symmetric groups).

Proposition 4. *An increase in the number of groups makes delay more likely.*

Proof. See Appendix. ■

Contrary to the more equivocal result in Proposition 3, a larger number of groups decreases the likelihood of cooperation. The proof shows that the condition for immediate cooperation (when all uncertainty is resolved in the next period) boils down to an inequality that decreases in J . The intuition behind this proposition is simple. As the number of groups increases, every group becomes poorer and thus more vulnerable during a slump. Simplifying the model helped to show that this is driven by the uncertainty arising from the lack of executive constraints (which we now no longer denote c but implicitly define through $p_i(y_j)$, since a uniform shock with a threshold does not capture the J -group case).

So far we did not explicitly consider political power or political relevance. Instead, we assumed that all groups start from inside the political safe zone, matter equally for the decision to cooperate, and may only fall into political irrelevance as a consequence of the slump. Keeping the decision mechanism fixed, we now reflect on what this implies for different power (group) configurations. We do so with an eye to the concepts that we can empirically capture in the next section. With this in mind, we summarize the role of political relevance in the last comment as follows.

Comment 4. *More politically relevant groups make delay more likely, while politically irrelevant groups do not matter. Conversely, this implies that if a group dominates or monopolizes the decision making process, delay becomes less likely.*

Here, political relevance refers to being a party to the negotiations and thus being included in this model. Note that this separates the issue of political relevance (and the number of groups) from the strength of constraints on the executive. For example, there may be a dominant group which is unconstrained and thus poses a threat for smaller groups, or there may be a dominant group whose hands are tied so that it cannot expropriate smaller groups. Since these cases are distinct, we should still observe an independent effect of both the number of groups represented in the executive (or different qualitative assessment of their access to executive power), and constraints on the executive *per se*.

5 Empirics and Discussion

Decline spells. Our dependent variable is the duration of the decline segment during deep economic slumps. The identification of the duration of these negative growth spells is not trivial and beyond the scope of this chapter. We only briefly summarize the method here. The details are discussed in Chapter 3 of this dissertation.

The procedure involves several steps. First, we fit a restricted partial structural change model with two breakpoints to each GDP per capita series in the *Penn World Table 7.0*. We impose sign restrictions on the model parameters, so that we only find major economic slumps. Second, we estimate candidates for the endogenous breakpoints and conduct a bootstrap Monte Carlo test of their significance. Third, we keep only breaks that are significant at the 10%-level and run the procedure again on the remaining data (before the first and after the second break) until all breaks have been found or the sample gets too small. Fourth, for each slump, we identify the empirical trough (the lowest point in the GDP series after the beginning of the slump) and then compute the duration of the decline segment (denoted \hat{t}).

This algorithm yields 58 slumps from 1950 to 2008. The basic correlations are as expected. Poorer countries have longer and deeper declines than richer countries; countries in Sub-Saharan Africa have the longest spells, followed by the Middle East and North Africa, and then Latin America. OECD countries experience only few, shallow and short spells.

Measuring institutions. Our core measure of political institutions is the variable *Executive Constraints* from the Polity IV data set. The variable directly measures the degree of institutionalized constraints placed on the political executive. It is coded unity when there is “unlimited executive authority” and seven when there is “executive parity or subordination”; intermediate values represent some constraints. We believe that this variable corresponds well with the parameter c in our model. The Polity IV project has information on executive constraints annually from 1800 (or the year of independence) until 2010. We do not use this wealth of time variation, since political institutions may endogenously respond to the slump. We only rely on the degree of executive constraints in the last year *before* the slump and denote this variable $XCONST_0$.

Measuring heterogeneity. We rely on two data sources to capture very different aspects of ethnic heterogeneity. The first source is a set of measures computed by Desmet et al. (2012) on the basis of the *Ethnologue* data. This data does not measure ethnicity directly but captures *linguistic* diversity. Fearon (2003) shows that linguistic (cultural) diversity coincides well with ethnic heterogeneity in some regions, notably Sub-Saharan Africa, but not so well in others. Together with the *Atlas Narodov Mira* data gathered by Soviet ethnographers in the 1960s, it is a standard source for data on ethnic heterogeneity and considerably more up-to-date than the former. Desmet et al. (2012) compute linguistic diversity at different levels of the language tree to capture the historical depth of ethnic divisions. We only make use of the most disaggregate level, since they also show that current divisions are correlated with economic growth more strongly than historical cleavages. The second data source is the *Ethnic Power Relations* (EPR) data presented in Wimmer et al. (2009), as well as Cederman et al. (2010). The EPR data has several advantages over other measures of linguistic or ethnic diversity, particularly for our application. It provides time series information on the degree of access to *executive* power of ethno-political groups from 1946 to 2010. Contrary to the *Ethnologue* data, it is not restricted to linguistic cleavages existing today. Instead, expert coders identified the most relevant division which may be ethnic, linguistic, racial or religious depending on the country and time period. The data contains information on the power status of each group, so that it allows us to focus on politically relevant groups; that is, groups with some form of representation in the presidency, cabinet, or other senior posts in the administration or army.

Our primary measure of heterogeneity is the commonly used index of ethno-linguistic fractionalization (e.g. Easterly and Levine, 1997). It is defined as

$$ELF_i = 1 - \sum_{j=1}^J \left(\frac{n_{ij}}{N_i} \right)^2 \quad (9)$$

where n_{ij}/N_i is the population share of group j in country i ($j = 1, 2, \dots, J$, n_{ij} is the number of people in group j , and N_i the size of the population in country i). We employ two versions of this index: one computed by Desmet et al. (2012) and one computed using the EPR data (denoted ELF_0). We scale all heterogeneity indices by 100 to give changes on the right hand side a percentage point interpretation.

Another important dimension of diversity is the degree of polarization of a society into two (opposing) groups. The literature on ethnic conflict often stresses that fractionalization and polarization have very different effects (e.g. see Esteban and Ray, 2011). We capture polarization with an index developed by Esteban and Ray (1994):

$$POL_i = k \sum_{j=1}^J \left(\frac{n_{ij}}{N_i} \right)^{1+\alpha} \left(1 - \frac{n_{ij}}{N_i} \right) \quad (10)$$

where $\alpha = 1$ (as they show in an auxiliary theorem) and $k = 4$ to scale the index between zero and one. Again, we use a version computed by Desmet et al. (2012) and one we compute for the EPR data (denoted POL_0).

While the polarization index captures the extent of bimodality of a distribution⁷, it is not a measure of asymmetries (such as the existence of one large and many small groups).

⁷It attains its maximum at a symmetric bimodal distribution.

Table 1 – Definitions of Variables

Symbol	Description	Source and Notes
<i>Dependent Variable</i>		
\bar{t}	Duration of decline segment	From Chapter 3 computed using structural break model with a significance level of 10%. Underlying GDP per capita data is from the Penn World Table 7.0.
<i>Independent Variables</i>		
$XCONST_0$	Constraints on the executive	From Polity IV data. Measures <i>de facto</i> independence of the executive. Scaled from 1 (no constraints) to 7 (fully constrained). Fixed at last year before slump.
ELF	Ethno-linguistic fractionalization	From Desmet et al. (2012), the original source is the Ethnologue data (15 th edition). Cross-section.
ELF_0	Fractionalization of ethno-political groups	From Ethnic Power Relations data version 3.01 (Wimmer et al., 2009). Fixed at last year before slump.
POL	Ethno-linguistic polarization	From Desmet et al. (2012) using the Esteban and Ray (1994) measure with $\alpha = 1$ and $k = 4$. The original source is the Ethnologue data (15 th edition). Cross-section.
POL_0	Ethno-political polarization	Computed using Ethnic Power Relations data version 3.01 (Wimmer et al., 2009) and Esteban and Ray (1994) measure with $\alpha = 1$ and $k = 4$. Fixed at last year before slump.
ELA_0	Asymmetries between ethno-political groups (relative to fractionalization at equal sizes).	Computed using Ethnic Power Relations data version 3.01 (Wimmer et al., 2009). Fixed at last year before slump.
$GROUPS_0$	Number of groups	————
$EGIPGRPS_0$	Number of included groups	————
$EXCLGRPS_0$	Number of excluded groups	————
$DOMPOP_0$	Dominant population (in %)	————
$MONPOP_0$	Monopoly population (in %)	————
<i>Control Variables</i>		
GDP per capita	Log of initial real GDP per capita	From the Penn World Table 7.0. Fixed at last year before slump.
Regional dummies	Africa, Americas, Asia, Europe, Oceania	UN classification. Oceania is base.
Decade dummies	1950s, 1960s, 1970s, 1980s, 1990s, and 2000s.	Coded at beginning of slump. 2000s is base.

To capture these, we propose another simple measure of ethno-linguistic asymmetries:

$$ELA_i = \frac{\sum_{j=1}^J \left(\frac{n_{ij}}{N_i}\right)^2 - \frac{1}{N_i}}{1 - \frac{1}{N_i}} = \frac{N_i}{N_i - 1} \left[\sum_{j=1}^J \left(\frac{n_{ij}}{N_i}\right)^2 - \frac{1}{N_i} \right], \quad \forall N_i > 1 \quad (11)$$

and $ELA_i = 1$ if $N_i = 1$. The ELA index is simply the (normalized) difference between fractionalization with equal shares and observed fractionalization; it's a normalized Herfindahl index. We only compute this index for the EPR data (denoted ELA_0). Desmet et al. (2012) do not use this measure. Recall that for any particular number of groups, the ELF measure attains its maximum at an equal allocation. The global maximum is reached when, in the limit, each person constitutes an ethnic group. Contrary to the polarization or fractionalization measure, the ELA index is zero when the groups are of equal sizes and approaches unity as a single group becomes dominant. For the empirical analysis that follows, using the index of group asymmetries together with the number of groups allows us to analyze the effect of these two components of ethnic heterogeneity separately and investigate the more subtle aspects of the theoretical model.

Table 2 – Summary Statistics

VARIABLES	Obs	Mean	Std. Dev.	Min	Max
$XCONST_0$	57	3.42	2.47	1.00	7.00
ELF	58	45.39	33.71	0.07	95.98
ELF_0	57	36.00	25.71	0.00	80.39
POL	58	40.04	24.98	0.14	85.99
POL_0	57	19.35	16.42	0.00	56.95
ELA_0	57	48.75	33.40	0.10	100.00
$GROUPS_0$	57	4.19	6.43	0.00	47.00
$EGIPGRPS_0$	57	1.37	1.33	0.00	7.00
$EXCLGRPS_0$	57	2.33	6.17	0.00	46.00
$MONPOP_0$	57	0.21	0.36	0.00	0.97
$DOMPOP_0$	57	0.21	0.34	0.00	0.98
(Log) GDP per capita	58	8.53	1.21	5.87	10.63

We also obtain several additional variables from the EPR data. $GROUPS_0$ is the number of relevant (active) ethno-political groups. $EGIPGRPS_0$ is the number of included ethno-political groups at the last year before the slump; that is, groups with have some level access to executive power. $EXCLGRPS_0$ is the number of ethno-political groups without access to the political executive. Finally, $DOMPOP_0$ and $MONPOP_0$ are the population shares of the dominant or monopoly groups (the two highest levels of political power). All of these variables are fixed at the last year before the slump to rule out any feedback from the duration to the group composition. Table 1 describes all variables and lists the underlying data sources. Table 2 presents the associated summary statistics.

Empirical approach. Our approach is to examine partial correlations and test whether these are consistent with the proposed theory. While we cannot rule out all forms of endogeneity, we do take care to ensure temporal precedence. We employ standard event history techniques to study the duration of the decline phase.

To estimate the partial correlations, we run log-normal accelerated failure time (AFT) regressions of the form:

$$\ln \tilde{t} \equiv \ln(t - t_0) = \beta_0 + \beta_1 XCONST_0 + \beta_2 H + \beta_3 (XCONST_0 \times H) + \mathbf{x}'_0 \boldsymbol{\xi} + \epsilon_t \quad (12)$$

where \tilde{t} is analysis time, t_0 is the last year before the slump, $XCONST_0$ is executive constraints, H is a measure of group (ethnic) heterogeneity, \mathbf{x}_0 is a vector of controls, and $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon)$. Variables which could endogenously react to a prolonged duration of declines are kept fixed at t_0 to rule out any such feedback; we drop the subscript if they have no time dimension. All parameters are estimated using Maximum Likelihood and the standard errors are clustered on the country level to account for repeated spells.

Our main parameters of interest are β_1 , β_2 , and β_3 . In several regressions, we impose $\beta_3 = 0$ to estimate first-order effects before examining a possible interaction effect. The vector \mathbf{x}_0 may include additional heterogeneity measures, the (log of) initial GDP per capita before the slump, region effects and decade dummies.

Accelerated failure time models are so named due to their interpretation. A coefficient greater than zero implies that time passes more slowly, so that the exit of the decline phase is prolonged. A coefficient less than zero implies that time passes more quickly and hence that the recovery starts sooner. Alternatively, we may simply read the effects as elasticities (or semi-elasticities) of the expected duration with respect to the variables on the right hand side. While duration models have the main benefit of accounting for unfinished spells (ongoing declines), their interpretation is otherwise identical to log-linear OLS when they are cast in the AFT form.

Results. Table 3 presents the first set of results relating mainly to the empirical predictions derived from the baseline model. We compute two sets of estimates. One using the *Ethnologue* data which focuses on linguistic diversity and one based on the EPR data which incorporates only politically-relevant groups divided along the predominant social cleavage (ethnic, linguistic, racial, etc.).

Columns (1) to (3) use the *Ethnologue* data. Column (1) establishes that stronger constraints on the political executive shorten the expected duration of the decline phase and that greater linguistic heterogeneity has an adverse effect on the expected duration. The effects are statistically significant at the 1%-level and economically meaningful. A one point improvement in executive constraints (before the slump) leads to an approximate 18.7% reduction in the duration until the trough. Conversely, a one percentage point change towards greater linguistic heterogeneity prolongs the decline phase by about 1.7%. Column (2) allows for a conditional effect and strongly suggests that the effect of political institutions depends on the level of linguistic diversity (and *vice versa*). Whenever we introduce an interaction term, we first center the two constituent variables on their average. This shifts the coefficients of the two base levels into a meaningful range, but leaves the magnitude and statistical significance of the interaction term unaffected. Holding the other variable constant, the coefficient on either base variable now measures the effect of a one unit change away from the average. As a result, the interaction effect can be ignored; it has to be taken into account only when both variables change. The interaction between executive constraints and linguistic fractionalization is significant at the 5%-level and comparatively large. The specification predicts that at perfect homogeneity the median decline lasts about 2 years, while at perfect heterogeneity it lasts about 15 years. These estimates cover nearly all of the observed differences between

Table 3 – Baseline – Executive Constraints, Heterogeneity and Interactions

VARIABLES	<i>Dependent Variable: $\ln \tilde{t}$</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Ethnologue</i>			<i>Ethnic Power Relations</i>		
$XCONST_0$	-0.187*** (0.063)	-0.291*** (0.092)	-0.171*** (0.064)	-0.187*** (0.067)	-0.262*** (0.085)	-0.170** (0.067)
ELF	0.017*** (0.004)	0.019*** (0.004)	0.023*** (0.006)			
ELF_0				0.020*** (0.007)	0.023*** (0.007)	0.025*** (0.007)
$XCONST_0 \times ELF_{15}$		-0.004** (0.002)				
POL			-0.011 (0.007)			
$XCONST_0 \times ELF_0$					-0.004* (0.002)	
POL_0						0.012 (0.009)
	<i>Control sets</i>					
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
	<i>Summary stats</i>					
Exits	47	47	47	47	47	47
Spells	57	57	57	57	57	57
Years of Decline	346	346	346	346	346	346
Log- \mathcal{L}	-73.677	-71.621	-72.679	-76.294	-74.952	-75.597
Pseudo- R^2	0.149	0.173	0.161	0.119	0.134	0.127

Note(s): The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

declines in Western Europe and Sub-Saharan Africa. The results in columns (1) and (2) are consistent with our theoretical predictions; greater constraints on the executive shorten the expected duration unless the society is nearly homogeneous. The partial effect of executive constraints is not statistically different from zero for low ELF values. Column (3) adds the linguistic polarization measure to the specification in column (1). The literature on civil conflict stresses that polarization matters; e.g. Esteban and Ray (2011) show theoretically that conflict over public goods is driven by polarization and conflict over private goods by fractionalization. Contrary to this literature but in line with our model, we find no evidence in favor of the hypothesis that polarization is an issue for (the lack of) cooperation during declines, while the coefficient on fractionalization is robust to this perturbation. In other words, existence of two equally powerful groups does not predict longer declines than, say, three equally-sized linguistic groups.

Measures of linguistic diversity tend to describe Sub-Saharan Africa as more diverse in comparison to other regions than alternative diversity measures. This begs the question

if we are just estimating an “Africa effect”. Columns (4) to (6) use the EPR data which addresses this issue by alternating the relevant cleavage by country (from racial over linguistic to religious). This changes the relative location of Sub-Saharan Africa, which is only the second most diverse region on this measure, after South Asia, contrary to being the most linguistically diverse region based on the *Ethnologue* data. The EPR data also only codes politically relevant groups, so that the level of heterogeneity – no matter the measure – is generally lower. Note that we compute the heterogeneity measures for all politically relevant groups, not just the included groups. Strikingly, the results are virtually unchanged. Column (4) shows that the first order effects of executive constraints are the same, and the effect of fractionalization is well within one standard error of the previous estimate. Moreover, the sign and size of the interaction effect in column (5) is nearly identical to the one in column (2). Only the statistical significance of the interaction effect is a bit lower (cluster robust t-stat = -1.71). Column (5) shows that we also find no evidence in favor of *ethno-political* polarization affecting the duration of declines, just as with linguistic polarization. Contrasting these results to the *Ethnologue* data, it seems safe to conclude that we are not only explaining that declines in Sub-Saharan Africa last longer than elsewhere because the subcontinent is the most linguistically diverse, but that this holds true when we account for political relevance and the prevalence of different divisions in different parts of the world.

Overall, Table 3 provides significant evidence that there is a robust partial correlation of the duration of the decline phase with executive constraints on the one hand and with ethnic diversity on the other hand. In addition, the effect of weak constraints on the executive seems to be conditional on the degree of (ethno-political) fractionalization.

In Table 4 we “unpack” these statements further and examine what type of group configurations give rise to the cooperation problem we are analyzing. These results relate mainly to the empirical content of the model extensions. We now only use the EPR data, as it provides the necessary detail on the number of groups, their power status and more. Column (1) presents a result that may appear puzzling at first sight. If we measure heterogeneity simply by the number of ethno-political groups, then we find no evidence of an effect on the duration of declines. However, this coefficient amalgamates two effects. The EPR data distinguishes between *included* groups, which have access to executive power, and *excluded* groups, which lack power at the state level or are (at worst) discriminated against. In the model presented earlier, only the former are relevant players and thus we also expect that only the former category is empirically relevant. Column (2) confirms this expectation. The effect of included groups is statistically significant at the 1%-level and economically very large: an additional group increases the duration of the decline phase by about 53%. On the contrary, the effect of excluded groups is estimated to be near zero and has a comparatively tight 99% confidence interval centered near zero. In line with the theory, these results suggest that only ethnic groups with some degree of access to political power matter for the duration of declines.

We have not yet isolated whether this adverse effect of heterogeneity is due to several equally powerful groups co-existing in society or due to particular asymmetries in political power. Columns (3) to (6) represent different attempts towards empirically answering this question. In Comment 4, we translated the theoretical results regarding political concentration from Proposition 3 and the number of groups from Proposition 4 into the concepts of monopoly groups and dominant groups. Column (3) is the empirical counterpart. Here we relate the share of population represented by a group that either monopolizes or dominates the political executive to the duration of the decline phase.

Table 4 – Extensions – Number of Groups, Political Relevance, and Asymmetries

VARIABLES	<i>Dependent Variable: $\ln \bar{t}$</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
$XCONST_0$	-0.225*** (0.070)	-0.241*** (0.063)	-0.215*** (0.065)	-0.179*** (0.066)	-0.210*** (0.070)	-0.200*** (0.073)
$GROUPO_0$	-0.008 (0.018)			-0.031** (0.014)		
$EGIPGRPS_0$		0.426*** (0.095)			0.290** (0.124)	0.300*** (0.111)
$EXCLGRPS_0$		-0.012 (0.013)			-0.021* (0.012)	-0.014 (0.012)
$DOMPOP_0$			-0.702* (0.361)			
$MONPOP_0$			-1.140** (0.484)			
ELF_0				0.022*** (0.007)	0.013 (0.009)	
ELA_0						-0.013** (0.006)
<i>Control sets</i>						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
<i>Summary stats</i>						
Exits	47	47	47	47	47	43
Spells	57	57	57	57	57	53
Years of Decline	346	346	346	346	346	334
Log- \mathcal{L}	-81.069	-75.062	-77.647	-75.457	-73.253	-66.669
Pseudo-R ²	0.064	0.133	0.103	0.129	0.154	0.166

Note(s): The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). *** p<0.01, ** p<0.05, * p<0.1.

The results are unambiguous. Both variables are associated with substantially shorter declines. Columns (4) and (5) try to explicitly tackle the issue of the number of groups versus group asymmetries. In column (4), we include the number of ethno-politically relevant group together with the index of ethno-political fractionalization. This leads to an interesting *ceteris paribus* condition. Increasing the degree of fractionalization by one percentage point while holding constant the number of groups necessarily implies that political concentration is decreasing; that is, the groups are becoming more alike. Once again, recall that for any given number of groups, fractionalization is maximized at equal shares. The estimates thus suggest that less political concentration leads to longer declines. Column (5) again distinguishes between included and excluded groups to illustrate that only the former are relevant. The coefficient on the ELF_0 measure loses significance, suggesting that the number of included groups may drive the effect of ethnic heterogeneity and that group imbalances hardly matter. However, column (6) addresses this issue more directly by using our index of ethnic asymmetries and

provides the same answer as column (4). Now the effect is easy to interpret, negative and significant at the 5% level. A one percentage point move towards greater asymmetries (political concentration) shortens the duration by about 1.3%. Note that the effect of executive constraints remains robust throughout, fluctuating around a 20% reduction in the duration of declines for a one point improvement.

To summarize, Table 4 adds several valuable insights about the effect of ethnic diversity on the duration of declines. Fractionalization of linguistic or ethno-political groups masks two effects: 1) the expected duration is increasing in the number of politically relevant groups, and 2) the expected duration is decreasing in greater group asymmetries (political concentration). Both theory and evidence suggest that this is not an issue of polarization, but rather an issue of adding a smaller, potentially irrelevant, group to any multi-modal distribution of power as opposed to adding another powerful group (an additional mode).

Table 5 – Robustness – Region and time effects

VARIABLES	<i>Dependent Variable: $\ln \tilde{t}$</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
$XCONST_0$	-0.245*** (0.079)	-0.216*** (0.065)	-0.211*** (0.077)	-0.171** (0.068)	-0.181*** (0.066)	-0.145** (0.064)
ELF	0.020*** (0.004)	0.021*** (0.003)				
ELF_0			0.017** (0.007)	0.015** (0.006)		
$XCONST_0 \times ELF$	-0.003** (0.001)	-0.004*** (0.001)				
$XCONST_0 \times ELF_0$			-0.003 (0.002)	-0.003* (0.002)		
$EGIPGRPS_0$					0.298*** (0.097)	0.212* (0.119)
$EXCLGRPS_0$					0.024* (0.013)	0.009 (0.026)
<i>Control sets</i>						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	No	Yes	No	Yes	No	Yes
<i>Summary stats</i>						
Exits	47	47	47	47	47	47
Spells	57	57	57	57	57	57
Years of Decline	346	346	346	346	346	346
Log- \mathcal{L}	-62.777	-55.812	-67.966	-64.108	-68.705	-66.701
Pseudo-R ²	0.275	0.355	0.215	0.260	0.207	0.230

Note(s): The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5 selects three key specifications, one for each data source, and then subjects

them to two robustness checks. First, we return to the issue of whether we are estimating an “Africa effect” by including region dummies in each specification. Second, we control for temporal heterogeneity by including a dummy for the decade in which the slump began in every other specification, since the 1970s, 1980s and 1990s exhibit significantly higher volatility than the other decades.

The Africa dummy is significant in all specifications, capturing that declines take substantially longer on the African continent. Nevertheless, we still find comparable effects. Column (1) uses the *Ethnologue* data and shows that our two variables of interest are robustly correlated with within region differences in the duration of declines. Column (2) adds that this is still the case when we also control for temporal heterogeneity. Using the *EPR* data, columns (4) and (5) verify that the same holds for ethno-political fragmentation. The standard error of the interaction term becomes somewhat wider, leading to a loss of significance, but the estimated coefficient is extremely stable. The last two columns show that this also hold for the effect of the number of included groups. In general, there is significant evidence of regional heterogeneity (a χ^2 -test always rejects the null of no heterogeneity at the 5%-level), but there is somewhat less evidence of temporal heterogeneity (on top of duration dependence).⁸ Throughout Table 5 the coefficient of political institutions and the coefficients of the various measures of ethnic heterogeneity remain statistically significant at conventional levels and well within their usual range.

This last set of empirical findings reveals the following: first, ethnic heterogeneity and constraints on the political executive are robust determinants of the length of the decline phase during major economic slumps. Second, this result is not due to regional differences in ethnic heterogeneity but holds when we only use within region variation. Taken as a whole, we believe that these results operationalize the key parameters of the model and demonstrate that there is robust evidence consistent with the theory outlined here.

6 Concluding remarks

This chapter presents a political economy theory of declines, highlighting a commitment problem between winners and losers of the recovery process after a crisis, and then analyzes empirical implications of this theory. We show that ethno-political heterogeneity coupled with weak constraints on the political executive can bring about delayed cooperation during the decline phase of a slump and hence explain why we observe such long declines in some countries and relatively short declines in others.

Both the theory and the empirical analysis suggest that ethnic heterogeneity is indeed harmful for getting groups to agree on a response to a crisis when political institutions are weak. More subtle predictions show that this is mostly an issue of having many powerful groups in the society and does not apply to the same degree when there is a politically dominant group. The overarching policy implication here is not that ethnic diversity is necessarily a problem, but that political institutions can be designed to contain the adversarial element of ethnic heterogeneity in particular and political heterogeneity in general. While not restricted to understanding declines in Sub-Saharan Africa, we would like to once again emphasize that we believe these insights are particularly important for understanding the political economy of declines on the subcontinent. Sub-Saharan

⁸A χ^2 -test rejects the null of no temporal heterogeneity at the 1%-level in column (1), at the 5%-level but not the 1%-level in column (4), and fails to reject the null at conventional levels in column (6).

Africa is home to the longest and deepest declines, politics shaped by ethnicity, and weak institutions governing executive power. While we still need to better understand why ethnic diversity tends to coincide with weak political institutions and how one shapes the other, we find that there is ample room for managing this heterogeneity better so that welfare gains are not lost in the next crisis.

This line of research is far from complete. Fruitful avenues for future research would be to extend these models further by integrating a richer description of the executive decision-making process, altering the decision rules, treating the quality of political institutions as endogenous to the decline, or modeling details of the policy response. On the empirical side, richer data on cabinet allocations, ethnicity and executive power would help to trace out the proposed mechanism more carefully.

Bibliography

- Acemoglu, D. and J. A. Robinson (2006). *Economic Origins of Democracy and Dictatorship*. Cambridge University Press.
- Aghion, P., A. Alesina, and F. Trebbi (2004). Endogenous political institutions. *Quarterly Journal of Economics* 119(2), 565–611.
- Ahlerup, P. and O. Olsson (2012). The roots of ethnic diversity. *Journal of Economic Growth* 17(2), 71–102.
- Alesina, A., R. Baqir, and W. Easterly (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics* 114(4), 1243–1284.
- Alesina, A. and A. Drazen (1991). Why are stabilizations delayed? *American Economic Review* 81(5), 1170–1188.
- Alesina, A. and E. L. Ferrara (2005). Ethnic diversity and economic performance. *Journal of Economic Literature* 43(3), 762–800.
- Alesina, A., C. Gennaioli, and S. Lovo (2014, September). Public goods and ethnic diversity: Evidence from deforestation in Indonesia. Working Paper 20504, National Bureau of Economic Research.
- Alesina, A. F., S. Michalopoulos, and E. Papaioannou (2012, November). Ethnic inequality. Working Paper 18512, National Bureau of Economic Research.
- Ashraf, Q. and O. Galor (2013). The “Out of Africa” hypothesis, human genetic diversity, and comparative economic development. *American Economic Review* 103(1), 1–46.
- Bates, R. H. (2000). Ethnicity and development in Africa: A reappraisal. *American Economic Review* 90(2), 131–134.
- Berg, A., J. D. Ostry, and J. Zettelmeyer (2012). What makes growth sustained? *Journal of Development Economics* 98(2), 149–166.
- Besley, T. and T. Persson (2011a). Fragile states and development policy. *Journal of the European Economic Association* 9(3), 371–398.
- Besley, T. and T. Persson (2011b). The logic of political violence. *Quarterly Journal of Economics* 126(3), 1411–1445.
- Bueno de Mesquita, B., A. Smith, and R. M. Siverson (2005). *The logic of political survival*. MIT press.
- Caselli, F. and W. J. Coleman (2013). On the theory of ethnic conflict. *Journal of the European Economic Association* 11, 161–192.
- Cederman, L.-E., A. Wimmer, and B. Min (2010). Why do ethnic groups rebel? New data and analysis. *World Politics* 62(1), 87–119.
- Cerra, V. and S. C. Saxena (2008). Growth dynamics: The myth of economic recovery. *American Economic Review* 98(1), 439–57.
- Collier, P. (2000). Ethnicity, politics and economic performance. *Economics & Politics* 12(3), 225–245.
- Cox, G. W. and M. D. McCubbins (1997). Political structure and economic policy: The institutional determinants of policy outcomes. In S. Haggard and M. D. McCubbins (Eds.), *Presidents, Parliaments, and Policy*, pp. 21–63. Cambridge University Press.
- Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2012). The political economy of linguistic cleavages. *Journal of Development Economics* 97(2), 322–338.
- Drazen, A. and V. Grilli (1993). The benefit of crises for economic reforms. *American Economic Review* 83(3), 598–607.
- Easterly, W. (2001). Can institutions resolve ethnic conflict? *Economic Development and Cultural Change* 49(4), 687–706.
- Easterly, W., M. Kremer, L. Pritchett, and L. Summers (1993). Good policy or good luck? *Journal of Monetary Economics* 32(3), 459–483.
- Easterly, W. and R. Levine (1997). Africa’s growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics* 112(4), 1203–1250.
- Eifert, B., E. Miguel, and D. N. Posner (2010). Political competition and ethnic identification in Africa. *American Journal of Political Science* 54(2), 494–510.
- Esteban, J., L. Mayoral, and D. Ray (2012). Ethnicity and conflict: An empirical study. *American Economic Review* 102(4), 1310–1342.
- Esteban, J. and D. Ray (1994). On the measurement of polarization. *Econometrica* 62(4), 819–851.
- Esteban, J. and D. Ray (2011). Linking conflict to inequality and polarization. *American Economic Review* 101(4), 1345–1374.

- Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth* 8(2), 195–222.
- Fearon, J. D. and D. D. Laitin (2003). Ethnicity, insurgency, and civil war. *American Political Science Review* 97, 75–90.
- Fernandez, R. and D. Rodrik (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *American Economic Review* 81(5), 1146–1155.
- Francois, P., I. Rainer, and F. Trebbi (2012, September). How is power shared in Africa? Working Paper 18425, National Bureau of Economic Research.
- Gehlbach, S. and E. J. Malesky (2010). The contribution of veto players to economic reform. *The Journal of Politics* 72(4), 957–975.
- Haggard, S. (2000). *The political economy of the Asian financial crisis*. Peterson Institute for International Economics.
- Hausmann, R., L. Pritchett, and D. Rodrik (2005). Growth accelerations. *Journal of Economic Growth* 10(4), 303–329.
- Hicken, A., S. Satyanath, and E. Sergenti (2005). Political institutions and economic performance: The effects of accountability and obstacles to policy change. *American Journal of Political Science* 49(4), 897–907.
- Labán, R. and F. Sturzenegger (1994a). Distributional conflict, financial adaptation and delayed stabilizations. *Economics & Politics* 6(3), 257–276.
- Labán, R. and F. Sturzenegger (1994b). Fiscal conservatism as a response to the debt crisis. *Journal of Development Economics* 45(2), 305 – 324.
- Michalopoulos, S. (2012). The origins of ethnolinguistic diversity. *American Economic Review* 102(4), 1508–39.
- Miguel, E. and M. K. Gugerty (2005). Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics* 89(11–12), 2325–2368.
- Posner, D. N. (2004). Measuring ethnic fractionalization in Africa. *American Journal of Political Science* 48(4), 849–863.
- Rainer, I. and F. Trebbi (2012, September). New tools for the analysis of political power in Africa. Working Paper 18424, National Bureau of Economic Research.
- Rodrik, D. (1999). Where did all the growth go? External shocks, social conflict, and growth collapses. *Journal of Economic Growth* 4(4), 385–412.
- Spolaore, E. (2004). Adjustments in different government systems. *Economics & Politics* 16(2), 117–146.
- Tsebelis, G. (1995). Decision making in political systems: Veto players in presidentialism, parliamentarism, multicameralism and multipartyism. *British Journal of Political Science* 25(03), 289–325.
- Tsebelis, G. (2002). *Veto players: How political institutions work*. Princeton University Press.
- Wimmer, A., L.-E. Cederman, and B. Min (2009). Ethnic politics and armed conflict: A configurational analysis of a new global data set. *American Sociological Review* 74(2), 316–337.

Appendix: Proofs

Proof of Proposition 1. The utility from cooperation in the first period when the other group cooperates is

$$v_j^1(C, c) = \frac{1}{1-\delta} \{(1-2c)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + c(g(0) + g(1))\} \quad (\text{A-1})$$

and the utility from choosing to delay cooperation one period when the other group cooperates is

$$v_j^1(D, c) = g((1-\Delta)y_j) + \frac{\delta}{1-\delta} \{(1-2p^2)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^2(g(0) + g(1))\} \quad (\text{A-2})$$

where $p^2 = c - x$; that is, half the probability of landing outside the safe zone in the second period.

The proof is by contradiction. We conjecture an equilibrium with immediate recovery, such that $v_j^1(C, c) \geq v_j^1(D, c)$. Using $p^2 = c - x$ and rearranging terms, we get

$$g((1-\Delta)y_j) \leq \mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - \left[c + \frac{\delta}{1-\delta}x \right] \{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - g(0) - g(1)\}. \quad (\text{A-3})$$

Note that concavity implies that $\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - g(0) - g(1)\} > 0$. Inequality (A-3) is contradicted whenever c , x or δ are large enough in relation to Δ , depending on the shape of the utility function $g(y_j)$ and its range, which completes the proof. ■

Proof of Proposition 2. First of all, it is useful to demonstrate that the difference in utility between recovery at any time period (t) and recovery at the subsequent period ($t+1$) decreases over time. For all $s > t$, we need to check whether

$$v_j^{t+1}(C, c) - v_j^t(C, c) > v_j^{s+1}(C, c) - v_j^s(C, c). \quad (\text{A-4})$$

Note that $v_j^{t+1}(C, c) = v_j^t(D, c)$.

Substituting the utilities and rearranging the inequality, we get

$$\begin{aligned} & g((1-\Delta)y_j) - (1-2p^t)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - p^t(g(0) + g(1)) + \\ & \frac{\delta}{1-\delta} \{2(p^t - p^{t+1})\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (p^{t+1} - p^t)(g(0) + g(1))\} > \\ & g((1-\Delta)y_j) - (1-2p^s)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - p^s(g(0) + g(1)) + \\ & \frac{\delta}{1-\delta} \{2(p^s - p^{s+1})\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (p^{s+1} - p^s)(g(0) + g(1))\}. \end{aligned} \quad (\text{A-5})$$

Recall that $p^t = c - (t-1)x$ implies $p^{t+1} - p^t = -x$, so the second and third terms cancel and the inequality reduces to

$$\begin{aligned} & (1-2p^t)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^t(g(0) + g(1)) < \\ & (1-2p^s)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^s(g(0) + g(1)). \end{aligned} \quad (\text{A-6})$$

Substituting $p^t = c - (t-1)x$ again, it is straightforward to show that this inequality is

always satisfied when $s > t$.

Having established this, setting the utility of choosing to cooperate in period t equal to the utility of recovering in period $t + 1$ results in an equation that will deliver a potentially non-integer t , such that the smallest higher integer ($\lceil t \rceil$) is the equilibrium time to recovery:

$$\begin{aligned} \frac{1}{1-\delta} \{ (1-2p^t)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^t(g(0) + g(1)) \} &= g((1-\Delta)y_j) + \\ \frac{\delta}{1-\delta} \{ (1-2p^{t+1})\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^{t+1}(g(0) + g(1)) \}. & \end{aligned} \quad (\text{A-7})$$

Inserting the linear process on $p^t = c - (t-1)x$ yields

$$\begin{aligned} \frac{1}{1-\delta} \{ (1-2(c-(t-1)x))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-(t-1)x)(g(0) + g(1)) \} &= \\ g((1-\Delta)y_j) + \frac{\delta}{1-\delta} \{ (1-2(c-tx))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-tx)(g(0) + g(1)) \}. & \end{aligned} \quad (\text{A-8})$$

Isolating the first term of the geometric series gives

$$\begin{aligned} \{ (1-2(c-(t-1)x))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-(t-1)x)(g(0) + g(1)) \} + \\ \frac{\delta}{1-\delta} \{ (1-2(c-(t-1)x))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-(t-1)x)(g(0) + g(1)) \} = \\ g((1-\Delta)y_j) + \frac{\delta}{1-\delta} \{ (1-2(c-tx))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-tx)(g(0) + g(1)) \} \end{aligned} \quad (\text{A-9})$$

and after canceling the common terms, we have

$$\begin{aligned} (1-2c)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + c(g(0) + g(1)) + \\ tx\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - (g(0) + g(1))\} \\ = g((1-\Delta)y_j) + \frac{1}{1-\delta}x\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - (g(0) + g(1))\}. \end{aligned} \quad (\text{A-10})$$

Solving for t^* and simplifying gives

$$t^* = \frac{g((1-\Delta)y_j) - \mathbb{E}[g(w_j)|w_j \in \mathcal{A}]}{x\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - (g(0) + g(1))\}} + \frac{c}{x} + \frac{1}{1-\delta}. \quad (\text{A-11})$$

The proposition follows directly from comparative statics w.r.t. to c

$$\frac{\partial t^*}{\partial c} = \frac{1}{x} > 0 \quad (\text{A-12})$$

that is, stronger executive constraints (smaller c) shorten the time to cooperation. This completes the proof for the interior case. Note that it can also be the case that recovery happens at the point when all uncertainty is resolved, i.e. the point where the probability of being outside the safe zone is zero and no longer changes. If this is the case it is straightforward to see that the time to recovery is shorter with stronger constraints on the executive. This follows directly from the fact, that the time it takes until all uncertainty is resolved is shorter with smaller c . ■

Proof of Proposition 3. Recall that for the asymmetric case only one group risks falling outside the political safe zone. Hence, for there to exist an equilibrium with recovery in the first period, the following condition needs to be true

$$\frac{1}{1-\delta} \{(1-p_1(y_j))y_j + p_1(y_j)z\} \geq (1-\Delta)y_j + \frac{\delta}{1-\delta}y_j \quad (\text{A-13})$$

which simplifies to

$$\Delta y_j + \frac{1}{1-\delta} \{p_1(y_j)(z - y_j)\} \geq 0. \quad (\text{A-14})$$

An decrease in concentration (asymmetry) makes delay more likely if the left hand side of the inequality is a decreasing function of y_j . This is true when the derivative of the left hand side is negative:

$$\Delta + \frac{1}{1-\delta} \left\{ \frac{dp_1(y_j)}{dy_j} (z - y_j) - p_1(y_j) \right\} < 0 \quad (\text{A-15})$$

which completes the proof. ■

Proof of Proposition 4. Remember that there always exists an equilibrium with recovery in period two in the sub-game that starts in period two after delay in period one. If all of the other groups decide to cooperate in period 1, it is optimal for the remaining group to cooperate if the following condition holds

$$\frac{1}{1-\delta} \left\{ (1 - Jp_1(y_j)) \frac{1}{J} + (J-1)p_1(y_j) \left(\frac{1}{J} + \frac{1}{(J-1)J} \right) + p_1(y_j)z \right\} \geq (1-\Delta) \frac{1}{J} + \frac{\delta}{1-\delta} \frac{1}{J}. \quad (\text{A-16})$$

The second term inside the curly braces simplifies to $p_t(y_j)$, so that inequality (A-16) becomes

$$\frac{1}{1-\delta} \left\{ (1 - Jp_1(y_j)) \frac{1}{J} + p_1(y_j) + p_1(y_j)z \right\} \geq (1-\Delta) \frac{1}{J} + \frac{\delta}{1-\delta} \frac{1}{J} \quad (\text{A-17})$$

or

$$\frac{1}{1-\delta} \left\{ \frac{1}{J} + p_1(y_j)z \right\} \geq (1-\Delta) \frac{1}{J} + \frac{\delta}{1-\delta} \frac{1}{J} \quad (\text{A-18})$$

and, after some algebraic manipulation, this simplifies to

$$\frac{\Delta}{J} + \frac{1}{1-\delta} p_1(y_j)z \geq 0. \quad (\text{A-19})$$

Note that $p_1(y_j)$ is increasing in J , as symmetry implies $y_j = 1/J$, and $z < 0$ by definition. As a result, the inequality becomes harder to satisfy if the number of groups increases, which completes the proof. ■

Part II

Growth, distribution and poverty

Chapter 5

The pace of poverty reduction

*A fractional response approach**

Abstract

The pace of poverty reduction through growth v. redistribution is at the heart of current debates on equitable development. In this chapter, we argue that empirical poverty decompositions should build in the inherent boundedness of the poverty headcount ratio directly. As a solution, we propose a fractional response approach to estimating poverty decompositions, and present extensions dealing with unobserved heterogeneity, measurement error and unbalancedness. Using a large new data set, we estimate income and inequality (semi-)elasticities of poverty for the \$2 a day and \$1.25 a day poverty lines. The models fit the data remarkably well over the entire data range. We highlight the relevance of focusing on semi-elasticities for policy purposes and, building on the improved accuracy of the fractional response results, we present poverty projections from 2010 through 2030. Finally, we discuss some implications of these results for the post-2015 development agenda.

Keywords: poverty, inequality, fractional response models, income growth

JEL Classification: I32, C25, O10, O15

*This chapter is based on Bluhm, R., de Crombrugghe, D., and A. Szirmai. 2013. The pace of poverty reduction - A fractional response approach," MERIT Working Papers 051, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

1 Introduction

The pace of poverty reduction through economic growth is at the heart of ongoing debates on inclusive growth and equitable development. Given the salience of this issue to both policy makers and academics, an increasingly large literature analyzes the impact of changes in incomes and inequality on poverty, as well as their respective contributions towards poverty reduction in the past decades (Ravallion and Chen, 1997; Besley and Burgess, 2003; Bourguignon, 2003; Kraay, 2006; Kalwij and Verschoor, 2007; Bresson, 2009; Chambers and Dhongde, 2011). Collectively, these studies established not only that income growth is crucial to achieving sustained decreases in poverty but also that the benefits of income growth strongly depend on the initial levels of income and inequality. In fact, poverty is linked to income and distribution through a decomposition identity that should guide empirical studies of poverty (Datt and Ravallion, 1992; Kakwani, 1993; Bourguignon, 2003). In this chapter, we revisit and extend this literature in several ways.

First, and most fundamentally, we argue that the previous literature based on linear models of poverty changes ignores the bounded nature of the poverty headcount ratio and either disregards or awkwardly reintroduces the information provided by income and inequality levels. This leads to poorly fitting approximations of the underlying decomposition identity and estimates of the relevant elasticities that can take on implausible values. Instead, we propose that empirical models of poverty should capture the boundedness of the poverty headcount ratio directly, and thus build in the inherent non-linearity of the relationship and the non-constancy of the elasticities. To this end, we derive fractional response models of the poverty headcount ratio. We present extensions of these models that deal with unobserved heterogeneity, measurement error and unbalanced panel data (see Papke and Wooldridge, 1996, 2008; Wooldridge, 2010a). A key result is the greatly improved accuracy of the poverty decomposition. Our specifications fit the data remarkably well and predict the observed poverty headcount ratio with less than 2.5 percentage points error on average. For comparison, we also reproduce the traditional approach and highlight some of its empirical shortcomings.

Second, using a new data set of 809 nationally representative surveys covering 124 countries from 1981 to 2010, we estimate income and inequality elasticities of the poverty headcount ratio for different regions and time periods. Our findings generally confirm the result of previous studies that the average income elasticity of poverty is around two. However, in order to understand the speed of poverty reduction across different regions and time periods, we are particularly interested in reliable estimates of values other than the overall cross-country mean. Our method provides considerably more precise regional and temporal estimates of the income and inequality elasticities of poverty that often contradict earlier studies. For example, we find universally higher income elasticities in Latin America and Eastern Europe and Central Asia but lower income elasticities in both South Asia and Sub-Saharan Africa than reported earlier (Kalwij and Verschoor, 2007). Furthermore, since income or inequality elasticities of poverty are concepts of relative change, they may give the misleading impression that richer countries are becoming ever better at reducing poverty even though the underlying absolute changes are small. Hence, we also stress the importance of semi-elasticities which capture the absolute change in poverty for a given rate of income growth or proportional change in the income distribution (Klasen and Misselhorn, 2008).

Last but not least, based on the much improved accuracy of our new estimates, we present projections of the poverty headcount ratio for the \$2 a day and \$1.25 a day poverty

lines until 2030. We find that absolute poverty in Sub-Saharan Africa and, as a not too distant second, South Asia remains the primary development challenge of the twenty-first century. In all other regions, poverty is projected to nearly disappear or fall to much less than 10% of the population by 2030. These results are based on the assumption that per capita incomes and expenditures in each country continue to grow at the average pace of the last 15 years. While it is clearly far from guaranteed that this will hold true, other assumptions are easily evaluated using our proposed estimation framework. The projected poverty trends have direct bearing on the post-2015 development agenda and can be used to assess the viability of proposed development goals.

The remainder of this chapter is organized as follows. Section 2 reviews how the existing literature models poverty, derives our approach, and discusses the econometrics of fractional response models. Section 3 briefly outlines the data used in this chapter. Section 4 presents the estimation results, regional elasticities, and poverty projections until 2030. Section 5 concludes.

2 Modeling poverty and elasticities

2.1 Traditional approaches: linear models of poverty changes

Using micro-data, changes in poverty can be decomposed into changes in income and the income distribution up to some error (Datt and Ravallion, 1992; Kakwani, 1993). A key problem for cross-country studies of poverty is that we generally do not have access to micro data sets of incomes and/or consumption expenditures for all countries but usually estimate poverty based on grouped data.¹ To overcome this limitation, Bourguignon (2003) suggests approximating the entire income distribution of each country using a two-parameter log normal distribution – an approach that is theoretically grounded², simple and popular but not without its critics (e.g. Bresson, 2009).

Bourguignon assumes that income, y_t , is a log normal random variable, such that $\ln y_t \sim \mathcal{N}(\mu_t, \sigma_t^2)$, and mean income can be written as $\bar{y}_t = E[y_t] = \exp(\mu_t + \sigma_t^2/2)$. Then the poverty headcount ratio (henceforth, poverty headcount) at time t may be defined as

$$H_t = H(\bar{y}_t/z, \sigma_t) = \Phi\left(\frac{-\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t\right) \equiv \Pr[y_t \leq z] \quad (1)$$

where $\Phi(\cdot)$ denotes the standard normal cdf, inequality is measured as the standard deviation (σ_t) of log income, and \bar{y}_t/z is the (relative) distance of mean income (\bar{y}_t) to the (fixed) poverty line (z) – which we interpret as a ‘shortfall’ when $\bar{y}_t < z$ and ‘affluence’ when $\bar{y}_t > z$.

¹There have been some attempts to either collect all the available primary data or to estimate gaps in survey coverage with the help of national accounts. Milanovic (2002) compiles a global data set of household level data to study the evolution of inequality, Sala-i-Martin (2006) estimates a “world income distribution” via kernel density approximations based on grouped data, and Kraay (2006) fits three-parameter Lorenz curves to grouped data in order to estimate the entire income distribution for the country-poverty spells in his sample.

²Gibrat’s law, for example, illustrates how the log normal distribution can arise from a sequence of stochastic income shocks $\ln y_t = \ln y_{t-1} + e_t$, so that e_t is a random transitory shock in log income and as t grows the distribution of e_t defines the distribution of $\ln y_t$. Battistin, Blundell, and Lewbel (2009) recently argued that this process is better thought of in terms of permanent income and suggest that consumption is closer to a log normal distribution.

Eq. (1) can be interpreted as the probability that, at a particular time t , an individual randomly drawn from the population is poor. This formulation gave rise to a large literature deriving the income and inequality elasticities of poverty analytically and estimating econometric models inspired by their analytic counterparts (e.g. Bourguignon, 2003; Kalwij and Verschoor, 2007; Klasen and Misselhorn, 2008). To summarize the essence of this approach note that, based on eq. (1), we can derive the income elasticity ($\varepsilon_t^{H\bar{y}} = \frac{\partial H_t}{\partial \bar{y}_t} \frac{\bar{y}_t}{H_t}$) and inequality elasticity ($\varepsilon_t^{H\sigma} = \frac{\partial H_t}{\partial \sigma_t} \frac{\sigma_t}{H_t}$) of the poverty headcount as

$$\varepsilon_t^{H\bar{y}} = -\frac{1}{\sigma_t} \lambda \left(\frac{-\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2} \sigma_t \right) \quad (2)$$

and

$$\varepsilon_t^{H\sigma} = \left(\frac{\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2} \sigma_t \right) \lambda \left(\frac{-\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2} \sigma_t \right) \quad (3)$$

where we define the inverse Mills ratio ($\lambda(x) \equiv \phi(x)/\Phi(x)$) as the ratio of the standard normal pdf to the standard normal cdf, and we require $H_t > 0$.

The decomposition of the poverty headcount is often written as

$$\frac{dH_t}{H_t} \approx \varepsilon_t^{H\bar{y}} \frac{d\bar{y}_t}{\bar{y}_t} + \varepsilon_t^{H\sigma} \frac{d\sigma_t}{\sigma_t} \quad (4)$$

where dH_t/H_t is a small relative change in the poverty headcount, $d\bar{y}_t/\bar{y}_t$ is a small relative change in mean incomes, and $d\sigma_t/\sigma_t$ is a small relative change in the standard deviation of log incomes. The approximation follows from a linear Taylor expansion of H_t . Appendix A derives this result.³

Given log normality, the standard deviation is a monotone transformation of the Gini inequality coefficient, denoted G_t , and can be obtained via $\sigma_t = \sqrt{2\Phi^{-1}(G_t/2 + 1/2)}$.⁴ So, eqs. (2) and (3) can be used to predict the elasticities directly using observed values of income and inequality. With some additional algebra, we can also derive an expression for the Gini elasticity and rewrite eq. (4) accordingly – see eq. (A-6) in Appendix A.

However, the assumption of log normality is only an approximation and unlikely to hold exactly. The key observation motivating the econometric models is that both elasticities depend only on the initial levels of mean income and inequality (when the poverty line is fixed). In order to allow for misspecification of the functional form, the authors rely on a linear approximation of these intrinsically non-linear functions. They capture the dependence on initial levels by interacting both mean income and inequality with the ratio of initial mean income to the poverty line and with initial inequality. This model is sometimes called the “improved standard model” (Bourguignon, 2003) and it is usually formulated in (annualized) differences:

$$\begin{aligned} \Delta \ln H_{it} = & \alpha + \beta_1 \Delta \ln \bar{y}_{it} + \beta_2 \Delta \ln \bar{y}_{it} \times \ln(\bar{y}_{i,t-1}/z) + \beta_3 \Delta \ln \bar{y}_{it} \times \ln G_{i,t-1} \\ & + \gamma_1 \Delta \ln G_{it} + \gamma_2 \Delta \ln G_{it} \times \ln(\bar{y}_{i,t-1}/z) + \gamma_3 \Delta \ln G_{it} \times \ln G_{i,t-1} + \epsilon_{it} \end{aligned} \quad (5)$$

where Δ is the difference operator, α is a linear time trend and ϵ_{it} is an error term. We also added country subscripts.

³Also see Datt and Ravallion (1992), Kakwani (1993) and Bourguignon (2003).

⁴This formulation is the inverse of $G_t = 2\Phi(\sigma_t/\sqrt{2}) - 1$ due to Aitchison and Brown (1957).

Suppose eq. (5) is estimated via OLS or GMM, then it is straightforward to see that the implied elasticities approximate eqs. (2) and (3). The estimated income elasticity is $\hat{\varepsilon}_{it}^{Hy} = \hat{\beta}_1 + \hat{\beta}_2 \ln(\bar{y}_{i,t-1}/z) + \hat{\beta}_3 \ln G_{i,t-1}$ and the estimated inequality elasticity is $\hat{\varepsilon}_{it}^{HG} = \hat{\gamma}_1 + \hat{\gamma}_2 \ln(\bar{y}_{i,t-1}/z) + \hat{\gamma}_3 \ln G_{i,t-1}$. Clearly, the elasticities depend on the initial levels of income and inequality. These two elasticities are sometimes referred to as the ‘distribution-neutral’ income elasticity and the ‘growth-neutral’ inequality elasticity. They identify the partial effect of changing either income or inequality, contrary to the simple observed elasticity that confounds both effects.⁵ Yet, this approximation is extremely coarse and does not place any meaningful restrictions on the parameter space. Moreover, it is unclear which level equation this specification derives from.

In general, poverty elasticities can paint a distorted picture of poverty dynamics. The income elasticity, for example, gives the impression that richer countries become ever better at poverty reduction because a drop in the poverty headcount from 2% to 1% is treated just the same as a drop from 50% to 25%. Recognizing this shortcoming, Klasen and Misselhorn (2008) suggest to focus on absolute poverty changes instead. Removing the log from the headcount in eq. (5) turns it into a model of semi-elasticities and alters the interpretation.⁶ The coefficients now measure the percentage point change in the population that is below the poverty line for a given rate of change in income or inequality. Likewise, eqs. (2) and (3) can be written as semi-elasticities by replacing the inverse Mills ratio with the standard normal pdf. Contrary to elasticities, the semi-elasticities approach zero as mean income becomes large. Klasen and Misselhorn (2008) also report that their models fit the data better and suggest that the specification in absolute changes captures more of the inherent non-linearity. Nevertheless, given the underlying identity, the fit typically found in the literature ($.55 \leq R^2 \leq .73$) is not particularly high.

There are advantages to relying on a linear framework other than mere simplicity, such as the well-known robustness properties of popular estimators. Nevertheless, a specification like eq. (5) suffers from several problems. First, it completely disregards the information provided by poverty levels, most likely introduces negative serial correlation, and compounds pre-existing measurement error.⁷ Second, after differencing removes the time-constant unobserved effects, the added interaction terms reintroduce the unobserved effects present in the lagged levels. Hence, the coefficients are likely to be biased no matter if the model is estimated using Ordinary Least Squares (OLS), Instrumental Variables (IV) or Generalized Methods of Moments (GMM) if time invariant measurement differences exist.⁸ *Third and most fundamentally*, the intrinsic non-linearity of the (semi-) elasticities is due to the bounded nature of the dependent variable which is not taken into account by linear specifications. The poverty headcount is a proportion and thus only takes on values in the unit interval ($H_{it} \in [0, 1]$). As a result, the linear estimates are

⁵For recent estimates of the simple empirical elasticity ($d \ln H_{it} / d \ln rgdpc_{it}$, where $rgdpc_{it}$ is real GDP per capita), see Ram (2013).

⁶The term semi-elasticity refers here to the quantity $\frac{\partial y}{\partial \ln x} = \frac{\partial y}{\partial x} x$ rather than the more usual $\frac{\partial \ln y}{\partial x} = \frac{\partial y}{\partial x} \frac{1}{y}$.

⁷If the errors in the original process (say, $y_{it} = \alpha + \beta(x_{it} + \nu_{it}) + \epsilon_{it}$, where ν_{it} is a mean zero process uncorrelated with x_{it}) are not autocorrelated, then differencing introduces correlation: $\Delta \epsilon_{it}$ has a term in common with $\Delta \epsilon_{i,t-1}$, so $E[\Delta \epsilon_{it} \Delta \epsilon_{i,t-1}] = E[-\epsilon_{i,t-1}^2] = -\sigma_\epsilon^2$. Furthermore, any attenuation bias is magnified by the first-difference transformation: $\text{plim}(\hat{\beta}_{FD}) = \beta \sigma_{\Delta x}^2 / (\sigma_{\Delta x}^2 + \sigma_{\Delta \nu}^2)$ where typically $\sigma_{\Delta \nu}^2 = 2\sigma_\nu^2$ but $\sigma_{\Delta x}^2 < 2\sigma_x^2$. Autocorrelation in the mismeasured variables further reduces the signal to noise ratio and increases the attenuation bias.

⁸Cf. Kalwij and Verschoor (2007), who first show a simpler linear model in differences to remove the unobserved effects and then estimate interaction models with unobserved effects reintroduced.

unlikely to work well for values distant from the mean and can easily take on implausible values (e.g. $\hat{\varepsilon}_{it}^{Hy} > 0$) for some combinations of income and inequality. Any version of the linear model will only poorly approximate the non-linear shape. Yet for these estimates to be relevant for any particular country or region, we are precisely interested in temporal and/or regional elasticities and not just the overall cross-country mean elasticities.

2.2 Alternative approaches: non-linear models of poverty

We propose modeling the conditional expectation of the poverty headcount using non-linear parametric fractional response models. To the best of our knowledge, we are the first to use such an approach for estimating poverty elasticities.

Some weaknesses of the linear specifications have been highlighted before. Chambers and Dhongde (2011) argue that the functional form is the source of the non-linearity. They estimate non-parametric models of the conditional mean of the headcount ($H_{it} = m(\bar{y}_{it}, G_{it}) + \epsilon_{it}$) and then obtain average elasticities. However, non-parametric techniques are often inefficient, cannot handle many covariates (due to the curse of dimensionality), and cannot easily deal with measurement error. It is generally difficult to test more involved hypotheses in a non-parametric framework. Nevertheless, we fully agree with Chambers and Dhongde (2011) that non-linear estimation of the conditional mean matters.

We follow Papke and Wooldridge (1996), who first suggested modeling proportions using fractional logit or fractional probit to estimate models of the form $E[y_i | \mathbf{x}_i] = F(\mathbf{x}'_i \boldsymbol{\beta})$, where $y_i \in [0, 1]$ and $F(\cdot)$ is the logistic or normal cdf. Applying this approach to our problem, we may approximate eq. (1) with

$$E[H_{it} | \bar{y}_{it}, G_{it}] = \Phi(\alpha + \beta \ln \bar{y}_{it} + \gamma \ln G_{it}) \quad \text{for } i = 1, \dots, N; t = 1, \dots, T \quad (6)$$

where $\ln z$ is absorbed into the constant and we also take the logarithm of the Gini coefficient (mostly for convenience). Naturally, we expect $\beta < 0$ and $\gamma > 0$. In motivating this model, we temporarily assume away all econometric complications such as unobserved heterogeneity, endogeneity and unbalancedness. These assumptions are relaxed in the next subsection.

Since $F(\cdot)$ is invertible, it can be used as a ‘link function’ in the spirit of the GLM literature (e.g. MacCullagh and Nelder, 1989). Thus, we may also write eq. (6) as $\Phi^{-1}(E[H_{it} | \bar{y}_{it}, G_{it}]) = \alpha + \beta \ln \bar{y}_{it} + \gamma \ln G_{it}$. In other words, the inverse normal cdf linearizes the conditional mean. Figures B-1 and B-2 in Appendix B use this property to plot $\Phi^{-1}(H_{it})$ against $\ln \bar{y}_{it}$ and $\ln G_{it}$ for each region, including a regression line. The result is striking. This simple transformation is extremely successful in removing the intrinsic non-linearity of the poverty headcount.

It is now straightforward to define the estimated income elasticity as

$$\hat{\varepsilon}_{it}^{Hy} = \frac{\partial \hat{E}[H_{it} | \bar{y}_{it}, G_{it}]}{\partial \bar{y}_{it}} \frac{\bar{y}_{it}}{\hat{E}[H_{it} | \bar{y}_{it}, G_{it}]} = \hat{\beta} \times \lambda \left(\hat{\alpha} + \hat{\beta} \ln \bar{y}_{it} + \hat{\gamma} \ln G_{it} \right) \quad (7)$$

and the estimated Gini elasticity as

$$\hat{\varepsilon}_{it}^{HG} = \frac{\partial \hat{E}[H_{it} | \bar{y}_{it}, G_{it}]}{\partial G_{it}} \frac{G_{it}}{\hat{E}[H_{it} | \bar{y}_{it}, G_{it}]} = \hat{\gamma} \times \lambda \left(\hat{\alpha} + \hat{\beta} \ln \bar{y}_{it} + \hat{\gamma} \ln G_{it} \right). \quad (8)$$

Contrary to the linear model, the elasticities in eqs. 7 and 8 closely mimic the properties and structure of the analytical elasticities in eqs. 2 and 3. The non-linearity arises simply from the bounded functional form and is not artificially captured by interaction terms. However, it is important to note that we *do not require log normality* but implicitly assume that an unspecified two-parameter distribution sufficiently describes the poverty headcount up to statistical error. We simply derived a more natural model of the poverty headcount as a function of income and inequality which happens to look a lot like its theoretical counterpart.

This approach has several advantages. We directly take the information provided by poverty levels into account, the model predictions of the poverty headcount ratio will strictly lie in the unit interval, and the same model is able to estimate both elasticities and semi-elasticities. As a result of respecting the bounded nature of the headcount, the elasticities will also approach zero when the inverse Mills ratio becomes vanishingly small and are likely to closely approximate values further away from the mean of the covariates. Hence, they share important properties of their theoretical counterparts based on the log normal assumption. Furthermore, we can directly predict the poverty headcount for interesting combinations of income and inequality rather than going the roundabout way of estimating the elasticities in one model and then projecting poverty separately.

2.3 Econometrics of fractional response models

Since there is no free lunch in econometrics, the apparent gains over the linear approach must come at a price. In non-linear models it is generally harder to deal with unobserved heterogeneity, measurement error and unbalancedness. While until a few years ago there was relatively little research on this issue, we can now draw on an increasingly well developed framework. Papke and Wooldridge (2008) extend fractional response models to balanced panels with unobserved heterogeneity and endogenous covariates, Wooldridge (2010a) develops the theory for unbalanced panels, and Wooldridge (2012) derives a general set-up for one-step estimators in non-linear models. Other contributions to the field are Loudermilk (2007), Cook, Kieschnick, and McCullough (2008), Ramalho, Ramalho, and Murteira (2011), and Tiwari and Palm (2011).

To simplify the exposition, we stack the coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)'$ and covariates $\mathbf{x}_{it} = (x_{it,1}, x_{it,2}, \dots, x_{it,k})'$. The ideal model we would like to estimate is

$$E[H_{it}|\mathbf{x}_i, \mu_i] = E[H_{it}|\mathbf{x}_{it}, \mu_i] = \Phi(\mathbf{x}'_{it}\boldsymbol{\beta} + \mu_i) \quad \text{for } i = 1, \dots, N; t = 1, \dots, T \quad (9)$$

where $\mathbf{x}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{iT})$ are the covariates in all periods. We assume that the covariates are strictly exogenous conditionally on the unobserved effects (μ_i), and that the panel is balanced. We introduced unobserved country-level effects to capture time-persistent differences in measurement or deviations from a two-parameter distribution, which may be arbitrarily correlated with the elements in \mathbf{x}_{it} .

The key problem with such an approach is that the unobserved effects are not identified when T is fixed and $N \rightarrow \infty$, leading to biased estimates of the parameter vector. This is the incidental parameters problem of Neyman and Scott (1948).⁹ In addition, the partial effects needed for calculating the elasticities are not identified either. Papke

⁹This bias tends to become small as T gets large, but there are no benchmark simulations for the fractional probit case that we know of and in our case \bar{T} is small. Papke and Wooldridge (2008) explain why replacing the standard normal cdf by the logistic cdf is not a good solution for this problem.

and Wooldridge (2008) suggest to solve this problem by imposing some structure on the correlation between the unobserved effects and the covariates using a device developed by Mundlak (1978) and Chamberlain (1984). This approach is generally known as correlated random effects (CRE). Concretely, we let

$$\mu_i | (\mathbf{x}_{i1}, \dots, \mathbf{x}_{iT}) \sim \mathcal{N}(\varphi + \bar{\mathbf{x}}_i' \boldsymbol{\theta}, \sigma_u^2) \quad (10)$$

where $\bar{\mathbf{x}}_i = T^{-1} \sum_{t=1}^T \mathbf{x}_{it}$ is the time average of all the included time-varying regressors, \mathbf{x}_{it} no longer contains a constant, and $u_i = \mu_i - \varphi - \bar{\mathbf{x}}_i' \boldsymbol{\theta}$ with $u_i | (\mathbf{x}_{i1}, \dots, \mathbf{x}_{iT}) \sim \mathcal{N}(0, \sigma_u^2)$. The covariates are still strictly exogenous conditionally on the unobserved effects. In linear models, this specification is equivalent to the traditional ‘fixed effects’ model and thus, in terms of accounting for unobserved effects, achieves the same aim as specifying a difference equation.

Plugging eq. (10) into eq. (9), we can rewrite our model of interest as

$$E[H_{it} | \mathbf{x}_i, u_i] = \Phi(\varphi + \mathbf{x}'_{it} \boldsymbol{\beta} + \bar{\mathbf{x}}_i' \boldsymbol{\theta} + u_i) \quad (11)$$

$$E[H_{it} | \mathbf{x}_i] = E[\Phi(\varphi + \mathbf{x}'_{it} \boldsymbol{\beta} + \bar{\mathbf{x}}_i' \boldsymbol{\theta} + u_i) | \mathbf{x}_i] = \Phi(\varphi_u + \mathbf{x}'_{it} \boldsymbol{\beta}_u + \bar{\mathbf{x}}_i' \boldsymbol{\theta}_u) \quad (12)$$

where the subscript u denotes scaling of the coefficients by the factor $(1 + \sigma_u^2)^{-1/2}$. Going from eq. (11) to eq. (12) applies iterated expectations and the last equality follows from mixing (compounding) independent mean-zero normals.

If these assumptions hold, then the scaled coefficients and average partial effects (APEs) of all time-varying covariates are identified. However, survey-specific (non-classical) measurement error in income is likely to lead to overestimating the income elasticity in absolute value (Ravallion and Chen, 1997). In addition, classical measurement error may attenuate the income coefficient and thus work in the opposite direction. Suppose we do not observe true income but $\ln \bar{y}_{it} = \ln \bar{y}_{it}^* + v_{it}$, where $\ln \bar{y}_{it}^*$ is the true value of log mean income/ expenditure and v_{it} is a composite error process with a classical and a non-classical component. We can view this as an omitted variable problem. Going back to the simplest model, we have $E[H_{it} | \mathbf{x}_{1it}, \bar{y}_{it}^*, \mu_i] \neq E[H_{it} | \mathbf{x}_{1it}, \bar{y}_{it}, \mu_i] = \Phi(\mathbf{x}'_{1it} \boldsymbol{\beta} + \psi(\ln \bar{y}_{it}^* + v_{it}) + \mu_i) = \Phi(\mathbf{x}'_{1it} \boldsymbol{\beta} + \psi \ln \bar{y}_{it} - \psi v_{it} + \mu_i)$, where \mathbf{x}_{1it} is \mathbf{x}_{it} without the mismeasured $\ln \bar{y}_{it}$, and v_{it} is also potentially correlated with the time-constant unobserved effects ($\text{cov}(v_{it}, \mu_i) \neq 0$). Inference using observed income will lead to underestimating or overestimating the effect depending on which type of error is stronger.

Building on Rivers and Vuong (1988) and the general result from Blundell and Powell (2004), Papke and Wooldridge (2008) suggest a simple two-step control function estimator for such endogeneity problems. Provided we have an $m \times 1$ vector of time-varying instruments (\mathbf{z}_{it}) that are relevant but not correlated with v_{it} , we can estimate a log-linear model $\ln \bar{y}_{it} = \varphi_1 + \mathbf{x}'_{1it} \boldsymbol{\beta}_1 + \mathbf{z}'_{it} \boldsymbol{\gamma}_1 + \bar{\mathbf{x}}_i' \boldsymbol{\theta}_1 + v_{it}$ in the first step and then obtain the residuals \hat{v}_{it} . It is important to note here that we redefine $\bar{\mathbf{x}}_i = (\bar{\mathbf{x}}_{1i}, \bar{\mathbf{z}}_i)$; that is, $\bar{\mathbf{x}}_i$ now contains the time averages of *all* strictly exogenous variables including the instruments. In the second step, we specify the residual-augmented model of $E[H_{it} | \mathbf{x}_{1it}, \mathbf{z}_{it}, \bar{y}_{it}, v_{it}] = \Phi(\varphi_k + \mathbf{x}'_{1it} \boldsymbol{\beta}_k + \bar{\mathbf{x}}_i' \boldsymbol{\theta}_k + \psi_k \ln \bar{y}_{it} + \rho_k \hat{v}_{it})$. Here, too, the Chamberlain-Mundlak device concerns both the strictly exogenous variables and the excluded instruments. The subscript k denotes a new scale factor $(1 + \sigma_k^2)^{-1/2}$. The solution is to simply condition on an estimate of the omitted variable (\hat{v}_{it}). A test of $\hat{\rho}_k = 0$ corresponds to a test of exogeneity and does not depend on the first step under the null (see Hausman, 1978).

The asymptotic standard errors must be adjusted for the uncertainty of the first step and can be derived via the delta method or approximated with the panel bootstrap.

Accounting for unbalancedness adds another layer of complication. Contrary to linear models with CRE, where most estimators need little or no practical adjustments to work with unbalanced panels, estimates from non-linear CRE models are inconsistent if applied to unbalanced panel data. The main problem is that both the estimates and the variances of the correlated random effects can differ depending on the sample size. Wooldridge (2010a) proposes to also directly model this dependence. Assuming that selection is conditionally independent, we can extend our model by letting the unobserved effects vary depending on the panel size: $E[\mu_i|\mathbf{w}_i] = \sum_{r=1}^T \delta_{T_i,r} \varphi_{kr} + \sum_{r=1}^T \delta_{T_i,r} \bar{\mathbf{x}}_i' \boldsymbol{\theta}_{kr}$, where \mathbf{w}_i is a vector of functions of the conditioning variables sufficient to represent the distribution $D[\mu_i|(s_{it}, s_{it}\mathbf{x}_{1it}, s_{it}\mathbf{z}_{it}, s_{it} \ln \bar{y}_{it})] = D[\mu_i|\mathbf{w}_i]$; further, s_{it} is a selection indicator and $\delta_{T_i,r}$ is the Kronecker delta which is equal to unity if $T_i = r$ and zero otherwise. The coefficients are still scaled by $(1 + \sigma_k^2)^{-1/2}$. Without further assumptions, this implies that we cannot use the observations where $T_i = 1$ as these have no separately identifiable panel dimension. Hence, they drop out of the estimating equation. Additionally, we also let the conditional variance depend on the panel size such that $\text{var}(\mu_i|\mathbf{w}_i) = \sigma_\mu^2 \exp(\sum_{r=2}^{T-1} \delta_{T_i,r} \omega_r)$, where the omegas represent unknown variance parameters and σ_μ^2 is the variance of the unobserved heterogeneity when $T_i = T$.^{10,11} The result is a variable scale factor: $(1 + \sigma_\mu^2 \exp(\sum_{r=2}^{T-1} \delta_{T_i,r} \omega_r))^{-1/2}$.

A convenient reparameterization arises when we treat the overall variance as heteroskedastic and assume that $D[\mu_i|\mathbf{w}_i]$ is normal (see Wooldridge, 2010a). Dividing the conditional expectation by $\exp(\sum_{r=2}^{T-1} \delta_{T_i,r} \tilde{\omega}_r)$, where $\tilde{\omega}_r$ denotes a new set of unknown parameters for the *overall* variance, we again obtain a constant scale factor. Then, the reparameterized two-step unbalanced CRE model is

$$E[H_{it}|\mathbf{x}_i, \nu_{it}, \mathbf{w}_i] = \Phi \left(\frac{\mathbf{x}'_{1it} \boldsymbol{\beta}_h + \psi_h \ln \bar{y}_{it} + \rho_h \hat{\nu}_{it} + \sum_{r=2}^T \delta_{T_i,r} \varphi_{hr} + \sum_{r=2}^T \delta_{T_i,r} \bar{\mathbf{x}}_i' \boldsymbol{\theta}_{hr}}{\exp \left(\sum_{r=2}^{T-1} \delta_{T_i,r} \tilde{\omega}_r \right)^{1/2}} \right) \quad (13)$$

where the explanatory variables at t are $(1, \mathbf{x}'_{1it}, \ln \bar{y}_{it}, \hat{\nu}_{it}, \delta_{T_i,2} \bar{\mathbf{x}}_i', \dots, \delta_{T_i,T} \bar{\mathbf{x}}_i')$ and the variance depends on a set of dummy variables shifting from $T_i = 2$ to $T_i = T - 1$ with $T_i = T$ as the base. The subscript h denotes the new scale factor. The specification nests the balanced case. If the panel is balanced, the numerator has only one set of time averages and a constant in addition to the time-varying covariates, while the denominator is unity. The first estimation step is also augmented to accommodate the varying panel sizes. We obtain the residuals via $\hat{\nu}_{it} = \ln \bar{y}_{it} - \mathbf{z}'_{it} \hat{\gamma}_1 - \mathbf{x}'_{1it} \hat{\beta}_1 - \sum_{r=2}^T \delta_{T_i,r} \hat{\varphi}_{1r} - \sum_{r=2}^T \delta_{T_i,r} \bar{\mathbf{x}}_i' \hat{\boldsymbol{\theta}}_{1r}$, where \mathbf{x}'_{1it} includes time dummies and $\bar{\mathbf{x}}_i'$ their time averages. The heterogeneity related to the instruments (the $\bar{\mathbf{z}}_i$ time averages in $\bar{\mathbf{x}}_i$) is interacted with the panel size dummies and thus enters the first step flexibly. This, too, simplifies back to the earlier result in

¹⁰As Wooldridge (2010a) points out, it is possible to model the conditional expectation and variance even more flexibly by allowing for additional intercepts, trends, and variances/covariances to approximate the non-parametric relationship from Altonji and Matzkin (2005).

¹¹We could also let the conditional variance depend on inequality (which can be motivated by assuming log normality of income). This relaxes an implicit assumption, namely that the marginal proportional rate of substitution ($MPRS_t = -\hat{\varepsilon}_t^{H\bar{y}}/\hat{\varepsilon}_t^{HG}$) is constant. The models fit marginally better but the substantive implications change very little. Additional results are available on request.

the balanced case if we remove the redundant variables (i.e. averages of the time effects).

Obviously this is a complicated model to fit but it can be estimated by any software that has a heteroskedastic probit implementation without any restrictions on the dependent variable (Wooldridge, 2010a).¹² Since this is a quasi-maximum likelihood estimator (QMLE), the standard errors based on the inverse information matrix will be too conservative and need to be adjusted for clustering at the country level (for details see Papke and Wooldridge, 1996; 2008; Wooldridge, 2010b). Apart from the assumptions made to restrict the unobserved heterogeneity and endogeneity, fractional probit only requires correct specification of the conditional mean irrespective of the true distribution of the dependent variable (Gourieroux, Monfort, and Trognon, 1984). Hence, it is as robust as non-linear least squares but potentially more efficient.

We still need to define the average partial effects (APEs) and elasticities. Both can be derived from the average structural function (ASF) computed over the selected sample (see Blundell and Powell, 2004; Wooldridge, 2010a) which makes clear that *only the APEs of time-varying covariates are in fact identified*. Let the linear predictors inside the cumulative normal be $\mathbf{m}'_{it1}\hat{\boldsymbol{\xi}}_1$ for the main equation and $\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2$ for the variance equation. Then, $\widehat{\text{ASF}}(\mathbf{x}_t) = N^{-1} \sum_{i=1}^N \Phi \left(\mathbf{m}'_{it1}\hat{\boldsymbol{\xi}}_1 / \exp(\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2)^{1/2} \right)$, where \mathbf{x}_t refers to all time-varying covariates including mismeasured income, and the coefficients are the scaled QMLE estimates. We always need to average over the cross-section dimension in order to get rid of the unobserved effects, varying panel sizes, and endogeneity/ measurement error. The APE at time t of a particular continuous variable is simply the derivative of the ASF with respect to that variable. We usually plug in interesting values for \mathbf{x}_t and obtain the APEs assuming the entire sample shares these characteristics.

Analogously, the elasticity with respect to any $x_k \in \mathbf{x}_t$ (provided that x_k is in logs and does not show up in the variance equation) is

$$\hat{\varepsilon}_t^{Hx_k} = \hat{\beta}_k \times N^{-1} \sum_{i=1}^N \exp \left(-\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2/2 \right) \lambda \left(\mathbf{m}'_{it1}\hat{\boldsymbol{\xi}}_1 / \exp(\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2)^{1/2} \right) \quad (14)$$

and the semi-elasticity ($\hat{\eta}_t^{Hx_k}$) is the derivative of the ASF with respect to x_k ; that is, the average partial effect (APE). We also average over time in order to obtain a scale factor.

The basic structure is exactly the same as in the simpler versions derived in the previous section with the addition of a variance equation adjusting for the degree of unbalancedness. If the panel is balanced, the non-redundant sums inside the linear predictors simplify and we again obtain the CRE analogues of eqs. (7) and (8).

3 Data

Based on the World Bank's PovcalNet database¹³, we compile a new and comprehensive data set consisting of 809 nationally-representative surveys spanning 124 countries from 1981 to 2010.¹⁴ Smaller panels of this data have been used in previous studies (e.g. Chambers and Dhongde, 2011; Kalwij and Verschoor, 2007; Adams, 2004) and the World

¹²However, most implementations (e.g. Stata's `hetprobit`) only allow binary dependent variables. We implement the estimator in a new module called `fhetprobit` (see Bluhm, 2014, *forthcoming*).

¹³The data is publicly available at <http://iresearch.worldbank.org/PovcalNet> (last accessed May 20 2013, last updated April 18 2013).

¹⁴Supporting materials and the panel data set are available at www.richard-bluhm.com/data/.

Bank's methodology is described in more detail in Chen and Ravallion (2010). Here we only briefly summarize the main features.

All data originate from household surveys. Our primary measure of poverty is the headcount index (H_{it}) given a poverty line (z) of \$2 a day or \$60.83 a month. The poverty headcount at the \$1.25 a day or \$38 a month poverty line is used as a secondary measure. The data also contains measures of mean monthly per capita household income or consumption expenditures (\bar{y}_{it})¹⁵, the Gini coefficient of inequality (G_{it}) for income or consumption expenditures, the surveyed population (pop_{it})¹⁶ and a set of indicators distinguishing if the survey uses income or consumption as a welfare measure and whether unit-level or grouped-level (deciles or finer quantiles) data are used. About 63% of the data come from expenditure surveys and about 74% are estimated from grouped data. All monetary quantities are in constant international dollars at 2005 PPP-adjusted prices.

Some countries¹⁷ do not conduct nationally representative surveys but instead report urban and rural data separately. We simply weigh the poverty and income data using the relative urban/rural population shares to construct national series. Since the Gini is not subgroup decomposable, we employ a mixing of two log normal distributions approximation to estimate a national Gini coefficient (Young, 2011).¹⁸ If only one urban or rural survey is available in any given year, we usually drop the survey, except in the case of Argentina where urbanization is near or above 90% for most of the sampled period and we thus consider the urban series nationally representative. This results in an unbalanced panel of 124 countries, with an average time series (\bar{T}) of about 6.5 surveys for a total of 809 observations, spanning 30 years. Table 1 provides summary statistics for the entire panel. In Appendix B, Table B-1 presents summary statistics by region, and List B-1 lists the countries and the corresponding numbers of surveys in the sample.

Table 1 – Summary statistics

	Mean	Std. Deviation	Min	Max	N
<i>Main variables</i>					
H_{it} – Headcount (\$2)	0.303	0.286	.0002	.9845	809
H_{it} – Headcount (\$1.25)	0.182	0.219	.0002	.9255	789
G_{it} – Gini coefficient	0.424	0.102	.2096	.7433	809
\bar{y}_{it} – Mean income or expenditure in \$ per month	194.59	125.90	14.93	766.78	809
PCE_{it}^P – Consumption (PWT) in \$ per month	338.64	234.59	14.39	1231.21	795
<i>Survey type dummies</i>					
Consumption (Grouped)	0.611	0.488	0	1	809
Consumption (Unit)	0.015	0.121	0	1	809
Income (Grouped)	0.132	0.339	0	1	809
Income (Unit)	0.242	0.429	0	1	809

For the linear models in log differences, we also construct a second data set of ‘poverty spells’ at the country level (data in changes). Manually defining poverty spells serves two

¹⁵Computed as a simple per capita average without equivalence scaling.

¹⁶Several entries in the PovcalNet population data are clearly mistaken, zero or missing. We fix the values/series using data from the World Development Indicators.

¹⁷China, India and Indonesia.

¹⁸PovcalNet omits weighting some recent data. To use a single consistent method, we apply Young's formula in all cases where separate urban and rural surveys are combined. The approximation is very accurate. The formula is $G = \sum_{i=1}^K \sum_{j=1}^K \frac{s_i s_j \bar{y}_i}{\bar{Y}} \left(2K \left[\frac{\ln \bar{y}_i - \ln \bar{y}_j + \frac{1}{2} \sigma_i^2 + \frac{1}{2} \sigma_j^2}{(\sigma_i^2 + \sigma_j^2)^{1/2}} \right] - 1 \right)$ where K is the total number of subgroups, s_i is the population share of the i -th subgroup and \bar{Y} is the population-weighted mean income across all subgroups.

purposes. First, we only use data from runs of the same survey type to avoid introducing artificial changes when there is a switch from income to expenditure surveys or vice versa (same for unit and grouped-level data; we assume the grouping does not change). Second, we annualize all differences to mitigate any biases arising from estimating elasticities over very different time periods. This leads to a smaller data set of 648 observations in 104 countries, as differencing requires $T_i \geq 2$. Most studies using poverty differences lose additional observations when the poverty headcount is zero at the beginning or end of a spell (e.g. Kalwij and Verschoor, 2007), implying a relative change in the poverty headcount of negative or positive infinity. This problem does not occur in our data, but more importantly, our approach does *not* require such ad hoc adjustments.

We add per capita consumption data from national accounts to the survey-based panel, which we later use as instruments for mean survey income and require for the poverty projections. Personal consumption expenditures are retrieved from both from the World Development Indicators (WDI) and Penn World Table 7.1 (PWT), denoted PCE_{it}^W and PCE_{it}^P respectively.¹⁹ The PWT version is preferred in the estimations; for the poverty projections a “merged” series (PCE_{it}) is constructed using the WDI series as a benchmark but replacing it with PWT data if coverage over 1981-2010 is better. Both series are in constant 2005 prices, but the PWT adjust the original 2005 ICP data and interpolate differently between benchmark years (Deaton and Heston, 2010). For the projections, we also use estimates of the total population from 2010 to 2030 based on the medium fertility variant from the World Population Prospects (the 2010 revision).

4 Results

4.1 Fractional response models

Table 2 presents our main results, with each specification progressively addressing more estimation issues (unobserved effects, unbalancedness and measurement error). In all specifications, we include time averages to proxy for time invariant measurement differences across different countries (unobserved effects). We also include survey type dummies (consumption or income, grouped or unit data) as reported poverty is typically lower in income surveys than consumption surveys and the availability of grouped versus unit-level data in PovcalNet may proxy for other systematic survey differences. In addition, time dummies allow for unspecified common year effects.

Column (1) includes correlated random effects but entirely ignores unbalancedness. As expected, the coefficient on average income is negative and the coefficient on inequality is positive. Since the estimated coefficients are arbitrarily scaled, the adjacent column reports the average partial effects (APEs) and we report the scale factor separately in the bottom panel. The APEs in column (1) imply the following semi-elasticities. On average, one *percent* income growth leads to a 0.284 *percentage point* reduction in the population that is poor and a corresponding one *percent* increase in inequality leads to a 0.232 *percentage point* increase in poverty. Turning to elasticity concepts, the average income elasticity across the entire estimation sample is about -1.83 ($SE = 0.084$) and the average Gini elasticity is about 1.5 ($SE = 0.167$). Hence, a one *percent* increase in

¹⁹Monthly PCE_{it}^P is computed as $(kc_{it}/100 \times rgdpl_{it}/12)$, where kc_{it} is the consumption share and $rgdpl_{it}$ is GDP per capita (Laspeyres) in 2005 constant prices. Similarly, PCE_{it}^W is household final consumption expenditure in 2005 prices divided by the population and converted to monthly figures.

income leads to about a 1.83 *percent* decrease in poverty. These two estimates are located near the lower bound of the results typically found in the literature.²⁰

Table 2 – Fractional probit models (QMLE) – Dependent variable: H_{it} , \$2 a day

	(1)		(2)		(3)	
	Regular H_{it}	APEs	Unbalanced H_{it}	APEs	Unbalanced + H_{it}	Two-Step APEs
$\ln \bar{y}_{it}$	-1.263 (0.054)	-0.284 (0.012)	-0.880 (0.048)	-0.281 (0.011)	-1.049 (0.198)	-0.339 (0.035)
$\ln G_{it}$	1.032 (0.114)	0.232 (0.026)	0.786 (0.098)	0.251 (0.026)	0.775 (0.163)	0.251 (0.032)
$\hat{\nu}_{it}$					0.133 (0.113)	
CRE (Corr. Rand. Effects)	Yes		Yes		Yes	
Survey type dummies	Yes		Yes		Yes	
Time dummies	Yes		Yes		Yes	
Panel size dummies	No		Yes		Yes	
Panel size dummies \times CRE	No		Yes		Yes	
Variance equation	No		Yes		Yes	
Scale factor	0.225		0.319		0.323	
$N \times \bar{T}$	789		789		775	
N	104		104		103	
pseudo R^2	0.992		0.996		0.997	
$\ln \mathcal{L}$	-219.3		-315.6		-313.4	
\sqrt{MSE}	0.0355		0.0238		0.0235	

Notes: 20 observations with $T_i = 1$ not used in estimation. In models (1) and (2), the standard errors of the coefficients are robust to clustering at the country level and the standard errors of the APEs are computed via the delta method. We include the time averages of the survey type and time dummies in (2) and (3), but constrain their coefficients to be equal across the panel sizes. The standard errors of the coefficients and the APEs in model (3) account for the first stage estimation step with a panel bootstrap using 999 bootstrap replications. The linear projection in the first stage uses $\ln PCE_{it}^P$ as an instrument for $\ln \bar{y}_{it}$. The first-stage cluster-robust F-statistic in (3) is 28.05. Model (3) also excludes West Bank and Gaza entirely (2 observations) and 12 observations from ECA countries pre-1990 for lack of PCE data.

Our first specification could be biased due to the strong unbalancedness of the panel and the presence of time-varying measurement error in income and inequality. Column (2) addresses unbalancedness by including panel size dummies, interactions of the time averages with the panel size dummies, and a separate variance equation. We consider this our best specification *without* correcting for measurement error. The substantive conclusions change very little. The APE of income is virtually unchanged and the APE of inequality increases by about one standard error. Clearly, varying panel sizes introduce little bias on average. Nonetheless, they may still have a non-negligible effect on the (semi-)elasticities at particular points in time.

²⁰The typical range for the income elasticity in earlier studies is from about 2 to 5, while the range for the Gini elasticity is much wider. Newer studies suggest the income elasticity is closer to 2. This is largely owed to parameter instability in linear approximations and changing data coverage. However, as Chambers and Dhongde (2011) report, estimates of the income elasticity using the 2005 PPPs are also universally lower (in absolute value) than estimates based on the earlier 1993 PPPs.

Our preferred specification, column (3), is the empirical counterpart of the two-step estimator presented in eq. (13). We account for measurement error in income by instrumenting survey mean incomes or expenditures with per capita consumption from the national accounts (PCE_{it}^P). The main identifying assumption is that any measurement error in per capita consumption from the national accounts is orthogonal to survey-based measurement error in income or expenditures. As both expenditure measures are estimated very differently in practice, this is a plausible and common identification strategy (Ravallion, 2001). Figure B-3 in Appendix B highlights the strength of the first stage relationship. It shows a partial regression plot of mean incomes or expenditures from the surveys against per capita consumption from the National Accounts, after taking out the variation in the Gini, the time averages of the Gini and PCE_{it}^P , panel size dummies, survey dummies and time dummies.

Ignoring the first-stage variability of income, we find some evidence of measurement error in income (cluster robust t-stat ≈ 1.69).²¹ After accounting for the two-step nature the evidence for measurement error is considerably weaker (panel bootstrap t-stat ≈ 1.18). The APE of income, on the other hand, is larger in absolute value, hence we conclude that the coefficient of income in models (1) and (2) is attenuated towards zero. These results suggest that classical attenuation bias is more of a problem than systematic survey bias due to under-reporting or over-reporting of incomes and expenditures. Since the biases arising from these two types of measurement error are likely to run in opposing directions, our estimates imply that they are not fully offsetting as in Ravallion and Chen (1997).

The unbiased average-income semi-elasticity of poverty estimated in column (3) is considerably larger in absolute value than in the previous two specifications. A one percent increase in incomes leads to a 0.339 percentage point reduction in the population that is poor. Likewise, the average elasticity increased ($\bar{\varepsilon}^{Hy} \approx 2.21$, SE = 0.156) and is now closer to the conventional estimates of about two. The inequality elasticity remains about the same ($\bar{\varepsilon}^{HG} \approx 1.64$, SE = 0.188). In fact, there may also be non-negligible measurement error in observed inequality as measured by the Gini coefficient. However, since inequality is most often estimated from household surveys and estimates based on alternative sources such as tax records are not available on a cross-country basis, we are lacking a corresponding instrument for the Gini coefficient.

Is there evidence of additional non-linearity missing in the non-linear functional form of the poverty headcount? To examine this possibility, we add squares ($\mathbf{m}'_{it1} \hat{\boldsymbol{\xi}}_1$)² and cubes ($\mathbf{m}'_{it1} \hat{\boldsymbol{\xi}}_1$)³ of the linear predictors in columns (1) and (2) for a RESET-type test as suggested by Papke and Wooldridge (1996). In the most basic column (1), this yields a robust χ^2_2 -statistic of 4.65 for a p -value of 0.098, giving no reason for concern. For the heteroskedastic model in column (2), the RESET test provides some evidence of additional non-linearity (robust $\chi^2_2 = 9.15$, $p > 0.01$). However, there is no theoretical reason to expect additional powers or interactions of income and inequality to enter the model.

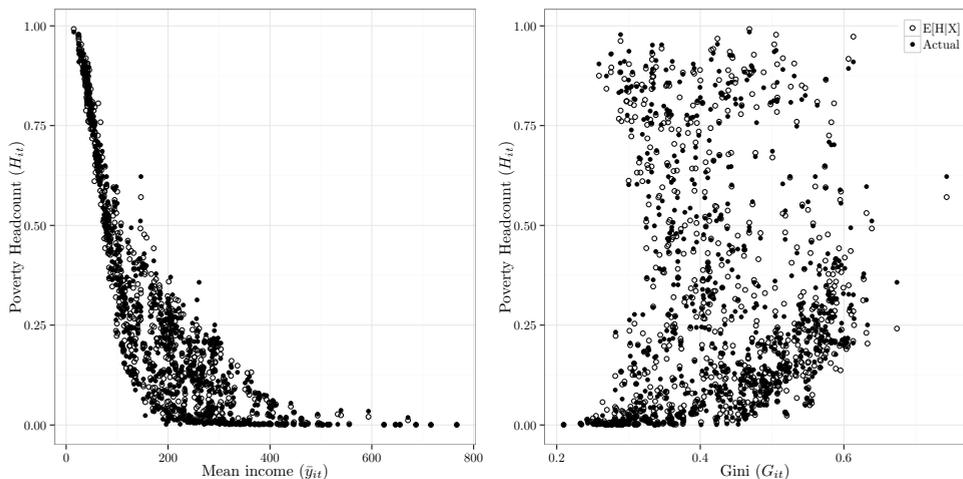
At first sight, all three models may suggest an only moderately larger effect of income growth when holding inequality constant relative to changes in inequality, and *vice versa*. This does not imply that both variables have the same scope for change or have to change independently for that matter. We estimate the impact of each component and not its contribution to overall poverty reduction (for estimates of the contributions see Kraay, 2006). While there is substantial variation in inequality, it shows no systematic

²¹As mentioned earlier, the Hausman (1978) test does not depend on the first stage under the null.

trend over the sample period from 1981 to 2010.²² In contrast, incomes and expenditures have increased substantially in all regions over the same time span (see Table B-5 in Appendix B). Moreover, the effect of income growth is not constant. In these models, it depends strongly on the levels of inequality and income. Thus there is a ‘double dividend’ to improvements in distribution (Bourguignon, 2003) and substantial heterogeneity in the estimated poverty (semi-)elasticities across time and space – an issue to which we return shortly.

Perhaps the most striking fact about all three specifications is how well they fit. For more intuitive comparisons, the last row shows the square root of the mean squared residual for each column – a model metric suggested by Wooldridge (2010a). Already in the first model, we predict the observed poverty headcount for each country-year with about three and a half percentage points accuracy and with better than two and a half percentage points accuracy in the preferred specification. This truly reflects an identity relationship. A simple pseudo- R^2 measure of the correlation between the observed and fitted values for models (1) to (3) suggests near perfect fit ($R^2 > 0.99$). Figure 1 illustrates this point and shows the shape of the estimated effects. Using our preferred specification, we plot both the observed headcount and the predicted headcount over the range of observed mean income or expenditures (left panel) and inequality (right panel). The quality of the non-linear approach is readily apparent as the fit is very close at either bound (near unity or near zero) and the model does not predict nonsensical values. Further, the variation in the observed values is completely covered by the model predictions. In linear models, neither of these two outcomes is guaranteed.

Figure 1 – Data versus fitted values, preferred specification, \$2 a day



For comparison, Table B-2 in Appendix B reproduces the linear approach of the previous literature using the data in levels and the poverty spell data in differences. The differences in the estimated *average* elasticities are not large, as is typical for comparisons between linear and non-linear approaches. We do not discuss these results in detail since they suffer from the expected problems (see Section 2). First, when switching from fixed

²²In a simple regression of the Gini coefficient on time, we cannot reject the null hypothesis that the time trend is zero (cluster robust t-stat ≈ 0.07 and $p > 0.94$).

effects to annualized differences in the simple models with only income and inequality, measurement error increases and attenuates the income coefficient. Second, the models with interaction terms do not fit nearly as well as those reported earlier and many coefficients are insignificant. Third, the two-step GMM results for the interaction models are unstable and not able to convincingly solve the problem of measurement error. The last model, which is in the spirit of the preferred specification of Kalwij and Verschoor (2007), even implies a negative Gini elasticity and all coefficients are estimated with great imprecision. In sum, these models do not perform well in comparison to their fractional response counterparts and are thus unlikely to produce reliable estimates over a wide range of circumstances.

Conversely, the strength of the fractional response approach lies in its ability to deliver much more precise estimates of effects other than the overall mean. Table 3 and Table 4 illustrate this point by estimating the income elasticities and Gini elasticities over different time periods for the six geographic regions in our sample. They are computed according to eq. (14) by plugging in the different time period and region specific means of income ($\ln \bar{y}_{it}$) and inequality ($\ln G_{it}$), and then averaging over the entire sample population. The standard errors of the elasticities are computed via a panel bootstrap and thus account for the uncertainty of the first stage. We present regional and temporal elasticities here but also provide estimates for the semi-elasticities in Table B-3 and Table B-4 in Appendix B for comparison.

Table 3 – Predicted regional income elasticities, preferred specification, \$2 a day

	<i>Time period</i>				
	1981–1989	1990–1994	1995–1999	2000–2004	2005–2010
East Asia and Pacific	-0.991 (0.030)	-1.029 (0.033)	-1.237 (0.055)	-1.139 (0.043)	-1.578 (0.101)
Eastern Europe and Central Asia	-4.358 (0.555)	-2.892 (0.309)	-2.700 (0.277)	-2.846 (0.304)	-3.304 (0.384)
Latin America and Caribbean	-2.284 (0.243)	-2.374 (0.257)	-2.425 (0.271)	-2.349 (0.258)	-2.985 (0.366)
Middle East and North Africa	-2.176 (0.203)	-2.116 (0.188)	-2.024 (0.168)	-1.966 (0.161)	-2.501 (0.246)
South Asia	-0.548 (0.053)	-0.629 (0.048)	-0.810 (0.030)	-1.024 (0.032)	-1.192 (0.046)
Sub-Saharan Africa	-0.831 (0.027)	-0.437 (0.039)	-0.436 (0.040)	-0.592 (0.035)	-0.632 (0.033)

Notes: Standard errors obtained via a panel bootstrap using 999 replications. The predicted elasticities are based on estimated APEs at each region/time-period mean of $\ln \bar{y}_{it}$ and $\ln G_{it}$.

There is considerable regional and temporal heterogeneity in the estimated income elasticities. However, its origins are very mechanical. As the theoretical derivations in Section 2 show and our estimates make clear, the income elasticity is an increasing function of income. In other words, regional heterogeneity of the income elasticity is mainly due to income heterogeneity. More affluent regions (Eastern Europe and Central Asia, Latin America and the Caribbean, and the Middle East and North Africa) have higher income elasticities than poorer regions (East Asia and Pacific, South Asia and Sub-Saharan Africa). Income dynamics over time are also clearly visible. In Eastern Europe and Central Asia, for example, income is comparatively high before the post-communist

transition, sharply collapses throughout the 1990s and then recovers during the 2000s. Compared to earlier results (e.g. Kalwij and Verschoor, 2007), we find markedly higher average income elasticities in more affluent regions and lower elasticities in poorer regions. Throughout Table 3, the standard errors are small compared to the point estimates and *remain* very accurate for regions with more extreme values (e.g. very low income and above average inequality in Sub-Saharan Africa in the 1980s).

Table B-3 in Appendix B presents the region and time specific income semi-elasticities of poverty. There the picture is reversed. Comparatively more affluent regions have fewer people near the poverty line, and thus the poverty reduction potential from a one percent increase in incomes is much smaller in terms of the population lifted out of poverty. This dynamic is again best visible in Eastern Europe and Central Asia, where absolute poverty at the \$2 a day poverty line is almost non-existent just before the post-communist transition and then rises sharply in the 1990s as incomes decline. Correspondingly, the semi-elasticity is close to zero in the 1980s but then it increases as more people fall into poverty. Likewise, the biggest poverty reduction potential in 2005-2010 was in East Asia, South Asia and Sub-Saharan Africa. This highlights an important point. For development policy, we really care more about poverty reduction in terms of the percent of the population lifted out of poverty rather than relative changes in the poverty headcount. Hence, semi-elasticities are the pertinent concept (see also Klasen and Misselhorn, 2008).

Table 4 – Predicted regional Gini elasticities, preferred specification, \$2 a day

	<i>Time period</i>				
	1981–1989	1990–1994	1995–1999	2000–2004	2005–2010
East Asia and Pacific	0.732 (0.105)	0.760 (0.101)	0.914 (0.113)	0.841 (0.108)	1.165 (0.144)
Eastern Europe and Central Asia	3.219 (0.510)	2.136 (0.307)	1.994 (0.283)	2.102 (0.296)	2.440 (0.353)
Latin America and Caribbean	1.687 (0.186)	1.753 (0.198)	1.791 (0.199)	1.735 (0.189)	2.205 (0.269)
Middle East and North Africa	1.607 (0.197)	1.563 (0.198)	1.495 (0.196)	1.452 (0.185)	1.847 (0.253)
South Asia	0.405 (0.093)	0.464 (0.097)	0.598 (0.095)	0.756 (0.107)	0.880 (0.127)
Sub-Saharan Africa	0.614 (0.087)	0.322 (0.055)	0.322 (0.060)	0.437 (0.066)	0.467 (0.069)

Notes: Standard errors obtained via a panel bootstrap using 999 replications. The predicted elasticities are based on estimated APEs at each region/time-period mean of $\ln y_{it}$ and $\ln G_{it}$.

The region and time specific Gini elasticities in Table 4 show where the potential of redistributive policies in terms of proportionate reductions in the poverty headcount was largest over the last three decades. Unsurprisingly, these regions are Eastern Europe and Central Asia, Latin America and the Caribbean, and the Middle East and North Africa – all of which have above average inequality. Sub-Saharan Africa starts out with high inequality in the 1980s²³ but incomes are very low relative to the poverty line, so that the Gini elasticity is small. This is the flip side of the dependency on initial levels: countries

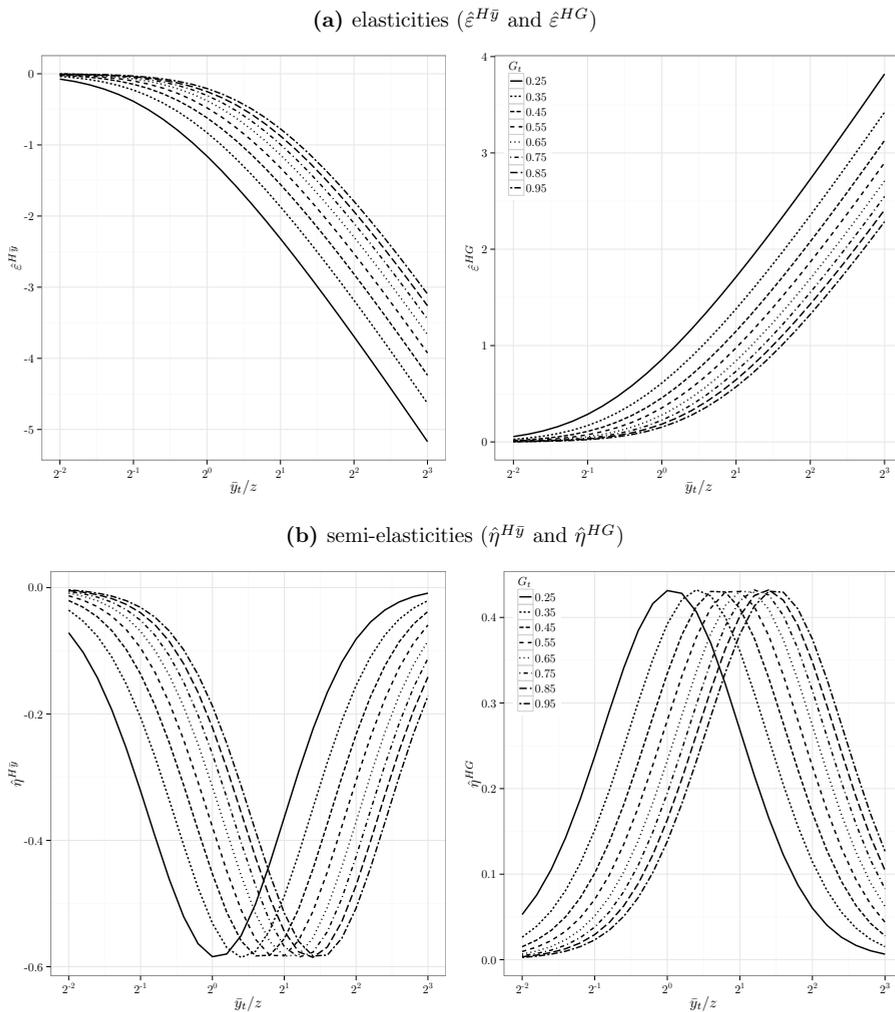
²³The population-weighted mean Gini in the 1980s is 0.4608.

can be so poor and unequal that the immediate effects of equalization and income growth on *relative* changes in the poverty headcount are comparatively small. However, here too, the semi-elasticities presented in Table B-4 in the Appendix help to clarify the picture. There the relative position of poorer and richer countries is reversed. The potential for reducing poverty through redistribution in terms of percent of the population that is poor was larger in poorer regions throughout the entire period from 1981 to 2010.

The ‘double dividend’ of reductions in inequality is illustrated in Figure 2 by graphing the estimated poverty elasticities or semi-elasticities over different combinations of income and inequality. Again, we compute these estimates according to eq. (14) by plugging in the different values for mean income or expenditures ($\ln \bar{y}_{it}$) and inequality ($\ln G_{it}$), and then averaging over the entire sample. As Figure 2a illustrates, on top of the direct poverty alleviating effect of income redistribution towards the poor, a lower level of inequality also raises the income elasticity in absolute value at every point. However, the magnitude of both elasticities is steeply increasing in the level of income; that is, the return to either income growth or equalization is bigger, the higher the income level. This may invite the conclusion that growth matters more at lower levels of income, while redistribution is only important for high income and high inequality countries. This, precisely, is the misleading feature of relative changes.

Figure 2b shows the predicted income and Gini semi-elasticities of poverty. The picture is very different and in many ways more intuitive. If the shortfall is too large – the mass of the income distribution is too far to the left of the poverty line – then both the income and the Gini semi-elasticities approach zero. However, if the country is very affluent – the mass of the income distribution is far to the right of the poverty line – then both semi-elasticities also approach zero. In between those two extremes, improvements in the income distribution can make a very large difference in terms of percent of the population lifted out of poverty, both, directly through redistribution and indirectly through growth. At $\bar{y}_t/z = 1$, for example, a Gini of 0.25 implies that one percent income growth leads to a 0.584 percentage point reduction in the poverty headcount and a Gini of 0.55 implies that one percent income growth leads to a 0.378 percentage point reduction in the poverty headcount. Especially at very low average income levels the initial income distribution is decisive. It practically determines whether there is any substantial potential for poverty alleviation through income growth at all (in terms of percent of the population that is poor). Moreover, as the Gini semi-elasticity also depends on the level of inequality, further improvements in the income distribution will have a larger effect on poverty reduction at lower levels of inequality. Contrary to Figure 2a, this suggests that poverty reduction strategies should focus both on income growth *and* equalization, especially in least developed countries and high inequality countries where the total returns to redistribution are large. Again, for policy questions, these relationships are much more pertinent than relative changes in the poverty headcount.

Could the decomposition be improved by allowing for other “more ultimate” determinants of poverty? If the assumption of log normality is justified, mean income and the Gini fully describe the distribution of incomes and expenditures, and there is logically no scope for other variables to enter. Yet this assumption is restrictive and we deliberately do not rely on log normality. In fact, we expect it to be violated at least for some cases (see, e.g., the host of alternative distributions analyzed by Bresson, 2009). More realistic distributions usually have more than one shape parameter in order to better capture skewness, long tails or the existence of multiple modes. “Ultimate factors” could thus be proxies for systematic deviations from equiproportional shifts in the distribution

Figure 2 – Predicted income and Gini elasticities and semi-elasticities of poverty, \$2 a day

of incomes and expenditures. Weak institutions, for example, may explain the fact that the rich receive more of the gains. Table B-6 in Appendix B extends the heteroskedastic fractional probit models with data on institutions, human capital, access to credit and trade openness. The APEs of income and inequality are not affected by the inclusion of additional covariates and the APEs of other determinants are virtually zero. Thus we conclude that with only two variables, several dummies and correlated random effects, these specifications are essentially saturated. The fractional response approach leaves little room for misspecification of the decomposition.

While the literature on poverty reduction has produced mixed results so far, it is largely consistent with this view. Prominent examples are two studies by Dollar and Kraay (2002, 2004), who find that trade, inflation and other factors influence the incomes of the poorest quintile, while several other variables do not. However, they emphasize

that these effects run predominantly through growth of GDP per capita. The interesting link is between some factor X and income or inequality, not between X and a measure of poverty. Usually such a “first stage” relationship has a dedicated literature that explicitly attempts to resolve causality issues and provides an appropriate theoretical background. Thus if we are interested in the effects of, say, institutions on poverty it is not only sufficient but, in our opinion, much more relevant to investigate the effects of institutions on income (e.g. see Acemoglu et al., 2001; Acemoglu and Johnson, 2005) and inequality (e.g. see Engerman and Sokoloff, 1997; Easterly, 2007). Separating these two estimation steps is important, as the impacts of income and distributional changes themselves depend on the initial levels of income and inequality.

4.2 Projecting poverty

Parts of the previous literature highlight that estimates of income and inequality elasticities or semi-elasticities are particularly useful for poverty simulations (e.g. Klasen and Misselhorn, 2008) and hence model fit is very important. Fractional response models provide a new, powerful and simple method of predicting poverty.

To illustrate the appeal and accuracy of this approach, we compare the predictions of our model for 2010 to the official World Bank data and then extrapolate poverty well into the medium-term future until 2030. Clearly, this is a hypothetical exercise and is not intended to replace any official estimates by the World Bank or national authorities. Rather it allows us to make somewhat more sophisticated predictions than back-of-the-envelope trend extrapolations and can provide a useful benchmark for setting global poverty reduction goals. Using fractional response models for this purpose has the added advantage that we can predict poverty responses to any combination of shifts in mean income and inequality. Further, these models have the attractive feature that the implied changes in the elasticities of poverty at different income and inequality levels are automatically taken into account.

In fact, even the official World Bank regional poverty figures involve a considerable amount of interpolation and extrapolation since most household surveys are not undertaken annually (for details see Chen and Ravallion, 2004). The basic steps are as follows. The World Bank first calculates poverty in the given survey year by fitting Lorenz curves to either the unit-level or grouped-level data. Then, average real household income is lined up to a reference year by interpolating between surveys or extrapolating with the growth rate of personal consumption expenditures per capita (PCE_{it}). Afterwards, the poverty headcount is recalculated using the new income level and the same Lorenz curve as before. If two surveys are available, one before and one after the reference year, the poverty headcount is a weighted average of the two estimates for the reference year.

Our method is similar in spirit. We proceed in four steps. First, we extrapolate the last available survey income to 2010 using actually observed country growth rates in PCE from the WDI, or PWT if the former is missing. Inequality is kept constant at the latest observed value. Second, we project mean income into the future using each country’s average growth rate of PCE over the last 15 years. We assume that growth is distribution neutral, which is in line with the absence of any significant correlation between changes in inequality and income growth (see Figure B-4 in the Appendix). Third, we predict the poverty headcount in 5-year intervals from 2010 to 2030 using our preferred specification without adjusting for measurement error in income (Column 2 in Table 2).²⁴ We typically

²⁴Note that although our preferred specification is only estimated on the sub-sample where $T_i \geq 2$,

do not need estimates of the measurement errors in income or inequality for forecasting purposes, but we implicitly assume that their contribution remains stable over time. Finally, the world total and the regional aggregates are estimated as population-weighted averages of our country level estimates using population data from the World Population Prospects. We do not provide standard errors for the point estimates since these are subject to fundamental uncertainty in the assumed PCE growth rates.

Table 5 shows the results and Table B-5 in Appendix B provides regional PCE growth rates highlighting the assumptions behind these forecasts. The comparison of the official World Bank poverty figures and our estimates in 2010 illustrates that our approach produces meaningful results. For three regions, our estimates are within a percentage point of the official figures, for another two regions, they are within less than 2.5 percentage points, and only for East Asia and Pacific, we estimate a much lower poverty level in 2010. Our results closely match the World Bank's results for the world total. Using the World Bank population data, our estimates imply 2,383.43 million people under the \$2 line worldwide in 2010 versus 2,395.21 million as reported by the World Bank.

Table 5 – Poverty projections ($\hat{H}_{it} \times 100$), preferred specification, \$2 a day

	2010 Official	2010 Estimate	2015 Estimate	2020 Estimate	2025 Estimate	2030 Estimate
East Asia & Pacific	29.73	26.70	16.80	10.53	6.92	4.86
Europe & Central Asia	2.35	2.98	1.88	1.17	0.73	0.46
Latin America & the Caribbean	10.37	10.62	8.96	7.59	6.46	5.52
Middle East & North Africa	12.04	14.57	11.36	8.86	6.92	5.41
South Asia	66.71	68.36	57.50	46.09	35.29	26.06
Sub-Saharan Africa	69.87	69.02	64.74	60.73	57.20	54.02
World Total	40.67	40.47	33.64	28.01	23.55	20.09

Notes: Regional aggregates are based on the World Bank classification of low and middle income countries in 1990. High income countries in 1990 are assumed to have no poor. The projections are for 123 countries. West Bank and Gaza is excluded as we lack both PCE and population data.

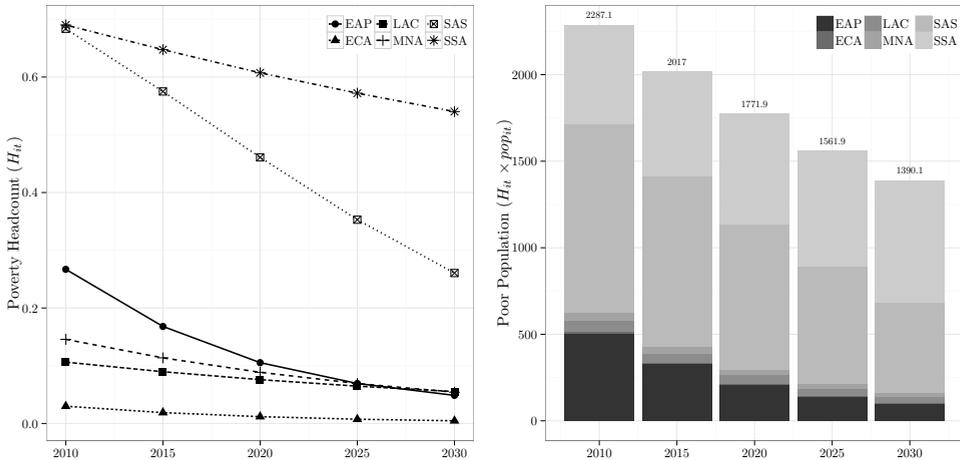
The anticipated regional trends from 2010 to 2030 are highlighted in Figure 3. The left panel shows the regional poverty rates and the right panel plots the regional distribution of the poor population. Given past growth trajectories, poverty in Sub-Saharan Africa and South Asia remains the fundamental development challenge of the twenty-first century. Estimated poverty in Sub-Saharan Africa is very high in 2010 (69.02%) and projected to remain high through 2030 (54.02%) on the entire subcontinent in spite of sustained income growth (about 2.3% p.a.). In South Asia, too, poverty is equally high in 2010 (68.36%) but projected to fall by more than half (to 26.06% in 2030). By 2030, about half of the world's poor will live in Sub-Saharan Africa, followed closely by South Asia.

Poverty in the East Asia and Pacific region, on the contrary, largely takes care of itself if incomes and consumption expenditures keep growing at the impressive rates of the last 15 years. We project poverty in East Asia (4.86%) to be below poverty levels in Latin America (5.52%) by 2030, and second only to Eastern Europe and Central Asia where absolute poverty virtually disappears (down to 0.46%). Most of the progress in East Asia

we can use the model estimates to predict poverty for the entire sample ($T_i \geq 1$). We only lack estimates of the panel size effects for $T_i = 1$, so we assign these observations to the adjacent group ($T_i = 2$).

is due to rapid income and expenditure growth in China. However, this prediction may not hold if a middle-income slow-down occurs in China as some observers suggest (see, e.g., Eichengreen, Park, and Shin, 2013).

Figure 3 – Predicted regional poverty, \$2 a day, 2010 to 2030



Notes: Regional aggregates are based on the World Bank classification of low and middle income countries in 1990. High income countries in 1990 are assumed to have no poor. The projections are for 123 countries. West Bank and Gaza is excluded as we lack both PCE and population data. Population data is the medium fertility variant from the World Population Prospects.

Progress in Latin America and the Caribbean, and the Middle East and North Africa is noticeably slower in spite of the assumption of robust yet moderate income growth (about 2.2% and 2.9% p.a., respectively) and comparatively large income elasticities. This suggests that the countries in these regions should reinforce their poverty alleviation efforts. However, for the Middle East and North Africa these numbers could be too optimistic given the recent social upheavals and volatile economic growth that ensued.

Part of this pattern follows directly from the process of “bunching up above \$1.25 a day and just below \$2 a day” occurring in East Asia and, to a lesser extent, in South Asia over the last two decades (Chen and Ravallion, 2010). These two regions have a relatively large population near the poverty line and hence most of the advances are projected to occur there. Latin America and the Caribbean, as well as the Middle East and North Africa, are richer and require stronger income growth to continuously reduce poverty. Sub-Saharan Africa, on the other hand, has a considerable proportion of the population far below the \$2 a day line in 2010 (with 48.47% poor at \$1.25 a day). It is facing a lower income elasticity and thus *requires exceptionally strong income growth* to make significant strides against poverty. As highlighted in the previous section, this heterogeneity in the income elasticity is mainly due to income differences.

We repeat the same exercise for the \$1.25 a day poverty line. Table B-8 and Figure B-5 in Appendix B show the results based on the estimates presented in Table B-7. The performance of our method is similar and, when compared to the World Bank approach, is just slightly less accurate for the 2010 baseline. The broad patterns are also similar but start from much lower poverty levels. It is worth noting that the gap between Sub-Saharan

Africa and South Asia is even wider for extreme poverty. All regions are predicted to have a poverty headcount below 7% in 2030, except Sub-Saharan Africa where we project poverty levels to remain at about 35%. In 2030, it is likely that the great majority of the world's extremely poor population will live on the Sub-Saharan subcontinent.

What do these results imply for the post-2015 development agenda? We suggest that a new goal to *at least* halve the \$2 a day poverty level within 20 years should be the bare minimum if we want to ensure steady progress. It could be combined with a more ambitious goal for extreme poverty (\$1.25 a day) and significant resources targeted at Sub-Saharan Africa and South Asia. Shifting the policy focus to a higher poverty line makes a lot of sense. The \$1.25 a day poverty line will become nearly irrelevant for most parts of the developing world. In fact, as long as incomes continue to grow, any absolute poverty measure will become less relevant over time when it is set too low and \$2 a day can hardly be described as generous. China, for example, recently updated its national poverty line from less than \$1 a day to about \$1.80 a day. Raising the headline poverty threshold ensures the goal remains relevant as time passes.

Even the lower bound of this poverty reduction goal will not be achieved automatically. A lasting slump in the developed world coupled with the possibility of China entering a middle-income trap could make it a challenge to preserve historical income growth rates throughout the medium-term future. In addition, greater equality can coincide with faster growth and anti-poverty policies will be more successful if they are accompanied by an improving income distribution. In any case, we now have a baseline to calculate goals against and to assess counterfactual assumptions. This can inform discussions on the post-2015 development agenda and help to identify reasonable benchmarks.

5 Concluding remarks

In this chapter we derive a new approach to decomposing the poverty headcount and show that this fractional response approach outperforms earlier linear approximations. Our main point is that the well-established non-linearity of the income and inequality elasticities of poverty arises primarily from the bounded nature of the poverty headcount ratio. Once this inherent non-linearity is taken into account, we can derive an empirical approximation of the poverty decomposition that implies income and inequality (semi-) elasticities with desirable properties.

We use this new approach to estimate income and inequality (semi-)elasticities of poverty based on a large new data set. Fractional response models fit the data extremely well. We provide evidence that the average income elasticity is around two and the average inequality elasticity is about one and a half. However, since these two averages are not very informative, we show that differences in income and inequality levels create strong regional heterogeneity in the estimated elasticities and semi-elasticities. Studies based on linear approximations do not fully capture this heterogeneity. Compared to earlier results, our approach provides estimates that are often substantially different, very stable and considerably more accurate. This holds for a wide range of different combinations of income and inequality. While our approach restricts the nature of the unobserved heterogeneity (measurement differences), it requires no distributional assumptions other than a correctly specified conditional mean. In addition, we show that classical measurement error in income attenuates the elasticity estimates and outweighs systematic survey bias pointing in the opposite direction.

Functional form matters a lot when estimating poverty decompositions. Elasticities and semi-elasticities of poverty estimated with fractional response models have properties closely resembling those of their theoretical counterparts derived under the assumption of log normality. Moreover, we emphasize that semi-elasticities rather than elasticities are the policy relevant metric. This non-linearity also has direct implications for a reduced-form literature interested in the poverty effects of more ultimate determinants or more policy-oriented variables. The focus should be first on how a particular variable of interest leads to changes in mean incomes and changes in distribution and, only in a second step, on how these changes bring about different poverty outcomes. Only in this manner, the non-linearity of the poverty-income-inequality relationship is properly taken into account.

To further illustrate the potential of the fractional response approach, we provide poverty projections from 2010 until 2030 based on the simple assumption that average historical consumption growth continues into the medium-term future. We show that the regional landscape of poverty is likely to change dramatically over the next two decades. Two findings stand out in particular. First, poverty in Sub-Saharan Africa and South Asia will remain the overarching challenge in the twenty-first century. Second, in all other regions poverty is projected to fall rapidly, so that there is a strong rationale for setting the post-2015 development goals on the basis of the \$2 a day poverty line.

It is tempting to interpret our findings as evidence of the primacy of growth. Yet, we are by no means arguing that income growth is all that matters for poverty reduction. It is important to emphasize that the causal effect of any particular policy on aggregate household income and the relative distribution of income cannot be discerned from a decomposition exercise such as this. What it does is help to identify how a *given* change in average income or in distribution translates into poverty outcomes, and not how that change is brought about. Hence, the importance of institutions, trade and a host of other factors for poverty alleviation remains undiminished. There is a potentially large ‘double dividend’ to be reaped if growth can be achieved in combination with simultaneous reductions in inequality.

Other important questions remain open. More work is needed on identifying viable paths of poverty alleviation that actually combine redistribution with growth. Such analyses require a more sophisticated political economy of redistribution and poverty than currently available. In addition, the issue of statistical discrepancies between expenditure surveys, national accounts and consumption proxies, and what these discrepancies imply for the confidence we place into poverty estimates, is just beginning to be explored. In our view, these are exciting areas for future research.

Bibliography

- Acemoglu, D. and S. Johnson (2005). Unbundling institutions. *Journal of Political Economy* 113(5), 949–95.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5), 1369–1401.
- Adams, R. H. (2004). Economic growth, inequality and poverty: Estimating the growth elasticity of poverty. *World Development* 32(12), 1989–2014.
- Aitchison, J. and J. A. C. Brown (1957). *The Lognormal Distribution: With Special References to Its Uses in Economics*. Cambridge University Press.
- Altonji, J. G. and R. L. Matzkin (2005). Cross section and panel data estimators for nonseparable models with endogenous regressors. *Econometrica* 73(4), 1053–1102.
- Barro, R. J. and J. W. Lee (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics* 104, 184–198.
- Battistin, E., R. Blundell, and A. Lewbel (2009). Why is consumption more log normal than income? Gibrat’s law revisited. *Journal of Political Economy* 117(6), 1140–1154.
- Beck, T., A. Demirgüç-Kunt, and R. Levine (2010). Financial institutions and markets across countries and over time: The updated financial development and structure database. *World Bank Economic Review* 24(1), 77–92.
- Besley, T. and R. Burgess (2003). Halving global poverty. *Journal of Economic Perspectives* 17(3), 3–22.
- Bluhm, R. (2014). fhctprob: A fast qmlc stata routine for fractional probit models with multiplicative heteroskedasticity. Forthcoming as UNU-MERIT working paper.
- Blundell, R. W. and J. L. Powell (2004). Endogeneity in semiparametric binary response models. *Review of Economic Studies* 71(3), 655–679.
- Bourguignon, F. (2003). The growth elasticity of poverty reduction: Explaining heterogeneity across countries and time periods. In T. S. Eicher and S. J. Turnovsky (Eds.), *Inequality and Growth: Theory and Policy Implications*, pp. 3–26. Cambridge, MA: MIT Press.
- Bresson, F. (2009). On the estimation of growth and inequality elasticities of poverty with grouped data. *Review of Income and Wealth* 55(2), 266–302.
- Chamberlain, G. (1984). Panel data. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume 2 of *Handbook of Econometrics*, pp. 1247–1318. Elsevier.
- Chambers, D. and S. Dhongde (2011). A non-parametric measure of poverty elasticity. *Review of Income and Wealth* 57(4), 683–703.
- Chen, S. and M. Ravallion (2004). How have the world’s poorest fared since the early 1980s? *The World Bank Research Observer* 19(2), 141–169.
- Chen, S. and M. Ravallion (2010). The developing world is poorer than we thought, but no less successful in the fight against poverty. *Quarterly Journal of Economics* 125(4), 1577–1625.
- Cook, D. O., R. Kieschnick, and B. D. McCullough (2008). Regression analysis of proportions in finance with self selection. *Journal of Empirical Finance* 15(5), 860–867.
- Datt, G. and M. Ravallion (1992). Growth and redistribution components of changes in poverty measures: A decomposition with applications to Brazil and India in the 1980s. *Journal of Development Economics* 38(2), 275–295.
- Deaton, A. and A. Heston (2010). Understanding PPPs and PPP-based national accounts. *American Economic Journal: Macroeconomics* 2(4), 1–35.
- Dollar, D. and A. Kraay (2002). Growth is good for the poor. *Journal of Economic Growth* 7(3), 195–225.
- Dollar, D. and A. Kraay (2004). Trade, growth, and poverty. *Economic Journal* 114(493), F22–F49.
- Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics* 84(2), 755–776.
- Eichengreen, B., D. Park, and K. Shin (2013, January). Growth slowdowns redux: New evidence on the middle-income trap. Working Paper 18673, National Bureau of Economic Research.
- Engerman, S. L. and K. L. Sokoloff (1997). Factor endowments, institutions, and differential growth paths among new world economies. In S. Haber (Ed.), *How Latin America Fell Behind*, pp. 260–296. Stanford, CA: Stanford University Press.
- Gourieroux, C., A. Monfort, and A. Trognon (1984). Pseudo maximum likelihood methods: Theory. *Econometrica* 52(3), 681–700.

- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica* 46(6), 1251–1271.
- Kakwani, N. (1993). Poverty and economic growth with application to Côte d’Ivoire. *Review of Income and Wealth* 39(2), 121–139.
- Kalwij, A. and A. Verschoor (2007). Not by growth alone: The role of the distribution of income in regional diversity in poverty reduction. *European Economic Review* 51(4), 805–829.
- Klasen, S. and M. Misselhorn (2008). Determinants of the growth semi-elasticity of poverty reduction. Working paper, Ibero America Institute for Economic Research.
- Kraay, A. (2006). When is growth pro-poor? Evidence from a panel of countries. *Journal of Development Economics* 80(1), 198–227.
- Loudermilk, M. S. (2007). Estimation of fractional dependent variables in dynamic panel data models with an application to firm dividend policy. *Journal of Business & Economic Statistics* 25(4), 462–472.
- MacCullagh, P. and J. A. Nelder (1989). *Generalized linear models* (2nd ed.). Chapman & Hall/ CRC Press.
- Milanovic, B. (2002). True world income distribution, 1988 and 1993: First calculation based on household surveys alone. *Economic Journal* 112(476), 51–92.
- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica* 46(1), 69–85.
- Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica* 16(1), 1–32.
- Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics* 11(6), 619–632.
- Papke, L. E. and J. M. Wooldridge (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics* 145(1–2), 121–133.
- Ram, R. (2013). Income elasticity of poverty in developing countries: updated estimates from new data. *Applied Economics Letters* 20(6), 554–558.
- Ramalho, E. A., J. J. Ramalho, and J. M. Murteira (2011). Alternative estimating and testing empirical strategies for fractional regression models. *Journal of Economic Surveys* 25(1), 19–68.
- Ravallion, M. (2001). Growth, inequality and poverty: Looking beyond averages. *World Development* 29(11), 1803–1815.
- Ravallion, M. and S. Chen (1997). What can new survey data tell us about recent changes in distribution and poverty? *World Bank Economic Review* 11(2), 357–382.
- Rivers, D. and Q. H. Vuong (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics* 39(3), 347–366.
- Sachs, J. D. and A. Warner (1995). Economic reform and the process of global integration. *Brookings Papers on Economic Activity* 1995(1), 1–118.
- Sala-i-Martin, X. (2006). The world distribution of income: Falling poverty and . . . convergence, period. *Quarterly Journal of Economics* 121(2), 351–397.
- Tiwari, A. and F. Palm (2011). Nonlinear panel data models with expected a posteriori values of correlated random effects. Working Paper 2011/13, CREPP HEC-Management School, University of Liège.
- Wacziarg, R. and K. H. Welch (2008). Trade liberalization and growth: New evidence. *World Bank Economic Review* 22(2), 187–231.
- Wooldridge, J. M. (2010a, May). Correlated random effects models with unbalanced panels. Manuscript.
- Wooldridge, J. M. (2010b). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). The MIT Press.
- Wooldridge, J. M. (2012, June). Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. Manuscript.
- Young, A. (2011, December). The gini coefficient for a mixture of ln-normal populations. Manuscript. The London School of Economics and Political Science, London, UK.

Appendix A: Derivations

Let the poverty line (z) be fixed and assume the poverty headcount is described by a two-parameter distribution, so that $H(\bar{y}_t/z, \sigma_t) = H(\bar{y}_t, \sigma_t) = H_t$. A Taylor linearization of $H(\cdot)$ about (\bar{y}_t, σ_t) gives

$$H(\bar{y}_t + d\bar{y}_t, \sigma_t + d\sigma_t) = H(\bar{y}_t, \sigma_t) + \frac{\partial H_t}{\partial \bar{y}_t} d\bar{y}_t + \frac{\partial H_t}{\partial \sigma_t} d\sigma_t + \xi_t \quad (\text{A-1})$$

where dx denotes a differential change in x , and ξ_t is a second-order remainder. This is easily extended to allow for a vector of Lorenz curve parameters as in Kakwani (1993).

Subtracting $H(\bar{y}_t, \sigma_t)$ from both sides, dropping the remainder by approximation, dividing through by H_t (provided $H_t > 0$), and multiplying the first (second) term by \bar{y}_t/\bar{y}_t (σ_t/σ_t), we arrive at eq. (4) from the main text:

$$\frac{dH_t}{H_t} \approx \left(\frac{\partial H_t}{\partial \bar{y}_t} \frac{\bar{y}_t}{H_t} \right) \frac{d\bar{y}_t}{\bar{y}_t} + \left(\frac{\partial H_t}{\partial \sigma_t} \frac{\sigma_t}{H_t} \right) \frac{d\sigma_t}{\sigma_t} = \varepsilon_t^{H\bar{y}} \frac{d\bar{y}_t}{\bar{y}_t} + \varepsilon_t^{H\sigma} \frac{d\sigma_t}{\sigma_t}. \quad (\text{A-2})$$

If we do not divide by H_t , we get a decomposition of the (non-relative) change of poverty in terms of income and inequality semi-elasticities (as in Klasen and Misselhorn, 2008).

Similar steps starting from $H(\bar{y}_t, G_t)$ lead to a decomposition in terms of mean income and Gini. Using the chain rule for elasticities, an expression for the Gini elasticity is

$$\varepsilon_t^{HG} = \varepsilon_t^{H\sigma} \left(\frac{dG_t}{d\sigma_t} \frac{\sigma_t}{G_t} \right)^{-1} \quad (\text{A-3})$$

enabling us to write

$$\frac{dH_t}{H_t} \approx \varepsilon_t^{H\bar{y}} \frac{d\bar{y}_t}{\bar{y}_t} + \varepsilon_t^{HG} \frac{dG_t}{G_t} = \varepsilon_t^{H\bar{y}} \frac{d\bar{y}_t}{\bar{y}_t} + \varepsilon_t^{H\sigma} \left(\frac{dG_t}{d\sigma_t} \frac{\sigma_t}{G_t} \right)^{-1} \frac{dG_t}{G_t} \quad (\text{A-4})$$

where eqs. (2) and (3) give $\varepsilon_t^{H\bar{y}}$ and $\varepsilon_t^{H\sigma}$ under log normality, but we still need an expression for $dG_t/d\sigma_t$ to get an explicit formula for ε_t^{HG} .

Even though we restricted our attention to one inequality parameter, the results thus far are quite general. Now if we also assume log normality, we arrive at an explicit form for the Gini elasticity. Using $\sigma_t = \sqrt{2}\Phi^{-1}(G_t/2 + 1/2)$, we have

$$\frac{dG_t}{d\sigma_t} = \frac{d[2\Phi(\sigma_t/\sqrt{2}) - 1]}{d\sigma_t} = \sqrt{2}\phi\left(\frac{\sigma_t}{\sqrt{2}}\right). \quad (\text{A-5})$$

Inverting and substituting eq. (A-5) together with eq. (3) from the main text into eq. (A-3) gives the Gini elasticity

$$\varepsilon_t^{HG} = \left(\frac{\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t \right) \left(\frac{\sigma_t}{G_t} \sqrt{2}\phi\left(\frac{\sigma_t}{\sqrt{2}}\right) \right)^{-1} \lambda\left(\frac{-\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t \right) \quad (\text{A-6})$$

where $\sigma_t = \sqrt{2}\Phi^{-1}(G_t/2 + 1/2)$. This result corrects for the missing σ_t/G_t in Kalwij and Verschuur (2007, p. 824). The Gini semi-elasticity (η_t^{HG}) is just eq. (A-6) with $\phi(\cdot)$ replacing $\lambda(\cdot)$. Clearly, both the Gini elasticity and the Gini semi-elasticity are highly non-linear functions, as illustrated in Figure 2.

Appendix B: Additional tables and figures

Figure B-1 – Transformed headcount (\$2 a day) and log mean income, by region

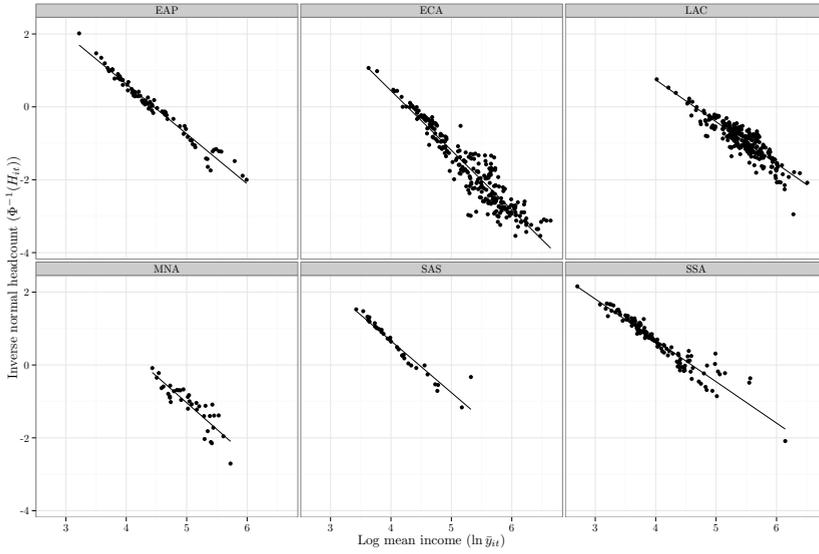


Figure B-2 – Transformed headcount (\$2 a day) and log Gini, by region

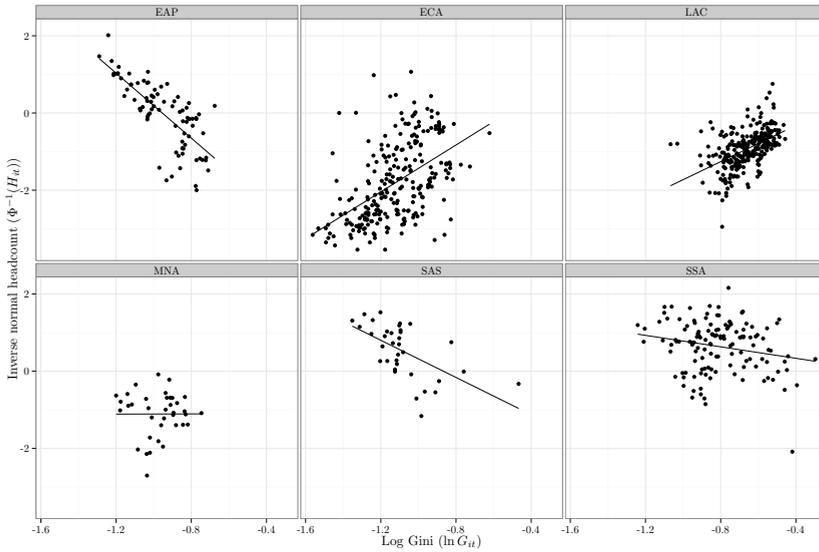


Table B-1 – Summary statistics by region (unweighted)

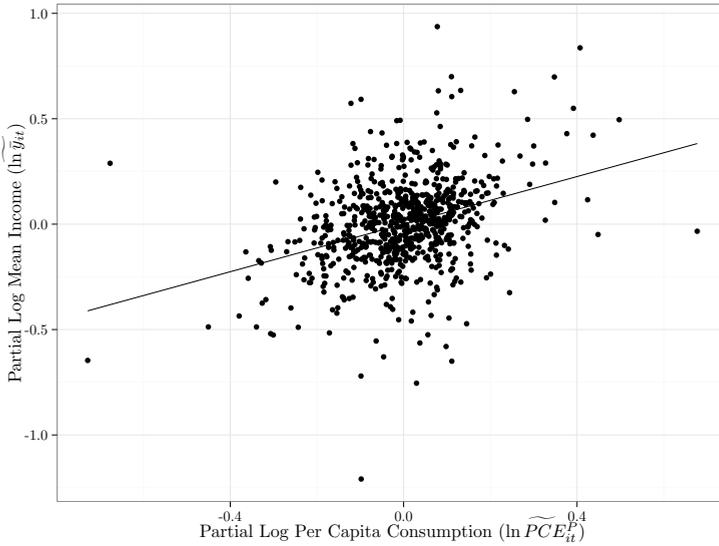
Variable	Mean	Standard Deviation	Min	Max
<i>East Asia and Pacific (N=80)</i>				
H_{it} – Headcount (2\$)	0.502	0.267	0.023	0.978
G_{it} – Gini coefficient	0.392	0.058	0.275	0.509
\bar{y}_{it} – Mean income or expenditure	107.86	78.39	25.02	399.76
<i>Eastern Europe and Central Asia (N=254)</i>				
H_{it} – Headcount (2\$)	0.110	0.169	0.000	0.857
G_{it} – Gini coefficient	0.330	0.056	0.210	0.537
\bar{y}_{it} – Mean income or expenditure	251.99	136.11	37.66	766.78
<i>Latin America and Caribbean (N=274)</i>				
H_{it} – Headcount (2\$)	0.204	0.122	0.002	0.775
G_{it} – Gini coefficient	0.523	0.054	0.344	0.633
\bar{y}_{it} – Mean income or expenditure	246.63	90.55	55.53	671.04
<i>Middle East and North Africa (N=37)</i>				
H_{it} – Headcount (2\$)	0.166	0.111	0.003	0.466
G_{it} – Gini coefficient	0.380	0.042	0.301	0.474
\bar{y}_{it} – Mean income or expenditure	165.26	56.59	84.02	306.33
<i>South Asia (N=35)</i>				
H_{it} – Headcount (2\$)	0.672	0.226	0.122	0.936
G_{it} – Gini coefficient	0.343	0.067	0.259	0.627
\bar{y}_{it} – Mean income or expenditure	67.78	39.20	30.71	204.98
<i>Sub-Saharan Africa (N=129)</i>				
H_{it} – Headcount (2\$)	0.708	0.202	0.018	0.985
G_{it} – Gini coefficient	0.453	0.087	0.289	0.743
\bar{y}_{it} – Mean income or expenditure	67.62	54.04	14.93	465.80

Notes: Mean income or expenditure in \$ per month. 809 observations, 124 countries in total, unbalanced sample from 1981 to 2010.

List B-1 – Included countries (number of surveys)

Albania (5), Algeria (2), Angola (2), Argentina (21), Armenia (11), Azerbaijan (3), Bangladesh (8), Belarus (14), Belize (7), Benin (1), Bhutan (2), Bolivia, Plurinational State of (11), Bosnia and Herzegovina (3), Botswana (2), Brazil (26), Bulgaria (7), Burkina Faso (4), Burundi (3), Cambodia (5), Cameroon (3), Cape Verde (1), Central African Rep. (3), Chad (1), Chile (10), China (16), Colombia (14), Comoros (1), Congo, Dem. Rep. of (1), Congo, Rep. of (1), Costa Rica (23), Cote D'Ivoire (9), Croatia (7), Czech Rep. (2), Djibouti (1), Dominican Rep. (16), Ecuador (12), Egypt (5), El Salvador (14), Estonia (9), Ethiopia (4), Fiji (2), Gabon (1), Gambia (2), Georgia (14), Ghana (5), Guatemala (8), Guinea (4), Guinea-Bissau (2), Guyana (2), Haiti (1), Honduras (20), Hungary (10), India (5), Indonesia (13), Iran, Islamic Rep. of (5), Iraq (1), Jamaica (7), Jordan (7), Kazakhstan (11), Kenya (4), Kyrgyzstan (10), Lao People's Dem. Rep. (4), Latvia (11), Lesotho (4), Liberia (1), Lithuania (9), Macedonia, Rep. of (10), Madagascar (6), Malawi (3), Malaysia (9), Maldives (2), Mali (4), Mauritania (6), Mexico (11), Micronesia, Federated States of (1), Moldova, Rep. of (15), Montenegro (4), Morocco (5), Mozambique (3), Namibia (2), Nepal (4), Nicaragua (4), Niger (4), Nigeria (5), Pakistan (8), Palestinian Territory, Occupied (2), Panama (13), Papua New Guinea (1), Paraguay (14), Peru (16), Philippines (9), Poland (18), Romania (15), Russian Federation (13), Rwanda (3), Saint Lucia (1), Sao Tome and Principe (1), Senegal (4), Serbia (9), Seychelles (1), Sierra Leone (1), Slovakia (7), Slovenia (6), South Africa (5), Sri Lanka (6), Sudan (1), Suriname (1), Swaziland (3), Syrian Arab Rep. (1), Tajikistan (5), Tanzania, United Rep. of (3), Thailand (14), Timor-Leste (2), Togo (1), Trinidad and Tobago (2), Tunisia (6), Turkey (11), Turkmenistan (3), Uganda (7), Ukraine (13), Uruguay (7), Venezuela, Bolivarian Rep. of (13), Vietnam (6), Yemen (2), Zambia (7).

Figure B-3 – Partial regression plot – first stage



Notes: The figure plots two residual series, so that the plotted slope is identical to the slope of $\ln PCE_{it}^P$ in the first stage. On the x-axis:

$$\ln PCE_{it}^P = \ln PCE_{it}^P - \mathbf{x}'_{1it}\hat{\beta}_1 - \sum_{r=1}^T \delta_{T_i,r}\hat{\varphi}_{1r} - \sum_{r=1}^T \delta_{T_i,r}\bar{\mathbf{x}}'_i\hat{\theta}_{1r}.$$

On the y-axis:
 $\ln \bar{y}_{it} = \ln \bar{y}_{it} - \mathbf{x}'_{1it}\hat{\beta}_1 - \sum_{r=1}^T \delta_{T_i,r}\hat{\varphi}_{1r} - \sum_{r=1}^T \delta_{T_i,r}\bar{\mathbf{x}}'_i\hat{\theta}_{1r}.$ In both cases, \mathbf{x}'_{1it} includes only the log of Gini but $\bar{\mathbf{x}}'_i$ contains the time averages of $\ln G_{it}$ and $\ln PCE_{it}^P$. Both regressions also contain survey type and time dummies, as well as their time averages.

Figure B-4 – Inequality changes and income growth, 1981–2010

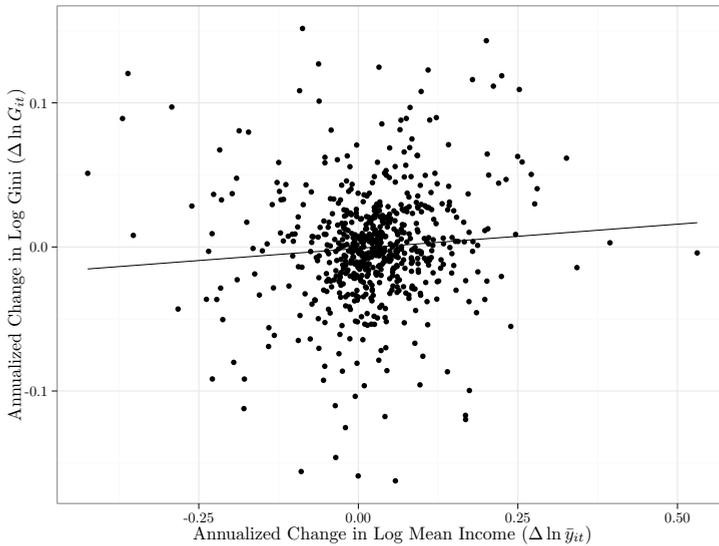


Table B-2 – Linear models – Dependent variable: $\ln H_{it}$, \$2 a day

	OLS			Two-Step GMM		
	(1)	(2)	(3)	(4)	(5)	(6)
	Within R+C '97	Differences R+C '97	Differences Bourg. '03	Differences R+C '97	Differences Bourg. '03	Differences K+V '07
$\Delta \ln \bar{y}_{it}$		-1.895 (0.170)	-0.268 (0.617)	-2.028 (0.271)	2.046 (1.043)	-0.362 (3.216)
$\Delta \ln \bar{y}_{it} \times \ln G_{i,t-1}$			1.108 (0.671)		3.445 (1.192)	2.097 (2.315)
$\Delta \ln \bar{y}_{it} \times \ln(\bar{y}_{i,t-1}/z)$			-0.552 (0.179)		-0.995 (0.258)	-0.517 (0.785)
$\Delta \ln G_{it}$		2.336 (0.311)	-0.527 (1.449)	1.664 (1.008)	1.257 (4.127)	-8.222 (11.185)
$\Delta \ln G_{it} \times \ln G_{i,t-1}$			-1.769 (1.586)		-1.416 (3.929)	-8.164 (8.296)
$\Delta \ln G_{it} \times \ln(\bar{y}_{i,t-1}/z)$			1.261 (0.427)		-0.315 (1.172)	-1.382 (1.996)
$\ln \bar{y}_{it}$	-2.114 (0.204)					
$\ln G_{it}$	3.024 (0.409)					
$\ln G_{i,t-1}$						-0.129 (0.134)
$\ln(\bar{y}_{i,t-1}/z)$						-0.023
$\bar{\varepsilon}^{H\bar{y}}$	-2.114	-1.895	-1.755	-2.028	-1.905	-2.684
$\bar{\varepsilon}^{HG}$	3.024	2.336	2.201	1.664	2.206	-2.345
$N \times \bar{T}$	648	648	648	641	641	641
N	104	104	104	102	102	102
Hansen's J (p-val.)	–	–	–	0.0418	0.579	0.639

Notes: All standard errors are robust to clustering at the country-level. The GMM results are estimated using two-step efficient GMM. Column (4) uses as instruments ΔPCE_{it} , $PCE_{i,t-1}$, $\ln \bar{y}_{i,t-1}$ and $\ln G_{i,t-1}$. Column (5) uses as instruments ΔPCE_{it} , $PCE_{i,t-1}$, $\Delta PCE_{it} \times \ln G_{i,t-1}$, $\Delta PCE_{it} \times \ln(\bar{y}_{i,t-1}/z)$, $\ln \bar{y}_{i,t-1}$, $\ln \bar{y}_{i,t-1} \times \ln G_{i,t-1}$, $\ln \bar{y}_{i,t-1} \times \ln(\bar{y}_{i,t-1}/z)$, $\ln G_{i,t-1}$ and $\ln G_{i,t-1} \times \ln G_{i,t-1}$. Column (6) uses the same instruments as column (5) but $\ln \bar{y}_{i,t-1}$ and $\ln G_{i,t-1}$ instrument for themselves. All models include a constant (not shown) and column (1) includes a time trend (not shown). Columns (2) and (4) are similar to Ravallion and Chen (1997) but we update their approach by also including the Gini as in Adams (2004), columns (3) and (5) are similar to the “improved standard model 2” in Bourguignon (2003), and column (6) is in the spirit of the preferred specification in Kalwij and Verschoor (2007). The latter also use the annualized log difference of the population ($\Delta \ln pop_{it}$) as an instrument and rely on real GDP per capita instead of real per capita consumption. A first-stage F -test shows that $\Delta \ln pop_{it}$ is an extremely weak instrument. Kalwij and Verschoor (2007) also use interactions of lagged inequality and lagged income with regional dummies as instruments. However, first stage diagnostics suggest a weak IV problem (the F -stat with regional dummy interactions is always lower than without) and thus we opt for a simpler instrument set. Further, in column (5) and equation (5) we do not include the lagged levels of income and inequality. Column (6) includes them for comparison with Kalwij and Verschoor (2007).

Table B-3 – Predicted regional income semi-elasticities, preferred specification, \$2 a day

	<i>Time period</i>				
	1981–1989	1990–1994	1995–1999	2000–2004	2005–2010
East Asia and Pacific	-0.568 (0.034)	-0.573 (0.036)	-0.585 (0.046)	-0.583 (0.042)	-0.552 (0.051)
Eastern Europe and Central Asia	-0.031 (0.008)	-0.214 (0.015)	-0.260 (0.020)	-0.225 (0.015)	-0.134 (0.010)
Latin America and Caribbean	-0.374 (0.028)	-0.348 (0.025)	-0.334 (0.024)	-0.355 (0.026)	-0.194 (0.013)
Middle East and North Africa	-0.405 (0.034)	-0.422 (0.037)	-0.447 (0.042)	-0.463 (0.043)	-0.313 (0.024)
South Asia	-0.418 (0.023)	-0.458 (0.019)	-0.526 (0.022)	-0.572 (0.036)	-0.585 (0.044)
Sub-Saharan Africa	-0.532 (0.024)	-0.354 (0.020)	-0.353 (0.020)	-0.440 (0.015)	-0.459 (0.015)

Notes: Standard errors obtained via a panel bootstrap using 999 replications. The predicted semi-elasticities are based on estimated APEs at each region/time-period mean of $\ln \bar{y}_{it}$ and $\ln G_{it}$.

Table B-4 – Predicted regional Gini semi-elasticities, preferred specification, \$2 a day

	<i>Time period</i>				
	1981–1989	1990–1994	1995–1999	2000–2004	2005–2010
East Asia and Pacific	0.419 (0.053)	0.423 (0.053)	0.432 (0.055)	0.431 (0.054)	0.408 (0.053)
Eastern Europe and Central Asia	0.023 (0.007)	0.158 (0.015)	0.192 (0.019)	0.166 (0.017)	0.099 (0.012)
Latin America and Caribbean	0.276 (0.046)	0.257 (0.043)	0.247 (0.043)	0.262 (0.045)	0.143 (0.029)
Middle East and North Africa	0.299 (0.041)	0.311 (0.040)	0.330 (0.041)	0.342 (0.044)	0.231 (0.025)
South Asia	0.309 (0.056)	0.338 (0.055)	0.389 (0.052)	0.423 (0.054)	0.432 (0.055)
Sub-Saharan Africa	0.393 (0.050)	0.261 (0.037)	0.261 (0.040)	0.325 (0.042)	0.339 (0.043)

Notes: Standard errors obtained via a panel bootstrap using 999 replications. The predicted semi-elasticities are based on estimated APEs at each region/time-period mean of $\ln \bar{y}_{it}$ and $\ln G_{it}$.

Table B-5 – Growth in personal consumption expenditures per capita (in %)

Region	Last 5 years	Last 10 years	Last 15 years	Last 20 years
East Asia and Pacific	6.843 (1.041)	5.962 (0.782)	5.647 (0.837)	5.967 (0.757)
Europe and Central Asia	4.532 (1.039)	6.033 (1.063)	4.793 (0.496)	2.856 (0.423)
Latin America and the Caribbean	3.364 (0.746)	2.399 (0.303)	2.222 (0.172)	2.267 (0.151)
Middle East and North Africa	2.705 (0.634)	3.778 (0.560)	2.911 (0.370)	2.499 (0.349)
South Asia	5.563 (0.865)	4.684 (0.556)	4.123 (0.537)	3.636 (0.438)
Sub-Saharan Africa	1.710 (1.599)	2.765 (0.682)	2.338 (0.577)	1.742 (0.414)
$N \times \bar{T}$	615	1222	1795	2332
N	123	123	123	123

Notes: Cluster robust standard errors in parentheses. Regional means are population weighted. Only the third column is relevant for the projections.

Table B-6 – Fractional probit models (QMLE) – Dependent variable: H_{it} , \$2 a day

	(1)		(2)		(3)		(4)	
	Institutions		Human Capital		Credit		Trade	
	H_{it}	APEs	H_{it}	APEs	H_{it}	APEs	H_{it}	APEs
$\ln \bar{y}_{it}$	-0.888 (0.050)	-0.285 (0.012)	-0.878 (0.060)	-0.284 (0.011)	-0.950 (0.036)	-0.289 (0.009)	-0.708 (0.032)	-0.302 (0.012)
$\ln G_{it}$	0.779 (0.107)	0.250 (0.028)	0.805 (0.104)	0.261 (0.027)	0.765 (0.102)	0.233 (0.027)	0.581 (0.097)	0.248 (0.033)
Executive Constraints	0.005 (0.005)	0.001 (0.001)						
Year of Schooling			-0.002 (0.017)	-0.001 (0.006)				
Private Credit / GDP					-0.007 (0.040)	-0.002 (0.012)		
Trade Openness							0.005 (0.017)	0.002 (0.007)
Scale factor	0.321		0.324		0.304		0.426	
$N \times \bar{T}$	678		705		697		385	
N	85		87		93		81	
AIC	894.8		914.1		887.6		552.5	
$\ln \mathcal{L}$	-276.4		-286.1		-282.8		-163.2	
\sqrt{MSE}	0.0203		0.0211		0.0201		0.0233	

Notes: The estimation samples are smaller due to less data coverage and all observations with $T_i = 1$ are not used in estimation. All models include time averages (CRE), time dummies, survey dummies, panel size dummies and interactions between the panel size dummies and the time averages (CRE). The time averages are recomputed for each sample size. The coefficients of the time average of the survey dummies and time effects are constrained to be equal across the panel sizes. The variance equation depends on the panel size. The standard errors of the coefficients are robust to clustering at the country level and the standard errors of the APEs are computed via the delta method. Data on *Executive Constraints* is from the Polity IV database. Human capital is measured as *Total Years of Schooling* from Barro and Lee (2013). We linearly interpolate the quinquennial data to an annual series. Financial development measured as *Private Credit / GDP* is from Beck, Demirgüç-Kunt, and Levine (2010). De jure *Trade Openness* is the binary measure developed by Sachs and Warner (1995) and extended by Wacziarg and Welch (2008).

Table B-7 – Fractional probit models (QMLE) – Dependent variable: H_{it} , \$1.25 a day

	(1)		(2)		(3)	
	Regular		Unbalanced		Unbalanced + Two-Step	
	H_{it}	APEs	H_{it}	APEs	H_{it}	APEs
$\ln \bar{y}_{it}$	-1.212 (0.056)	-0.216 (0.010)	-0.668 (0.038)	-0.218 (0.008)	-0.800 (0.180)	-0.263 (0.034)
$\ln G_{it}$	1.238 (0.121)	0.221 (0.022)	0.726 (0.074)	0.237 (0.020)	0.714 (0.180)	0.235 (0.032)
$\hat{\nu}_{it}$					0.104 (0.104)	
CRE (Corr. Rand. Effects)	Yes		Yes		Yes	
Survey type dummies	Yes		Yes		Yes	
Time dummies	Yes		Yes		Yes	
Panel size dummies	No		Yes		Yes	
Panel size dummies \times CRE	No		Yes		Yes	
Variance equation	No		Yes		Yes	
Scale factor	0.179		0.326		0.329	
$N \times \bar{T}$	768		768		754	
N	103		103		102	
pseudo R^2	0.988		0.995		0.995	
$\ln \mathcal{L}$	-172.4		-244.7		-243.7	
\sqrt{MSE}	0.0339		0.0214		0.0220	

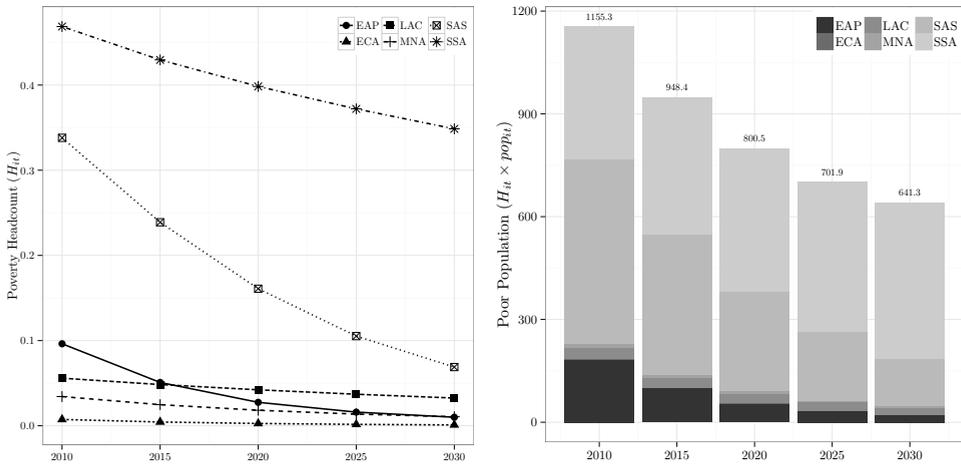
Notes: The \$1.25 a day sample is smaller as for 20 observation we only have data at the \$2 a day line. 21 observations with $T_i = 1$ are not used in estimation. In columns (1) and (2), the standard errors of the coefficients are robust to clustering at the country level and the standard errors of the APEs are computed via the delta method. We include the time averages of the survey type and time dummies in columns (2) and (3), but constrain their coefficients to be equal across the panel sizes. The standard errors of the coefficients and the APEs in model (3) account for the first stage estimation step with a panel bootstrap using 999 bootstrap replications. The linear projection in the first stage uses $\ln PCE_{it}^P$ as an instrument for $\ln \bar{y}_{it}$. The first-stage cluster-robust F-statistic in column (3) is 24.40. Column (3) also excludes West Bank and Gaza entirely (2 observations) and 12 observations from ECA countries pre-1990 for lack of PCE data.

Table B-8 – Poverty projections ($\hat{H}_{it} \times 100$), preferred specification, \$1.25 a day

	2010 Official	2010 Estimate	2015 Estimate	2020 Estimate	2025 Estimate	2030 Estimate
East Asia & Pacific	12.48	9.63	5.09	2.76	1.60	1.00
Europe & Central Asia	0.66	0.74	0.44	0.27	0.16	0.10
Latin America & the Caribbean	5.53	5.59	4.84	4.22	3.70	3.26
Middle East & North Africa	2.41	3.43	2.48	1.82	1.37	1.06
South Asia	31.03	33.81	23.89	16.09	10.53	6.89
Sub-Saharan Africa	48.47	46.87	42.95	39.84	37.20	34.86
World Total	20.63	20.44	15.82	12.66	10.58	9.27

Notes: Regional aggregates are based on the World Bank classification of low and middle income countries in 1990. High income countries in 1990 are assumed to have no poor. The projections are for 123 countries. West Bank and Gaza is excluded as we lack both PCE and population data.

Figure B-5 – Predicted regional poverty, \$1.25 a day, 2010 to 2030



Chapter 6

Poor trends

*The pace of poverty reduction after the Millennium Development Agenda**

Abstract

This chapter reviews the origins of the dollar-a-day poverty line, discusses historical poverty and inequality trends, and forecasts poverty rates until 2030 using a new fractional response approach. Three findings stand out. First, global poverty reduction since 1981 has been rapid but regional trends are heterogeneous. Second, the pace of poverty reduction at \$1.25 a day will slow down. Our optimistic scenarios suggest a poverty rate of 8-9% in 2030, far short of the World Bank's new 3% target. Third, rapid progress can be maintained at \$2 a day, with an additional one billion people crossing that line by 2030.

Keywords: poverty, inequality, consumption growth

JEL Classification: I32, O10, O15

*This chapter is based on Bluhm, R., de Crombrughe, D., and A. Szirmai. 2013. Poor trends - The pace of poverty reduction after the Millennium Development Agenda, MERIT Working Papers 006, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

1 Introduction

“We are at an auspicious moment in history, when the successes of past decades and an increasingly favorable economic outlook combine to give developing countries a chance, for the first time ever, to end extreme poverty within a generation” (Jim Yong Kim, World Bank President, speaking at Georgetown University, April 3, 2013)

Only 13 years after the Millennium Summit in September 2000 at which world leaders agreed on halving the 1990 global poverty rate at \$1.25 a day by 2015, the end of extreme poverty seems to be in sight. Recent estimates suggest that the first Millennium Development Goal (MDG) was already reached in 2010 and about 700 million people were lifted out of poverty. In 2013, the World Bank declared a new organizational goal of ending extreme poverty by 2030, that is, reducing the \$1.25 a day poverty rate to 3% by 2030. The last two decades clearly ushered in unprecedented success, but is 2030 really likely to mark the end of extreme poverty? Our main contribution is to demonstrate that this is unlikely.

In this chapter, we review the origins of the ‘dollar-a-day’ poverty line, discuss progress over the last three decades, and forecast \$1.25 and \$2 a day poverty rates until 2030. It is well-known that regional trends in poverty alleviation are very heterogeneous. In spite of rising inequality, rapid growth in China was the driving force behind global progress over the last two decades and accounts for more than three quarters of the reduction in the number of people living below \$1.25 a day. However, most of the poverty reduction potential coming from China is now exhausted. Poverty reduction in the developing world outside China has been considerably slower, although economic growth has accelerated significantly since 2000. In 2010, three-fourths of the extremely poor lived in Sub-Saharan Africa and South Asia, as opposed to approximately 40% in 1981. This changing regional composition of world poverty has important ramifications for future trends in poverty reduction. The number of inhabitants of historically fast growing countries make up less and less of the global poor.

Building on a new method for estimating poverty elasticities and predicting poverty headcount ratios developed in Chapter 5, we show that the pace of poverty reduction at \$1.25 a day is likely to slow down significantly after 2015. Extreme poverty barely falls below 8% in the most optimistic scenario. Ravallion (2013) first suggested the 3% target relying on the assumption that consumption in developing countries would continue to grow at the average post-2000 trend, or 4.5% per year. We find this ‘equal-growth’ assumption too optimistic. Poverty tends to be higher in countries with rapid population growth and lower than average consumption growth. None of our scenarios predicts a poverty rate near 3% once country-specific trends from 2000 to 2010 are used. However, the \$2 a day poverty rate may fall below 20% in 2030, while a slowdown happens only late during the forecast period or not at all. A distinct advantage of our approach is that it is computationally inexpensive. Hence, it can easily be used for benchmarking progress as new data become available.

The chapter is structured as follows. Section 2 discusses some of the controversies surrounding the setting and updating of international poverty lines. Section 3 is a data-driven review of global poverty and inequality trends with a particular focus on China, India, Brazil and Nigeria. Section 4 presents projections of global and regional poverty rates until 2030 at the \$1.25 and \$2 a day poverty lines using different growth and inequality scenarios. Section 5 concludes and offers some policy recommendations.

2 Drawing the line: international poverty lines

The dollar-a-day poverty line was first defined in a background paper to the 1990 World Development Report (in 1985 PPPs), then updated to \$1.08 (in 1993 PPPs) in 2000, and again updated to \$1.25 (in 2005 PPPs) in 2008. While the first update went by almost unnoticed, the most recent change has sparked a controversy. Redefining extreme poverty as living below \$1.25 a day raised the global poverty headcount by about 10 percentage points and reclassified approximately 450 million people as extremely poor (Chen and Ravallion, 2010). In this section, we briefly review the origins of the \$1 a day measure and discuss shortcomings of the current updating procedure.

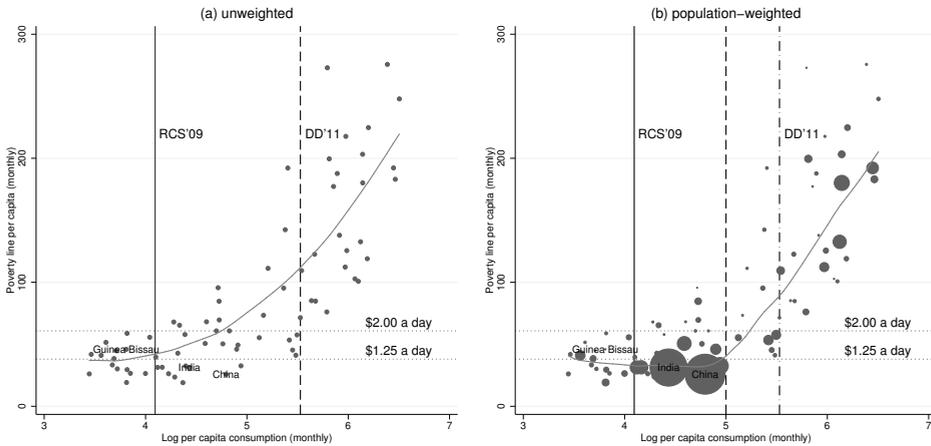
The problem of setting a global poverty line is far from trivial. Even if we could use a ‘basic needs’ or calorie-intake approach to devise a minimum consumption bundle for the entire world, it is inherently difficult to apply any such bundle in international comparisons. Subsistence needs, relative prices, and purchasing power vary across countries and over time. Faced with these problems, Ravallion, Datt, and van de Walle (1991, henceforth RDV) suggested an original solution. Since many national poverty lines are set using a basic needs or calorie-intake method, there should be a universal lower bound among the absolute poverty lines which may be recovered from the data. Converting 33 national poverty lines and the corresponding consumption levels from the 1970s and 1980s into international dollars, RDV showed that a poverty line of about \$31 per month (\$1.02 a day, in 1985 prices) was shared by the six poorest countries in their sample, while those of two other countries came close. They argued that a rounded-off poverty line of \$1 a day was a sensible threshold for measuring global poverty, since any one poverty line is likely to be estimated with error and the non-food allowance included in the subsistence basket varies across countries. RDV also estimated a lower line of \$23 per month (about 76 cents a day) for the poorest country in their sample. This lower line was close to India’s poverty line at the time and became widely used as the international poverty line during the 1970s and 1980s (e.g. Ahluwalia, Carter, and Chenery, 1979).

Setting the poverty line in international prices has the advantage that domestic inflation is typically taken into account when average incomes or expenditures from surveys are converted into (base year) international dollars, so that the line itself does not have to be explicitly updated annually. However, purchasing power parities (PPPs) change over time as countries grow richer (due to the Balassa-Samuelson effect). In addition, the quality of PPP estimates has been improving substantially with each round of the International Comparison Program (ICP), so that updates are needed approximately every decade. When the 1993 ICP data became available, Chen and Ravallion (2001) revised the \$1 a day line to \$1.08 in 1993 prices. Using the same data as in the original study, they found that \$1.08 a day was the median poverty line of the ten poorest countries. However, when the 2005 ICP was completed, instead of converting the old poverty line to 2005 prices, new data were collected and the poverty line was redrawn. Ravallion, Chen, and Sangraula (2009, henceforth RCS) compiled a data set of 74 national poverty lines to update the original analysis. They found that national poverty lines do not rise with per capita consumption until a certain turning point (about \$60 per month) but increase strongly thereafter (left panel, Figure 1). RCS set the global line as the average poverty line of the 15 countries below this threshold, or \$1.25 a day in 2005 prices.

Deaton (2010), as well as Deaton and Dupriez (2011), take issue with this approach. They argue that updating the international poverty line based on new data leads to

“graduation effects” when countries move out of the reference group. They illustrate their case using India and Guinea Bissau as examples. India was part of the initial reference group in RDV, and both countries appear in RCS’s more recent reference group. India has a relatively low poverty line (\$0.90 a day in 2005 prices) and a population of more than a billion people, whereas Guinea Bissau has a higher poverty line (\$1.51 a day in 2005 prices) and is home to less than 1.5 million people. As average consumption in India grew considerably until 2000, it crossed the \$60 threshold and is no longer part of the reference group. Even though the average Indian has become richer, both the international poverty line and the global poverty headcount increased as a result of India dropping out of the average. With Guinea Bissau the case is reversed. Its poverty line is currently part of the average. A move out of the reference group would entail a fall in the global poverty line and a reduction in global poverty that is many times greater than the population of Guinea Bissau. The left panel of Figure 1 illustrates this relationship. The bold horizontal line marks the \$60 per month threshold (labeled RCS’09).

Figure 1 – Poverty lines and consumption levels around 2000



Notes: Authors’ calculations using the data reported in Ravallion et al. (2009) and following the illustration of Deaton (2010). The non-linear trends are estimated using a (weighted) local linear smoother with bandwidth 0.8.

A related issue is that the ICP data are primarily designed for comparing living standards of entire populations, not just poor people. The typical consumption basket of the poor, and the associated price level, may be very different than the reference basket used for computing PPPs.

To address both the graduation issue and the PPP issue, Deaton and Dupriez (2011) propose using an alternative procedure. Linking consumption surveys to ICP data for the 50 poorest countries, they simultaneously estimate the poverty line and PPPs of those near the poverty line (PPP for the poor, or P4s). This yields lower poverty lines in between \$0.92 and \$1.19 a day. However, the effect of the P4s on the global poverty counts – at similar poverty lines – is relatively small. The resulting estimates of global poverty are lower primarily due to the lower poverty lines and not due to differences in relative prices.

To an extent, the Deaton-Dupriez criticism can be addressed within the RCS approach by (1) weighting the national poverty lines by population sizes, and (2) extending the reference group of “poorest” countries. The Deaton-Dupriez proposal, labeled DD’11 below, is to select the 50 poorest countries to constitute a fixed reference group. Clearly, the threshold of 50 countries is arbitrary. A possible alternative is to replicate the RCS approach but estimate the consumption gradient using population weights instead of equal weights (right panel, Figure 1). This is the approach proposed here. Examining the plot to find the point where the slope of consumption begins to be positive, we visually identify a threshold of about 5 log dollars or \$148.41 per month (C^*), which is just under \$5 a day. There are 39 countries below this line, suggesting an alternative reference group consisting of those 39 countries. A population-weighted average of the national poverty lines for this group implies a global poverty line of \$1.06 a day whereas the simple average is \$1.46 a day (both in 2005 prices).

Table 1 – Estimates of the international poverty line (Z^*)

Estimation Sample	Consumption		International Poverty Lines (in 2005 PPPs)					
	$\ln(C^*)$	C^*	<i>Equally-weighted</i>		<i>Population-weighted</i>			N
			\bar{Z}	Z^*	$Q_{50}(Z)$	\bar{Z}	Z^*	
RCS’09	4.09	60.00	1.25	1.25	1.27	1.17	1.17	15
Alternative	5.00	148.41	1.46	1.46	1.38	1.06	1.06	39
DD’11	5.53	251.59	1.76	1.26	1.51	1.12	1.00	50
Full sample	6.50	668.31	2.91	1.46	2.00	1.84	1.06	74

Notes: Authors’ calculations using the data reported in Ravallion et al. (2009). \bar{Z} is the (unweighted or weighted) average poverty line. $Q_{50}(Z)$ is the (unweighted) median poverty line. The regression based columns estimate the average level of the poverty line before the consumption gradient turns positive. Following Ravallion et al. (2009), we obtain Z^* from $Z_i = Z^* I_i + (\alpha + \beta C_i) \times (1 - I_i)$, where Z_i is the national poverty line expressed in 2005 PPPs, Z^* is the mean poverty line for the reference group, C_i is average per capita consumption in 2005 PPPs, and I_i is one if $C_i \leq C^*$ and zero otherwise; that is, it indicates whether the country is in the reference group. This method imposes that the slope is zero until a monthly consumption level C^* .

Table 1 provides alternative estimates of the international poverty line using different approaches, references groups and weights. We can fully reproduce the main findings of RCS. The average poverty of the poorest 15 countries is \$1.25 a day and the full sample median is \$2 a day. Three other results stand out. First, all estimates are above or equal to \$1 a day in 2005 prices. Second, population weights lead to universally lower estimates of the poverty line. Third, the estimated poverty line is fairly sensitive to both the choice of C^* and the summary measure (mean, median, weighted or not). While this lends some support to the notion that \$1.25 is a relatively high upward revision, the choice of method, cutoff and weights remains subjective. Moreover, updating the old \$1.08 poverty line in 1993 prices to \$1.25 in 2005 prices implies a (global) inflation rate of about 1.2% per year. Precisely this point leads Chen and Ravallion (2010) to argue that “as long as it is agreed that \$1 in 1993 international prices is worth more than \$1 at 2005 prices, the qualitative result that the new ICP round implies a higher global poverty count is robust” (p. 1612). In the following section, we will consider the policy consequences of using both \$1.25 a day and the full-sample median of \$2 a day as absolute poverty lines (a common practice in the literature, e.g. see Chen and Ravallion, 2010).

Counting the global poor involves several difficult methodological choices on top of choosing an international poverty line that can have large effects on the estimated poverty rates.¹ The inherent difficulty of convincingly solving these issues led some to suggest that the current approach should be abandoned (e.g. Klasen, 2009; Reddy and Pogge, 2010). One alternative is to set national poverty lines in local currencies using the same method in each country. While this approach tries to sidestep the issue of purchasing power parity comparisons altogether, it would certainly raise new problems. Alas, a considerable degree of indeterminacy regarding the level of extreme poverty in the world seems unavoidable (but more is known about rates of change which we discuss in the next section).

3 Taking stock: poverty reduction over the past three decades

While the global Millennium Development Goal (MGD) of halving the \$1.25 a day poverty rate in 1990 by 2015 was reached in 2010 (Chen and Ravallion, 2013), progress has been very uneven across regions. Most poverty reduction over the last three decades occurred in East Asia and, to a lesser extent, in South Asia. Poverty in Sub-Saharan Africa, on the contrary, has hardly budged and the continent as a whole will – most probably – fail to meet the first MDG by 2015.

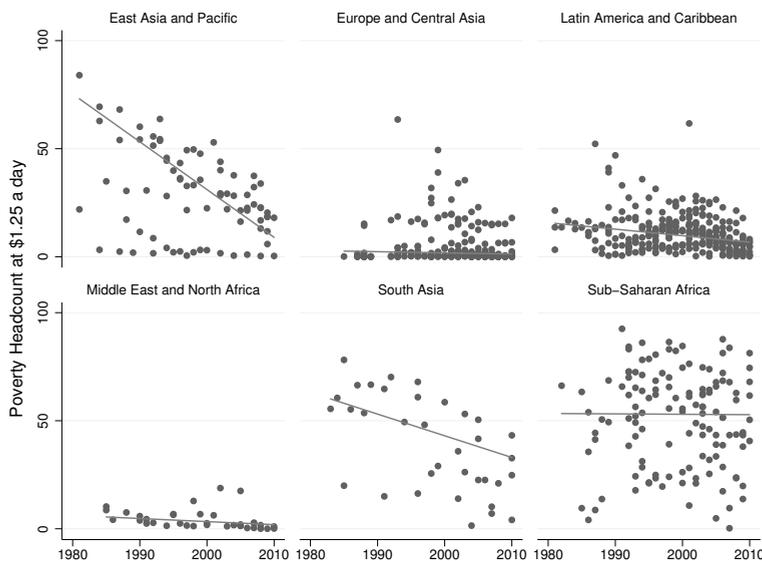
Figure 2 shows population-weighted time trends estimated for each region using the nationally representative household surveys available in the World Bank's *PovcalNet* database (using consumption-based data where available). A pooled weighted least squares regression of the poverty headcount ratio on time for all developing countries reveals that the poverty headcount at the \$1.25 a day line fell by an average 1.5 percentage points per year (cluster $t^2 = -4.11$). There is substantial regional heterogeneity. On average, poverty fell by 2.21 percentage points annually in East Asia and Pacific (cluster $t = -10.08$), by about 1 percentage point in South Asia (cluster $t = -4.82$), but only by 0.02 percentage points in Sub-Saharan Africa (cluster $t = -0.05$). In the other three developing regions progress has been slow but steady. In Europe and Central Asia, Latin America and Caribbean, and Middle East and North Africa the estimated slopes imply an expected annual fall of 0.05 (cluster $t = -1.91$), 0.30 (cluster $t = -5.76$) and 0.14 (cluster $t = -2.40$) percentage points, respectively. Using an alternative \$2 a day poverty line, the magnitudes and differences in speeds across regions remain broadly similar.³

An important question is whether consumption growth or redistribution is driving the decline in poverty. Estimating the historical contributions of growth and changes in distribution during the 1980s and 1990s, Kraay (2006) found that most poverty reduction was due to income or consumption growth. Our analysis broadly corroborates this finding

¹For example, it is not clear that the ICP 2005 provides an adequate picture of the consumption patterns in the 1980s or 1990s. Another issue is the use of survey means versus national accounts means. The use of national accounts typically leads to much lower poverty estimates. See Sala-i-Martin (2006), Pinkovskiy and Sala-i-Martin (2009), and, in particular, Dhongde and Minouin (2013) for a comparison of different methods.

²Throughout the text, “cluster t ” denotes a cluster-robust t -statistic, with clusters defined by countries.

³East Asia and Pacific (slope = -2.31, cluster $t = -7.33$), Europe and Central Asia (slope=-0.26, cluster $t=-2.33$), Latin America and Caribbean (slope=-.55, cluster $t = -5.74$), South Asia (slope = -0.72, cluster $t = -4.92$), and Sub-Saharan Africa (slope = -0.02, cluster $t = -0.07$).

Figure 2 – Population-weighted poverty trends by region, 1981 to 2010, \$1.25 a day

Notes: Authors' calculations based on surveys from *PovcalNet*.

(although we do not explicitly estimate contributions). The population-weighted growth rate of the survey means from 1981 to 2010 across all countries is a very robust 4% per year (cluster $t = 3.03$). Over the same period, within-country inequality, as measure by the Gini coefficient, actually increased slightly by about 0.7% per year (cluster $t = 1.64$). This implies that, on average, changes in distribution may have in fact moderately slowed the pace of poverty reduction. Poverty reduction over the last three decades has mostly been due to income and consumption growth. However, both the high average growth rate in the survey means and the apparent rise in within-country inequality are driven by China. Excluding China, the survey means grew about 1.8% per year (cluster $t = 2.45$) and inequality barely moved (increased 0.047% per year, cluster $t = 0.13$). In other words, poverty reduction in the developing world outside China has been steady but slow and has (on average) not been helped by improvements in distribution.⁴

These findings are also in line with estimates of poverty at the \$1.25 a day poverty line reported by the World Bank (see Appendix Table A-1). Chen and Ravallion's (2010) estimates indicate rapid progress in China, little improvement in Sub-Saharan Africa, and moderate poverty reduction elsewhere. The poverty headcount ratio in Sub-Saharan Africa only fell by about three percentage points over the entire period from 1981 to 2010, and actually exceeded its 1981 value for most of the period. Combining these trends with population growth rates reveals the dire absence of a robust positive trend in terms of the number of global poor outside of East Asia. While China has lifted an astonishing 680 million people out of poverty between 1981 and 2010, the rest of the world has only about 50 million fewer extremely poor people in 2010 than in 1981. This

⁴Excluding India in addition to China from the sample does not qualitatively alter this result.

trend is due to persistently high poverty rates coupled with strong population growth in Sub-Saharan Africa and India. This is most evident in Sub-Saharan Africa where the number of extremely poor has roughly doubled over three decades (in spite of the slight decrease in the headcount ratio). The rise of China from a poor to a middle income country also implies that the relative composition of world poverty is changing rapidly. In 1981 about 40% of the world's extremely poor lived in Sub-Saharan Africa and South Asia, by 2010 their share has risen to 75%.

A very intuitive approach to illustrating past progress (or lack thereof) is to approximate the shape of the income or expenditure distribution at various points in time and examine how the features of the distribution (esp. quantiles) shift over time. Figures 3 and 4 plot the lower tail (up to \$400) of the monthly income or expenditure distribution for the most populous country of the four poorest regions – East Asia, South Asia, Latin America, and Sub-Saharan Africa – in 1985, 1990, 2000 and 2010. The vertical lines are the \$1.25 and \$2 a day poverty lines in terms of monthly consumption. After lining up the survey data in time, we estimate the different density functions using a log-normal approximation.⁵ While the assumption of log-normality has its weaknesses⁶, it usually provides a useful first estimate of the shape of the income distribution. A key advantage is that it only requires knowledge of the mean and Gini coefficient.

We can illustrate a few essential concepts with these graphs. The area under the curve to the left of the poverty line gives the fraction of the population that is poor (the poverty headcount ratio), while the spread of the distribution reflects inequality. The raw difference between two such areas under the curve is the absolute change in the poverty headcount ratio in percentage points and the relative difference gives the percent change in the poverty headcount ratio. The sensitivity of poverty reduction to changes in income or inequality is often measured in the form of elasticities or semi-elasticities. The income elasticity of poverty is the percent change in poverty for a one percent increase in incomes, and the income semi-elasticity of poverty is the percentage point difference in poverty for a one percent increase in incomes. The inequality elasticity and semi-elasticity are defined analogously. An attractive feature of the semi-elasticity is that it first increases and then decreases again during the development process. It measures the pace of poverty reduction in terms of the percentage of the population lifted out of poverty. Hence, it is usually more informative for policy-makers and more useful than reporting relative changes.⁷

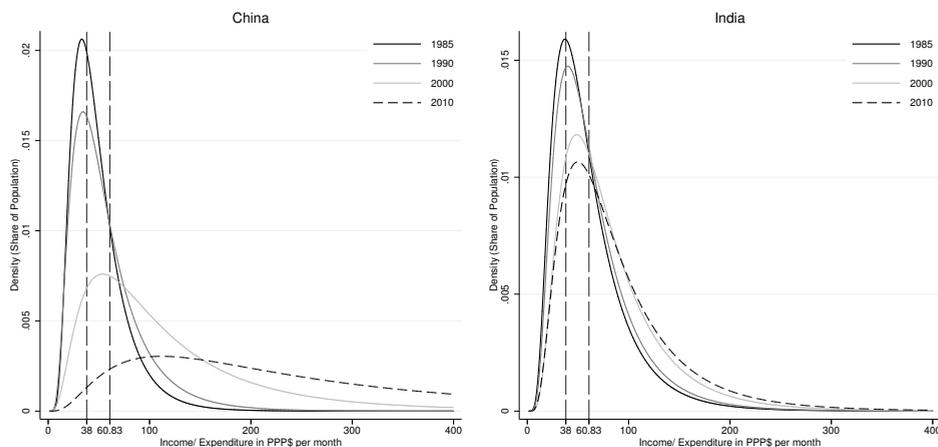
Figure 3 visualizes the tremendous progress in reducing poverty rates in China over the last three decades.⁸ As noted before, poverty in China at \$1.25 a day fell rapidly

⁵We interpolate and extrapolate the data as follows. First, we project mean consumption forward and backward using the corresponding growth rates of personal consumption expenditures from the national accounts. Second, we linearly interpolate between the available Gini coefficients and extrapolate beyond the first or last available measure by keeping inequality constant. The same data set (with all countries from *PoocalNet*) is later used for computing the inequality indices in the developing world.

⁶Log-normality typically works better with consumption surveys than with income surveys (Lopez and Serven, 2006), tends to underestimate the level of poverty (Dhondge and Minoiu, 2013), and overstates the pace of poverty reduction (Bresson, 2009).

⁷In relatively rich countries with low percentages of people below the poverty line, elasticities can be very misleading. Small reductions in the poverty headcount rate can manifest themselves as very high elasticities. For a more detailed discussion of the properties of elasticities and semi-elasticities of poverty see Bourguignon (2003), Klasen and Misselhorn (2008), and Chapter 5.

⁸The implied poverty rates for China correspond well with the official World Bank estimates. At the \$1.25 a day poverty line, our estimates imply a poverty rate of 60.56% in 1985, 56.92% in 1990, 31.97% in 2000 and 9.75% in 2010.

Figure 3 – Estimates of the expenditure distribution: China and India, 1985–2010

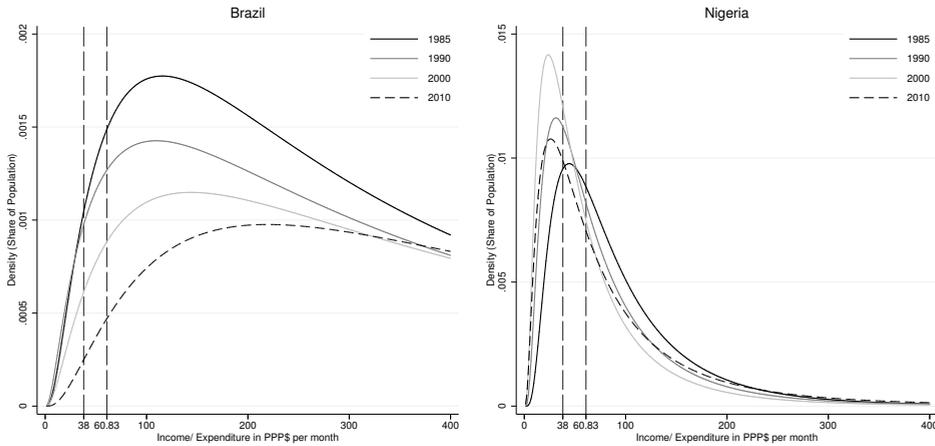
Notes: Author’s calculations. China’s expenditure distribution is estimated based on a weighted mean and a rural-urban ln-mixture for the Gini coefficient. China’s surveys in *PovcalNet* are consumption-based after 1987 and income-based before.

over the entire period. The biggest gains occurred early on, between 1985 and 2000, when the peak of the distribution was close to the \$1.25 and \$2 poverty lines. By 2010, the peak of the distribution has moved considerably to the right of both poverty lines and the overall spread has widened. A great many Chinese are now considered part of a developing country ‘middle class’ (if defined between \$2 and \$13 per day).⁹ However, this also implies that the poverty reduction potential from China is largely exhausted. The income semi-elasticity of the poverty headcount is far beyond its peak and steadily approaching zero. In addition, inequality has increased remarkably over the same period. In 1985, the Gini coefficient was 0.28 and by 2010 it has risen to 0.44.

In India, on the contrary, there remains much greater potential for poverty reduction in the medium-term future. While the mode of the income distribution was near the \$1.25 line around 1985 and 1990, the peak of the distribution in 2000 and 2010 is located between the two poverty lines. The process of “bunching up” in front of \$2 a day observed by Chen and Ravallion (2010) implies that, in the medium-term future, the pace of poverty reduction in India (defined as the absolute change in the headcount) will be particularly fast at the \$2 a day line and continue at a fast but decelerating pace at the \$1.25 line. Put differently, India’s income semi-elasticity around 2010 is very high and a moderate rate of growth will immediately have a large (but decreasing) effect on the poverty headcount ratio at both thresholds.

Figure 4 illustrates two very different cases. The left panel shows that from 1985 to 1990 poverty reduction in Brazil was very slow, with some progress at the \$2 a day line but a nearly unchanged poverty rate at the \$1.25 line. Yet, on average, Brazilians were already considerably better-off in the 1990s than their Chinese or Indian counterparts in

⁹Ravallion (2010) defines the size of the ‘middle class’ by developing country standards as the proportion of the population living on at least \$2 per day but less than \$13 per day, where the upper bound is the poverty line in the United States. Naturally, this is one of many possible definitions.

Figure 4 – Estimates of the expenditure distribution: Brazil and Nigeria, 1985– 2010

Notes: Authors' calculations. Brazil's distribution is based on incomes instead of expenditures. The 2010 data for Nigeria was revised in late 2014 (not incorporated here).

2010. After 1990, the pace of poverty reduction accelerates and by 2010 only 4.92% of the population were below the \$1.25 a day poverty line.¹⁰ Lifting the remaining people out of poverty will require sustained economic growth, as both the income and distribution semi-elasticities of poverty in Latin America as a whole are rather low (see Chapter 5). With a Gini of 0.56 in 1985 income inequality was initially very high in Brazil, peaked at 0.61 in 2000 and then fell again to 0.55 by 2010, thus positively contributing to poverty reduction after 2000. The right panel illustrates that poverty in Nigeria was considerably higher in 2000 or 2010 than in 1985. Nigeria's plight is characteristic for most of Sub-Saharan Africa in the 1980s and 1990s, as real consumption on the subcontinent was declining at a pace of about 0.82% per year. Only after 2000 did expenditures recover and the poverty headcount ratio began to decline. Yet even by 2010, the peak of the expenditure distribution is still noticeably to the left of the poverty line and the implied poverty rate at \$1.25 a day is 65.96%.¹¹ In addition, inequality in Nigeria increases over the observed period, starting from a Gini of 0.39 in 1985 to 0.49 in 2010.

Taken together, these four distributions exemplify the changing composition of global poverty and broadly represent the trends in their respective regions. Over the last three decades, most poverty reduction occurred in East Asia where consumption growth was fastest, some poverty reduction occurred in India where real consumption growth was steady, and little poverty reduction occurred in Sub-Saharan Africa where real consumption growth was slow and volatile. This suggests that without significantly faster growth in Sub-Saharan Africa than in the past, possibly coupled with improvements in the income or expenditure distribution, the global pace of poverty reduction will inevitably slow down in the near future.

Another essential aspect of poverty analysis is studying the evolution of inequality. In this part, we focus only on inequality among citizens of developing countries, as

¹⁰The World Bank estimated a poverty rate of 5.38% at \$1.25 a day for Brazil in 2010.

¹¹The World Bank estimated a poverty rate of 67.98% at \$1.25 a day for Nigeria in 2010.

our interest is the changing relative position of people in the developing world rather than their position vis-à-vis rich countries. Interestingly, many of the global trends are also evident even when we restrict our attention to this truncated distribution. We compute three measures of inequality by applying Young's (2011) mixture of log-normal distributions approach to the *PovcalNet* data. 'Overall inequality' is the Gini coefficient for citizens of developing countries regardless of their country of residency. 'Within inequality' is a population-weighted summary measure of inequality within each country. Last, 'between inequality' is the population-weighted Gini coefficient of average incomes among all developing countries. In other words, the first measure encompasses both the within-country and between-country components that make up overall inequality in the developing world. Naturally, global inequality – including the citizens of developed countries – is typically estimated to be considerably higher. Recent estimates of the global Gini suggest that it is around 0.65-0.70, and may be even higher if underreporting of top-incomes is taken into account (e.g. see Pinkovskiy and Sala-i-Martin, 2009; Milanovic, 2012; Chotikapanich, Griffiths, Rao, and Valencia, 2012; Lakner and Milanovic, 2013).

Table 2 – Inequality in the developing world, 1980–2010

Year	<i>Gini coefficient</i>			Mean Consumption	Population	N
	Overall	Within	Between			
1980	0.596	0.356	0.486	73.29	2907.8	83
1985	0.555	0.353	0.421	79.05	3223.2	86
1990	0.578	0.367	0.449	95.98	4049.3	104
1995	0.559	0.385	0.411	98.41	4555.2	114
2000	0.537	0.395	0.374	102.45	4931.4	121
2005	0.535	0.399	0.372	120.34	5285.6	123
2010	0.554	0.404	0.399	150.72	5625.1	123
Δ 1980-2010 (in %)	-7.19	13.63	-17.82	–	–	–
Δ 1990-2010 (in %)	-4.17	10.10	-11.08	–	–	–
Δ 2000-2010 (in %)	3.07	2.20	6.64	–	–	–

Notes: Authors' calculations. The sample size varies over the years. A total of 124 countries are recorded in *PovcalNet* but we lack PCE data for West Bank and Gaza. The results are very similar if we constrain the developing world to consist of the 104 countries from which we have (interpolated) data from 1990 onwards. Due to the lower coverage, the results for the 1980s should be interpreted with caution. For details on the ln-mixture calculations refer to Young (2011).

Table 2 reveals some interesting trends. Overall inequality in the developing world has been falling between 1990 and 2005, but it exhibits an increase in 2010. At the same time, within-country inequality has been rising steadily since the mid-1980s. Between-country inequality fell over most of the period but also shows a slight increase between 2005 and 2010. If we exclude China from the computations given that its weight is very high, then these trends are considerably muted or even non-existent.¹² Hence, two

¹²Overall inequality is estimated as 58.33 in 1990 and 58.36 in 2010, within-country inequality is estimated as 38.36 in 1990 and 39.12 in 2010, and between-country inequality is estimated as 45.04 in 1990 and 45.44 in 2010. Removing India in addition to China has little effect on the trends in the inequality measures.

developments drive the overall change. First, inequality of incomes within China has been increasing significantly and, second, its relative position among developing countries has been changing rapidly. Rising mean incomes in China from the 1980s onwards initially implied a reduction of between-country inequality as the average citizen in China was moving from the bottom towards the middle of the developing country ranks, but they now put upward pressure on overall inequality as incomes in China continue to grow and the distance from incomes in Sub-Saharan Africa increases.¹³

4 Going forward: poverty projections until 2030

As the expiration date of the MDGs is approaching quickly, new goals will have to be selected. Picking from among a wide range of possible benchmarks invariably involves formulating expectations towards a fundamentally uncertain future. Thus, it becomes important to ask: what can the current data and methods tell us about the prospects for poverty alleviation over the next two decades? The list of policy-relevant questions is long. What level of poverty do we expect to prevail in 2030? Will it be feasible to truly eradicate extreme poverty by 2030? Or, how quickly do we expect poverty rates under the \$2 a day poverty line to decrease? Here, we provide both a glimpse into several likely futures and some potential answers to these questions.

This section draws heavily on Chapter 5, where we develop a ‘fractional response approach’ for estimating income and inequality (semi-) elasticities of poverty. Among other things, we show that this new method can be used to easily forecast global poverty rates using only two variables (the survey mean and the Gini coefficient). A key advantage of this approach over, say, linear trend extrapolations, is that it builds in the non-linearity of the poverty-income-inequality relationship. Neither the income or inequality elasticity nor the income or inequality semi-elasticity is assumed to be constant. The method accounts for the fact that income growth will have an increasing effect in very poor countries, where the mass of the distribution is to the left of the poverty line, and less and less of an effect in rich countries, where the mass of the distribution is far to the right of the poverty line.¹⁴ Similarly, the effect of changes in distribution will indirectly depend on the prevailing levels of both income and inequality.

We are of course not the first to present poverty projections over the next two decades. Ravallion (2013), for example, outlines an aspirational scenario where an additional billion people are lifted out of extreme poverty by 2025-2030. Karver, Kenny, and Sumner (2012) discuss the future of the MDGs more generally and simulate poverty rates at the \$1.25 and \$2 a day poverty lines for 2030.¹⁵ Yet there are some important conceptual and

¹³This trend is corroborated by the literature on global inequality. According to Lakner and Milanovic (2013), average incomes in Sub-Saharan Africa were \$742 in 1988 and only \$762 in 2008 (in 2005 PPPs), while Chinese incomes increased by 228.9% and no longer make up a large part of the lower tail of the global income distribution. They also show that inequality within China has risen between 1988 and 2008.

¹⁴The inability to account for countries that have relatively high incomes and zero poverty at some point in time (typically the beginning or end of a spell) is a key weakness of studies investigating poverty elasticities.

¹⁵Karver et al. (2012) allow for country-specific growth rates but use older data (their *PovcalNet* reference year is 2008) and disregard the difference between GDP per capita growth and growth of the survey mean. This leads them to overestimate the speed of poverty reduction relative to our forecasts. A recent study by Chandy, Ledlie, and Penciakova (2013) echoes some of our results. They use GDP per capita rather than consumption expenditure data for most of the period, but apply a conversion factor,

methodological differences between our approach and these studies. First, the assumption that the developing world will continue growing at the accelerated 2000 to 2010 pace for another twenty years (our optimistic scenario) is questionable. There is a well-known instability of growth rates across decades that should not be ignored (Easterly et al., 1993), especially since the high average growth rates in the developing world were driven by rapid growth in China. A more conservative assumption is that countries will grow at rates much closer to their individual long-run growth path. Second, the changing composition of the countries contributing to global poverty matters a lot for the expected speed of global poverty reduction. Unless there is a persistent acceleration of consumption growth in Sub-Saharan Africa on top of the post-2000 growth rates and sustained consumption growth in India, we can show that the pace of poverty reduction at the \$1.25 line is likely to experience a pronounced slowdown in all of our forecast scenarios (defined below). Third, pro-poor growth can potentially make a sizable difference in the expected poverty rates, while a rise in within-country inequality will hasten the arrival of the slowdown. Fourth, our method approximates the ‘official’ *PovcalNet* results at a fraction of the computational cost, so that a variety of scenarios can be easily estimated (and frequently updated with the arrival of new data).

We define three different constant growth scenarios on the basis of the historical personal consumption expenditure (PCE) growth rates from the national accounts.¹⁶ An ‘optimistic’ scenario uses the average PCE growth rate of each country from 2000-2010, during which period growth rates were significantly higher than before 2000. A ‘moderate’ growth scenario uses the average PCE growth rate of each country from 1980 to 2010 – the long run average over the entire data set. Finally, a ‘pessimistic’ growth scenario uses the 1980 to 2000 average PCE growth rates. The latter scenario assumes that mean consumption in Sub-Saharan Africa is shrinking at a rate of about 0.82% per year.¹⁷ Table A-2 in the Appendix reports the population-weighted average regional growth rates over several different periods to illustrate the implied regional income dynamics.

For each growth scenario, we also simulate three different inequality patterns. ‘Pro-poor growth’ implies an annual decline in the Gini coefficient of approximately -0.5%, ‘distribution-neutral growth’ keeps inequality constant at the level prevailing in 2010, and ‘pro-rich growth’ implies an increase in the Gini coefficient of approximately 0.5% per year.¹⁸ As an illustration, if a country’s Gini coefficient is 0.40 in 2010 and we apply the pro-poor pattern, then by 2030 we project a Gini coefficient of about 0.36. If we apply the pro-rich pattern, then the Gini coefficient is about 0.44 in 2030. Changes of this magnitude are in line with the population-weighted regional trends obtained from the surveys.

We forecast the poverty rates until 2030 as follows. First, we estimate the model outlined in Chapter 5 for the \$1.25 a day poverty line using all nationally-representative surveys recorded in *PovcalNet* over the period from 1981 to 2010. Next, after lining up all surveys in 2010, we apply each of the nine growth and distribution scenarios to project the income and inequality data forward to 2030, country by country.¹⁹ Then, we predict

and report lower poverty estimates.

¹⁶The term ‘national accounts’ refers to data from the World Development Indicators or the Penn World Table 7.1, whichever has more data over the 30 year horizon.

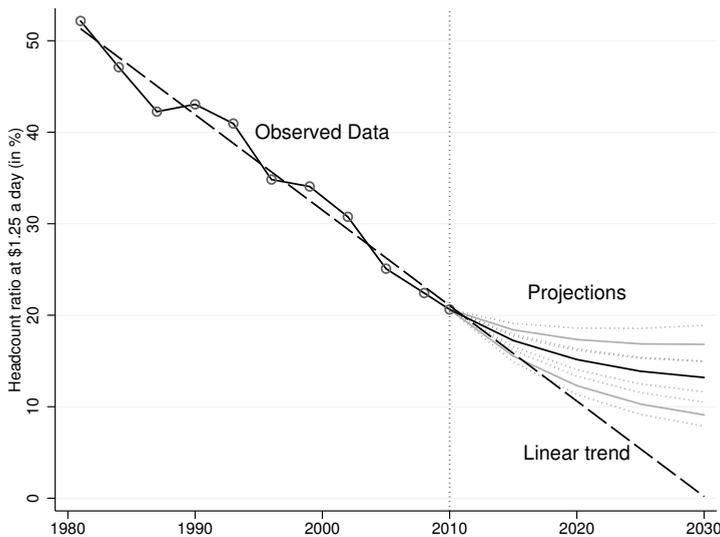
¹⁷Owing to the post-communist transition, consumption and incomes in Europe and Central Asia were shrinking over the same period. However, given the small number of poor in 2010, the influence of that region on the global poverty headcount in 2030 is minimal.

¹⁸All reported growth rates (in percent) are computed as log differences if not otherwise noted.

¹⁹To line up all surveys in 2010, we use the actually observed PCE growth rates from the national

the poverty headcount ratios in five-year intervals over the period 2015 to 2030, country by country. Finally, we calculate population-weighted regional poverty rates and apply these to the projected total population in each region. For consistency with *PovcalNet*, the population projections are also taken from the World Bank and the ‘developing world’ is defined as in 1990 – the countries targeted by the MDGs – no matter how high we forecast the average level of consumption to be in 2030. Contrary to the World Bank’s recent redefinition of the denominator, we still focus on the percent of poor population in the developing world and *not* the entire world.

Figure 5 – Actual and projected poverty headcount ratios at \$1.25 a day, 1981–2030



Notes: Authors’ calculations based on Chapter 5 and survey data from *PovcalNet*. The solid black line beyond 2010 refers to the moderate (distribution-neutral) growth scenario in Table 3, while the solid grey lines represent the distribution-neutral variants of the optimistic and pessimistic scenarios. The pro-poor and pro-rich variants are shown as grey dotted lines and are located above or below a solid line.

Figure 5 plots the historical evolution of the poverty headcount from 1981 to 2010, a linear trend fitted through the observed data and then extrapolated until 2030, and our different scenarios. The linear trend serves as a reference for the non-linear projections. Several points are noteworthy. First, only the linear extrapolation predicts a poverty rate in the vicinity of zero by 2030. Regressing the global poverty rate at \$1.25 a day on time one obtains a slope of about one percentage point per year (see also Ravallion, 2013).²⁰ As the global poverty rate was about 20.6% in 2010, the linear trend predicts

accounts to extrapolate the survey means from the latest available survey. In doing so, we keep inequality constant at the last observed Gini coefficient. In 2010, the average year when the last survey was conducted is 2006.7, so about 3 years prior to 2010. More than 40% of the last surveys were conducted in 2009 or 2010.

²⁰This differs from the 1.5 percentage points estimated in the previous section as the global poverty rate is measured by lining up and weighting all surveys at reference years (three year intervals from 1981 onwards), whereas in the previous section we were using an unbalanced panel of unequally-spaced,

that extreme poverty will have vanished by 2030. Second, all our projections show a decelerating rate of poverty reduction. Even in the most optimistic scenario, the pace of poverty reduction slows down. Most forecasts show a decelerating trend early on. In the optimistic scenario the slowdown only becomes noticeable by about 2020. Third, all scenarios but the optimistic pro-poor growth or optimistic distribution-neutral growth scenarios imply a poverty rate higher than 10% in 2030 at the \$1.25 a day line. The optimistic pro-poor growth and distribution-neutral scenarios suggest a poverty rate in 2030 of 7.9% and 9.1%, respectively. In a nutshell, 2030 is not likely to mark the end of extreme poverty, even under very optimistic assumptions. Our projections suggest that the World Bank's goal of 3% extreme poverty in 2030 is not likely to be reached.

Table 3 provides the corresponding regional and total poverty rates in 2030 including the expected number of poor in the various scenarios. Our moderate growth estimate suggests a global poverty rate of 13.2% in 2030, implying about 950 million poor versus 1.2 billion poor people in 2010. The pace of poverty reduction will have slowed significantly both in terms of relative changes and in terms of numbers of poor people. In this scenario, about 70% of the world's poor live in Sub-Saharan Africa and about 23% in South Asia by 2030. In contrast, the (distribution-neutral) optimistic result suggests a poverty rate of 9.11%, with about 655 million people remaining extremely poor. About 76% percent live in Sub-Saharan Africa and about 17% in South Asia. The pessimistic case suggests next to no progress at all. Given an unchanged distribution, the poverty headcount ratio is estimated at 16.82% and the world is still home to 1.2 billion extremely poor people. Even if growth rates in Sub-Saharan Africa were to *double* relative to the post-2000 trend, the global poverty rate in 2030 is still projected to be 6.50% with pro-poor growth, 7.67% with distribution-neutral growth, and 8.81% with pro-rich growth.

All of these estimates imply that it will take considerably longer than 2030 to lift the remaining 1.2 billion people out of poverty. The good news is that by 2030 extreme poverty in Europe and Central Asia, East Asia, Latin America, and Middle East and North Africa may virtually disappear (projected to be less than 5% in most forecasts). However, we predict a strong increase in the (relative) share of global poverty located in Sub-Saharan Africa, which suggests that a non-trivial fraction of extreme poverty may be concentrated in 'fragile states'. Whether these countries will overcome civil strife, political instability and corruption will ultimately decide whether there is a lower bound at which extreme poverty will continue to exist.

Gradual changes in inequality raise or lower the overall headcount in between 1.2 and 2.1 percentage points and account for about 100 million poor people more or less. Contrary to suggesting that inequality does not matter (we only assume slow changes), this finding hints at two crucial points. First, if the developing world as a whole is to truly maintain the impressive record in poverty reduction of the last decades, then this requires both sustained high growth at the level experienced since 2000 and improvements in distribution. Second, any systematic worsening of within-country inequality, particularly in large and largely poor countries like India or Nigeria, will reinforce the slowdown and thus more strongly decelerate the global rate of poverty reduction.

Readers may wonder why these results are so different from the projections reported in Ravallion (2013). Our results differ mainly because Ravallion (2013) uses the average growth rate of the developing world to project poverty in countries with very different track records, while we use country-specific average growth rates. Otherwise there are only minor differences in the data used and our method closely approximates results

population-weighted survey data with a wide yet somewhat selective coverage.

Table 3 – Projected poverty headcount ratios and poor population at \$1.25 a day in 2030, by region

	Average PCE Growth								
	Optimistic (2000-2010)				Pessimistic (1980-2000)				
	pro-poor	neutral	pro-rich	pro-poor	pro-poor	neutral	pro-rich	pro-poor	
	<i>Change in Inequality (Gini)</i>								
	<i>Panel (a) – Headcount at \$1.25 a day in 2030 (in percent)</i>								
East Asia and Pacific	0.65	0.93	1.31	0.76	1.07	1.48	0.94	1.29	1.74
Europe and Central Asia	0.12	0.16	0.21	1.21	1.45	1.71	5.17	5.74	6.44
Latin America and Caribbean	2.27	2.74	3.28	3.46	4.12	4.91	4.59	5.46	6.48
Middle East and North Africa	0.48	0.66	0.91	1.54	2.07	2.75	3.72	4.77	6.05
South Asia	4.19	5.54	7.24	8.48	10.89	13.79	12.76	15.99	19.77
Sub-Saharan Africa	32.09	35.69	39.37	43.62	47.17	50.70	51.75	55.12	58.47
Total	7.88	9.11	10.49	11.63	13.20	14.96	14.96	16.82	18.89
	<i>Panel (b) – Poor population at \$1.25 a day in 2030 (in millions)</i>								
East Asia and Pacific	14.05	20.23	28.56	16.59	23.29	32.22	20.44	28.00	37.87
Europe and Central Asia	0.59	0.76	0.97	5.72	6.86	8.11	24.47	27.16	30.45
Latin America and Caribbean	16.15	19.44	23.32	24.61	29.30	34.86	32.61	38.77	46.05
Middle East and North Africa	2.12	2.94	4.04	6.84	9.18	12.18	16.47	21.15	26.83
South Asia	83.47	110.38	144.35	169.13	217.08	275.00	254.40	318.89	394.16
Sub-Saharan Africa	449.54	499.97	551.61	611.15	660.85	710.32	725.03	772.26	819.22
Total	567.20	655.36	754.95	836.34	949.49	1076.37	1076.35	1209.95	1359.23

Notes: Authors' calculations. Population projections are from the World Bank's Health, Nutrition and Population Statistics database. The different scenarios are estimated using the fractional response approach outlined in Chapter 5 and the survey data reported in the World Bank's *PovertyNet* database.

obtained using *PovcalNet*. Ravallion (2013) calculates that a PCE growth rate of 4.5% per year may bring the global poverty rate down to 3% by 2027. However, he makes the (in our view implausible) assumption that all developing countries will continue to grow equally fast at this common rate of 4.5%. Likewise, the linear projection of the global poverty rate on time ignores all issues of aggregation and provides an overly optimistic picture of the medium-term future.²¹ Yet composition matters, even if we incorporate the optimistic assumption that the post-2000 trend will continue. The average hides that rapid growth is less likely in some countries than in others. As we have shown, this has direct consequences for when a slowdown will be observed and how strong the deceleration will be. However, even if we assume a uniform growth rate for all developing countries, a deceleration appears sooner or later within the next two decades (although it may actually be preceded by a brief acceleration if we assume growth rates in excess of 5% p.a.). It is comforting that, in line with Ravallion (2013), our method implies that if consumption in the entire developing world grew at a distributional-neutral pace of 7.6% per year, then extreme poverty would indeed virtually disappear by 2030 (fall to 1.1%). The 3% target can be reached with a uniform distribution-neutral growth rate of approximately 5.5% per year.²²

We repeat this exercise at the \$2 a day poverty line. The results are reported in Table 4 and Figure A-1 in the Appendix. Interestingly, the linear projection is a much better approximation of progress at the \$2 poverty line than at the \$1.25 poverty line. This is not due to a slower historical poverty reduction record: a regression of the global poverty rate at \$2 a day on time also yields a slope of approximately one percentage point per year. However, the composition of countries (or people) near the \$2 a day poverty line in 2010 is more reminiscent of its \$1.25 counterpart in the early 2000s. At the start of the decade in 2010, the total \$2 poverty rate is 40.67% – roughly double the \$1.25 poverty rate. Fast growing East Asia and moderately fast growing South Asia still make up more than half of global poverty, implying that progress in these two regions will have a large effect on the overall poverty headcount.

Our moderate growth scenario predicts that about 1.87 billion people (26%) live on less than \$2 a day in 2030 versus about 2.4 billion people in 2010. Considerably greater gains are possible. Global poverty at the \$2 line falls below 20% in the optimistic distribution-neutral and pro-poor scenarios. If this occurs in 2030, then more than one billion people will have left poverty at the \$2 a day line – undeniably a remarkable achievement. In most scenarios we also observe a slowdown at the \$2 a day line but this slowdown tends to occur later and is less pronounced than at the lower threshold. In the most optimistic scenario, the rate of poverty reduction actually accelerates somewhat to about 1.16 percentage points per year, while the moderate growth scenario gives a trend of 0.73 percentage points per year over the projection period.

Examining the regional distribution, we find that poverty in East Asia is likely to fall to around 5% by 2030, down from 29.7% in 2010. Nearly everyone in East Asia will have entered the middle class (by developing country standards), but this forecast partially hinges on fast growth in China. In fact, some observers suggest that there is

²¹We do not mean to imply that Ravallion (2013) is not aware of the aggregation issues. In fact, he uses *PovcalNet* precisely to confirm that his ‘optimistic scenario’ is possible once the intrinsic non-linearity of the poverty-income-inequality relationship is accounted for. Our point is rather that he envisions “the best possible world” to be used as a benchmark for future progress while we also focus on other, more likely, scenarios.

²²Interestingly, a recent working paper by Yoshida et al. (2014), independently and using different methods, comes to very similar conclusions.

Table 4 – Projected poverty headcount ratios and poor population at \$2 a day in 2030, by region

	Average PCE Growth								
	Optimistic (2000-2010)				Pessimistic (1980-2000)				
	pro-poor	neutral	pro-rich	pro-poor	neutral	pro-rich	pro-poor	neutral	pro-rich
	<i>Change in Inequality (Gini)</i>								
	<i>Panel (a) – Headcount at \$2 a day in 2030 (in percent)</i>								
East Asia and Pacific	3.79	4.61	5.55	4.03	4.90	5.90	4.39	5.31	6.36
Europe and Central Asia	0.45	0.56	0.69	2.80	3.13	3.49	9.35	10.28	11.31
Latin America and Caribbean	4.00	4.73	5.59	6.29	7.39	8.66	8.57	9.99	11.60
Middle East and North Africa	2.85	3.55	4.39	7.20	8.62	10.25	12.86	14.88	17.12
South Asia	19.83	23.12	26.74	31.60	35.88	40.39	40.35	45.00	49.73
Sub-Saharan Africa	51.62	54.56	57.46	63.67	66.36	68.98	70.93	73.32	75.63
Total	17.36	19.23	21.24	23.73	25.94	28.26	28.71	31.07	33.53
	<i>Panel (b) – Poor population at \$2 a day in 2030 (in millions)</i>								
East Asia and Pacific	82.42	100.20	120.61	87.49	106.39	128.23	95.44	115.26	138.11
Europe and Central Asia	2.12	2.63	3.26	13.24	14.79	16.49	44.25	48.63	53.50
Latin America and Caribbean	28.39	33.63	39.73	44.71	52.52	61.50	60.90	70.99	82.42
Middle East and North Africa	12.64	15.74	19.47	31.94	38.22	45.46	57.02	65.95	75.91
South Asia	395.33	461.03	533.21	629.97	715.46	805.25	804.56	897.12	991.46
Sub-Saharan Africa	723.19	764.47	805.00	892.01	929.76	966.48	993.71	1027.19	1059.59
Total	1249.19	1383.61	1528.08	1707.20	1866.01	2033.35	2065.12	2235.43	2412.30

Notes: Authors' calculations. Population projections are from the World Bank's Health, Nutrition and Population Statistics database. The different scenarios are estimated using the fractional response approach outlined in Chapter 5 and the survey data reported in the World Bank's *PovertyNet* database.

reason to believe that China runs a non-negligible risk of falling into a ‘middle-income trap’ (Eichengreen, Park, and Shin, 2013) which might make it harder to achieve less than 10% poverty at \$2 a day by 2030.²³ Progress in South Asia is also likely to be rapid. According to our moderate growth estimate the expected poverty rate is 35.9% in 2030, implying about 716 million poor, down from 66.7% and about 1.1 billion poor in 2010. In the optimistic pro-poor growth case, the headcount ratio falls by about one third to less than 20% and the number of poor decreases to less than 400 million. As a stark contrast, the \$2 a day poverty rate in Sub-Saharan Africa is expected to remain very high. Our moderate growth scenario predicts a poverty rate of about 66%, down from 69.9% in 2010, which at current population projections implies almost one billion poor in Sub-Saharan Africa alone. Even in the optimistic distribution-neutral growth scenario, we project a poverty rate of about 55% and more than 750 million poor. This is underlined by the analysis in the preceding section where we suggest that the mass of the consumption distribution is far to the left of the \$2 a day poverty line in 2010 for most of the subcontinent. Poverty alleviation in Sub-Saharan Africa remains the primary development challenge of the first half of the 21st century.

5 Concluding remarks and policy recommendations

The main contribution of this chapter is to forecast global poverty rates until 2030. To set the stage, we first highlighted that there is a fundamental uncertainty about the precise levels of extreme poverty. For the sake of comparison, we selected the standard poverty lines of \$1.25 and \$2 a day (in 2005 PPPs). We then discussed a robust set of global poverty and inequality trends. The global MDG of halving the 1990 extreme poverty level was reached in 2010 but this apparent progress hides substantial regional heterogeneity. Most of the global success was driven by rapid growth in China. Inequality among the citizens of developing countries and between developing countries was declining until 2005, while average within-country inequality was rising steadily until 2010.

The changing composition of global poverty has profound implications for the medium-term future. After 2010, fast growing East Asia will contribute less and less to global poverty reduction, while the share of the global poor residing in Sub-Saharan Africa and South Asia will continue to rise. All of our projections show that the global rate of poverty reduction at \$1.25 a day will slow down markedly between 2020 and 2025. None of our nine scenarios predicts a poverty rate near zero by 2030. This stands in stark contrast to earlier studies and the ‘3% by 2030’ target recently announced by the World Bank. The Bank’s target can only be reached if we make the unrealistic assumption of equally rapid growth in all developing countries. Once country-specific growth rates are used, even our most optimistic scenarios suggest a poverty rate of between 7.9% and 10.5%, depending on the evolution of inequality. At \$2 a day, the slowdown will occur much later and remarkable gains are possible if the post-2000 growth trends continue. An optimistic estimate suggests that the \$2 a day poverty rate may fall below 20% by 2030, implying one billion fewer poor people than in 2010.

We propose two new ‘twin targets’ on the basis of these findings. An aspirational but realistic benchmark for progress would be to *“reduce the proportion of the population living below \$1.25 to 8% by 2030 and reduce the proportion of the population living below*

²³However, our estimates suggest that this would require an exceptionally large slowdown. For poverty in East Asia to remain above 10% at \$2 a day, growth needs to be less than half of the 2000-2010 trend.

\$2 a day to 18% by 2030.” Both of these targets are firmly anchored in our optimistic pro-poor growth scenarios. The \$2 a day poverty line should receive more attention in the future to better track continued progress in East Asia and, later on, South Asia. Partly for the same reason, China has recently raised its own national poverty line to about \$1.80 a day.

These targets can be reached in a variety of ways but not only through a continuation of the current path. They will require either an additional acceleration of growth in poorer countries, or improvements in distribution. Reversing the trend of rising within-country inequalities would speed up the pace of poverty reduction and still ensure progress at more moderate growth rates. The returns to redistribution are increasingly high in East and South Asia, and remain relatively large in Latin America. However, in some regions growth takes precedent. Rapid poverty alleviation in Sub-Saharan Africa still requires a significant and sustained acceleration in consumption growth.

Bibliography

- Ahluwalia, M. S., N. G. Carter, and H. B. Chenery (1979). Growth and poverty in developing countries. *Journal of Development Economics* 6(3), 299–341.
- Bourguignon, F. (2003). The growth elasticity of poverty reduction: Explaining heterogeneity across countries and time periods. In T. S. Eicher and S. J. Turnovsky (Eds.), *Inequality and Growth: Theory and Policy Implications*, pp. 3–26. Cambridge, MA: MIT Press.
- Bresson, F. (2009). On the estimation of growth and inequality elasticities of poverty with grouped data. *Review of Income and Wealth* 55(2), 266–302.
- Chandy, L., N. Ledlie, and V. Penciakova (2013). The final countdown: Prospects for ending extreme poverty by 2030. Global Views Policy Paper 2013-04, Brookings Institution.
- Chen, S. and M. Ravallion (2001). How did the world’s poorest fare in the 1990s? *Review of Income and Wealth* 47(3), 283–300.
- Chen, S. and M. Ravallion (2010). The developing world is poorer than we thought, but no less successful in the fight against poverty. *Quarterly Journal of Economics* 125(4), 1577–1625.
- Chen, S. and M. Ravallion (2013). More relatively-poor people in a less absolutely-poor world. *Review of Income and Wealth* 59(1), 1–28.
- Chotikapanich, D., W. E. Griffiths, D. S. P. Rao, and V. Valencia (2012). Global income distributions and inequality, 1993 and 2000: Incorporating country-level inequality modeled with beta distributions. *Review of Economics and Statistics* 94(1), 52–73.
- Deaton, A. (2010). Price indexes, inequality, and the measurement of world poverty. *American Economic Review* 100(1), 1–34.
- Deaton, A. and O. Dupriez (2011). Purchasing power parity exchange rates for the global poor. *American Economic Journal: Applied Economics* 3(2), 137–166.
- Dhonde, S. and C. Minoui (2013). Global poverty estimates: A sensitivity analysis. *World Development* 44(C), 1–13.
- Easterly, W., M. Kremer, L. Pritchett, and L. H. Summers (1993). Good policy or good luck? *Journal of Monetary Economics* 32(3), 459–483.
- Eichengreen, B., D. Park, and K. Shin (2013, January). Growth slowdowns redux: New evidence on the middle-income trap. Working Paper 18673, National Bureau of Economic Research.
- Karver, J., C. Kenny, and A. Sumner (2012). MDGS 2.0: What goals, targets, and timeframe? *IDS Working Papers* 2012(398), 1–57.
- Klasen, S. (2009). Levels and trends in absolute poverty in the world: What we know and what we don’t. In E. Mack, M. Schramm, S. Klasen, and T. Pogge (Eds.), *Absolute poverty and global justice*, Chapter 1, pp. 21–36. London, UK: Ashgate.
- Klasen, S. and M. Misselhorn (2008). Determinants of the growth semi-elasticity of poverty reduction. Working paper, Ibero America Institute for Economic Research.
- Kraay, A. (2006). When is growth pro-poor? Evidence from a panel of countries. *Journal of Development Economics* 80(1), 198–227.
- Lakner, C. and B. Milanovic (2013, December). Global income distribution: From the fall of the Berlin Wall to the great recession. Policy Research Working Paper Series 6719, The World Bank.
- Lopez, H. and L. Serven (2006, January). A normal relationship? Poverty, growth, and inequality. Policy Research Working Paper Series 3814, The World Bank.
- Milanovic, B. (2012). Global inequality recalculated and updated: The effect of new PPP estimates on global inequality and 2005 estimates. *Journal of Economic Inequality* 10(1), 1–18.
- Pinkovskiy, M. and X. Sala-i-Martin (2009, October). Parametric estimations of the world distribution of income. NBER Working Papers 15433, National Bureau of Economic Research.
- Ravallion, M. (2010). The developing world’s bulging (but vulnerable) middle class. *World Development* 38(4), 445–454.
- Ravallion, M. (2013). How long will it take to lift one billion people out of poverty? *World Bank Research Observer* 28(2), 139–158.
- Ravallion, M., S. Chen, and P. Sangraula (2009). Dollar a day revisited. *World Bank Economic Review* 23(2), 163–184.
- Ravallion, M., G. Datt, and D. van de Walle (1991). Quantifying absolute poverty in the developing world. *Review of Income and Wealth* 37(4), 345–361.
- Reddy, S. G. and T. Pogge (2010). How to not count the poor. In S. Anand, P. Segal, and J. E. Stiglitz

- (Eds.), *Debates on the measurement of global poverty*, Chapter 2, pp. 42–85. New York, NY: Oxford University Press.
- Sala-i-Martin, X. (2006). The world distribution of income: Falling poverty and . . . convergence, period. *Quarterly Journal of Economics* 121(2), 351–397.
- Yoshida, N., H. Uematsu, and C. E. Sobrado (2014, January). Is extreme poverty going to end ? An analytical framework to evaluate progress in ending extreme poverty. Policy Research Working Paper Series 6740, The World Bank.
- Young, A. (2011, December). The gini coefficient for a mixture of ln-normal populations. Manuscript. The London School of Economics and Political Science, London, UK.

Appendix A: Additional tables and figures

Table A-1 – World Bank poverty estimates by region, 1981 to 2010 (selected years)

	<i>Year</i>				
	1981	1990	1999	2005	2010
<i>Panel (a) – Headcount ratio at \$1.25 a day (in percent)</i>					
East Asia and Pacific	77.18	56.24	35.58	17.11	12.48
China	84.02	60.18	35.63	16.25	11.62
Europe and Central Asia	1.91	1.91	3.79	1.33	0.66
Latin America and Caribbean	11.89	12.24	11.86	8.66	5.53
Middle East and North Africa	9.56	5.75	5.01	3.45	2.41
South Asia	61.14	53.81	45.11	39.43	31.03
India	59.83	51.31	45.62	40.82	32.67
Sub-Saharan Africa	51.45	56.53	57.89	52.31	48.47
Total	52.16	43.05	34.07	25.09	20.63
<i>Panel (b) – Poor population at \$1.25 a day (in millions)</i>					
East Asia and Pacific	1096.5	926.42	655.59	332.08	250.90
China	835.07	683.15	446.35	211.85	155.51
Europe and Central Asia	8.21	8.87	17.83	6.26	3.15
Latin America and Caribbean	43.33	53.43	60.10	47.60	32.29
Middle East and North Africa	16.48	12.96	13.64	10.47	7.98
South Asia	568.38	617.26	619.46	598.26	506.77
India	428.68	448.34	472.74	466.30	400.08
Sub-Saharan Africa	204.93	289.68	375.97	394.78	413.73
Total	1937.83	1908.45	1742.53	1389.2	1214.98
Total excl. China	1102.76	1225.30	1296.18	1177.35	1059.31

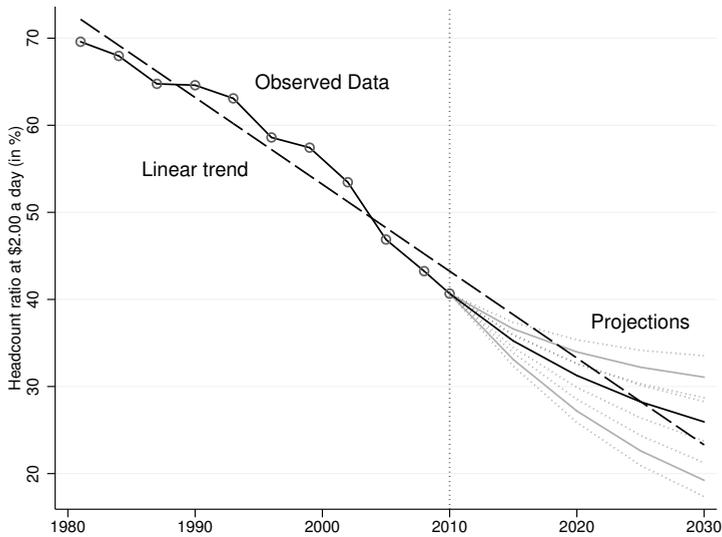
Notes: Based on *PovcalNet* and Chen and Ravallion (2010, 2013).

Table A-2 – Population-weighted regional PCE growth rates over various periods

	<i>Period</i>					
	2000 – 2010	1990 – 2010	1980 – 2010	1980 – 2000	1990 – 2000	
East Asia and Pacific	5.906 (0.813)	5.772 (0.653)	5.598 (0.725)	5.377 (0.677)	5.608 (0.508)	
Europe and Central Asia	6.085 (0.989)	2.755 (0.412)	2.558 (0.411)	-0.769 (0.916)	-1.225 (1.027)	
Latin America and Caribbean	2.444 (0.239)	2.219 (0.140)	1.445 (0.098)	0.677 (0.171)	1.931 (0.337)	
Middle East and North Africa	3.495 (0.443)	2.532 (0.440)	1.851 (0.293)	0.495 (0.545)	1.253 (0.648)	
South Asia	4.448 (0.489)	3.612 (0.388)	3.179 (0.351)	2.173 (0.284)	2.511 (0.294)	
Sub-Saharan Africa	2.382 (0.689)	1.419 (0.470)	0.698 (0.472)	-0.818 (0.540)	0.016 (0.688)	
Overall average	4.544 (0.152)	3.809 (0.132)	3.437 (0.114)	2.565 (0.161)	2.862 (0.225)	
N	123	123	123	122	122	
\bar{T}	10.99	20.64	27.16	16.30	9.730	
$N \times \bar{T}$	1352	2539	3341	1989	1187	

Notes: Authors' calculations. We use growth in per capita consumption from World Development Indicators or Penn World Table 7.1, depending on which series is longer. Cluster-robust standard errors are reported in parentheses.

Figure A-1 – Actual and projected poverty headcount ratios at \$2 a day, 1981–2030



Notes: Authors' calculations based on Chapter 5 and survey data from *PovcalNet*. The solid black line beyond 2010 refers to the moderate (distribution-neutral) growth scenario in Table 4, while the solid grey lines represent the distribution-neutral variants of the optimistic and pessimistic scenarios. The pro-poor and pro-rich variants are shown as grey dotted lines and are located above or below a solid line.

Chapter 7

Concluding remarks

The research presented in this dissertation asked two overarching questions:

- How do political institutions affect economic crises?
- How much does poverty react to changes in average income and distribution?

In the following, I briefly summarize the main findings from each chapter and outline how they together form an answer to these questions that (hopefully) exceeds the sum of its constituent parts.

The first part of this dissertation is devoted to the political economics of crises. An obvious question which arises right away is: why only study periods of negative growth and not levels or growth rates of GDP per capita directly? The starting point is an empirical observation. The previous literature has robustly linked institutional quality to *levels* of GDP per capita, but struggles to identify a role for political institutions in determining post-WWII *growth rates*. Similarly, studies of growth accelerations hardly find any evidence linking political factors to the start of an acceleration. In fact, often the converse is found. Democratization, for example, is often followed by periods of lower growth. This dissertation suggests that we have been looking in the wrong places. Growth accelerations occur relatively frequently and are caused by idiosyncratic macroeconomic factors that have little to do with the institutional environment. In contrast, crises are ubiquitous, often long and deep in the developing world, but comparatively rare, short and shallow in the developed world. This dissertation shows that it is not only the lack of rapid growth which creates the dividing line between catching up and falling behind, but also the failure to achieve sustained growth and to preserve earlier welfare gains. Weak institutions and social conflict play a significant role in how countries deal with negative growth shocks.

Chapter 2 is a first attempt to study research question one. It looks at stagnation episodes, defined as periods where GDP per capita is at a peak, then declines and then recovers to its previous level, so that there was no development. In studying the switching in and out of these episodes, it pays attention to two sub-questions: *i*) can we predict the onset of stagnation episodes with institutional and macroeconomic factors, and *ii*) do the variables which determine the onset of stagnation also affect the probability of continuation of a stagnation spell? The answers are mixed. On the one hand, we find no evidence that the quality of political institutions (measured by the Polity IV score) affects the onset or continuation of stagnation; that is, it does not seem to be linked to the incidence of stagnation. On the other hand, negative regime changes have the single

largest adverse effect on the probability of stagnation. This suggests that institutional shocks are indeed a problem for maintaining positive growth. Several macroeconomic factors are associated with the incidence of stagnation. Open economies, in terms of trade and finance, with an undervalued exchange rate seem to run a lower risk of experiencing a stagnation spell. Not so surprisingly, high inflation signals macroeconomic mismanagement and is associated with a higher probability of stagnation. Only trade openness has a differential impact. It reduces the chances of falling into stagnation but does not speed up the recovery. Chapter 2 also offers insights into what cannot be learned from panel studies of annual transitions in and out of an episode. Variables like institutional quality are slow moving, but the between-country variation will be wiped out by fixed effects or correlated random effects, leaving little within-country variation to work with.

Chapter 3 digs deeper. It relaxes several assumptions that were implicit in our study of stagnation spells. We no longer treat the decline and recovery phases as one, but only focus on the former. The underlying idea is that policy responses to declines depend on political institutions as well as the level of heterogeneity among ethno-political groups that have to agree on coordinated responses to an economic decline. In addition, we now focus only on large economic slumps to exclude business cycle fluctuations which are unlikely to be determined by political institutions. Much of the chapter is concerned with the econometric identification of economic slumps. We show that a restricted structural change approach coupled with a parametric bootstrap works well for finding and dating slumps. It turns out that severe downward volatility occurs often and has substantial welfare implications for low income countries. We then establish that stronger constraints on the executive shorten the duration of declines, while greater ethnic divisions prolong the duration. In fact, the positive impact of constraints on the executive on the duration of declines appears to depend on the level of ethno-linguistic heterogeneity; it is largest in very diverse societies. We also show that these two variables similarly affect the overall depth of declines and that their effects on depth run through the duration of the decline phase, as opposed to the average rate of decline. Macroeconomic factors still play a role, as in the previous chapter, but the effects of executive constraints and ethnic cleavages seem to operate independently of macroeconomic variables. Moreover, many structural features of modern economies are only likely to emerge together with strong institutions guaranteeing the security of transactions, among other things. An interesting stylized fact is that we observe broad-based institutional reforms following the onset of a slump. We interpret this finding both as evidence of the endogeneity of political institutions and as a sign that deep recessions offer windows of opportunity for institutional reforms.

Chapter 4 offers a theory of why we observe such long declines in heterogeneous countries with weak executive constraints. It models this situation as an infinite horizon cooperation game where political groups decide on whether to formulate a response to an exogenous shock. We show that group heterogeneity in a divided society and weak constraints on the political executive potentially lead to delayed cooperation during the decline phase of a slump. This is due to a commitment problem between winners and losers of the recovery process after a crisis. Uncertain post-recovery incomes and weak constraints on the executive create a ‘winner-take-all’ effect. Groups anticipate this uncertainty and may find it beneficial to hold out. We derive four theoretical results based on this setup. First, delay can occur in equilibrium. Second, stronger constraints on the executive shorten the time to recovery. Third, decreasing political concentration can make delay more likely under certain conditions. Fourth, an increase in the number

of groups makes delay more likely. The empirics are consistent with the proposed theory. Using detailed data on ethnic power configurations and the data set of slumps from Chapter 3, we show that heterogeneity is indeed a problem when the political executive is relatively unconstrained. Our empirics also suggest that this problem primarily arises as a consequence of an increase in the number of equally powerful groups, rather than political constellations where one or few groups wield all power. In fact, greater group asymmetries work in the opposite direction. This chapter has an important policy implication: ethnic diversity is not necessarily a problem. Political institutions can be designed to contain the adversarial element of ethnic heterogeneity and may help to ensure that recent welfare gains are not lost in the next crisis.

Taken together, the three essays offer an answer to the first research question. Chapter 2 shows that negative regime transitions raise the probability of a crisis occurring but otherwise political institutions do not seem to be linked to the incidence of stagnation spells. Chapter 3 then demonstrates a robust association of political institutions and ethnic diversity with the duration of economic slumps. Their effects on the expected time until the recovery starts are economically very large. Chapter 4 suggests that these long declines may occur due to a commitment problem among winners and losers of the crisis; there are no enforceable contracts where groups return what they expropriated from a weaker group after the recovery. Guaranteeing that this does not happen is precisely one of the roles played by strong political institutions. Given this body of empirical and theoretical evidence it seems safe to conclude that political institutions do indeed play a significant role during economic crises. Many important issues have not been addressed in this part of the dissertation. An integrated empirical comparison of the relative welfare effects of accelerations and growth collapses would be an interesting extension of this line of research. In addition, considerably more work is needed to arrive at a unified theory of the institutional causes of uneven growth.

The second part of this dissertation shifts the focus away from economic growth towards the poverty of people. It seeks to understand how different growth experiences translate into poverty reduction and how inequality moderates this relationship. This part of the dissertation first examines how changes in average income (consumption) and changes in inequality have historically translated into poverty reduction. At the micro-level, there is little mystery in this question. All we need to do is to estimate the Lorenz curve, take point derivatives with respect to mean income or a measure of inequality and we learn how much of a change in either of these two variables translates into poverty reduction. At the cross-country level, we typically only have average incomes and the Gini coefficient. Chapter 5 shows that with just these two data points, we can very closely approximate the underlying elasticities and semi-elasticities. Nevertheless, finding a particular average value for the income elasticity of poverty in the past does not imply that poverty will respond the same in the future. In Chapter 6, we turn to the question of how this process is likely to continue if we look forward to 2030.

Chapter 5 outlines a new econometric method for estimating the impacts of changes in average incomes and changes in inequality. We propose a fractional response approach which explicitly takes the bounded nature of the poverty headcount ratio into account; that is, the fact that it cannot exceed the unit interval. This approach allows us to derive estimates of the income (and inequality) elasticities and semi-elasticities that have desirable properties, such as being bounded at zero when incomes are so small that no one is lifted out of poverty due to one percent income growth. We then extend this approach to deal with unobserved differences in the underlying household surveys, measurement

error in income (consumption) and uneven spacing of surveys. We first show that the income elasticity of poverty from 1981 to 2010 was about two and the inequality elasticity was about one and a half. The chapter offers two additional findings: *i*) elasticities are not very informative and underrepresent the contribution of inequality, and *ii*) we learn very little from these historical averages as they apply to no country in particular. Instead, we argue that both the academic literature and policy makers should focus on semi-elasticities. Semi-elasticities measure how much of a one percent change in average income (inequality) translates into a change in the fraction of the total population that is poor (a percentage point change in the poverty rate). We illustrate that our method can trace out the entire distribution of these quantities. Semi-elasticities vary a lot over the development process. They express that the poverty reduction potential of a country is increasing until the mass of the distribution is right in front of the poverty line and then decreases to zero as fewer and fewer poor people are left. We also briefly explore how other variables, such as political institutions, can influence poverty reduction and show that they only do so indirectly through average incomes and inequality. In theory and practice, elasticities and semi-elasticities are only non-linear functions of initial levels and changes in both average incomes and inequality.

Chapter 6 turns to the future. It illustrates a fundamental insight from Chapter 5 with an important policy question. If the income and inequality semi-elasticities of poverty are falling after some point as countries grow richer, then the income level of poor countries near the global poverty line matters for the speed of poverty reduction. We use precisely this fact to show that the pace of poverty reduction at the \$1.25 a day poverty line is likely to slow down. Fast growing East Asia will contribute less and less to global poverty reduction in the future, while more and more of the world's poor will be concentrated in Sub-Saharan Africa and South Asia. While several countries on these two subcontinents have experienced substantially faster GDP per capita growth rates since the early 2000s, achieving the World Bank's new global goal of 3% poverty by 2030 at the \$1.25 a day poverty line requires unrealistic assumptions regarding income and consumption growth. All of our projections show that the pace of poverty reduction at \$1.25 a day will slow down. The global poverty rate only reaches about 8-9% in 2030 under very optimistic assumptions. In contrast, the situation at the \$2 a day poverty line is more reminiscent of the recent past. East Asia still has plenty of people living on incomes between \$1.25 and \$2 a day, so that the historical speed of poverty reduction of one percentage point per annum can be maintained. Our optimistic scenarios suggest that the \$2 a day poverty rate may fall below 20% by 2030 down from about 40% in 2010, implying at least one billion fewer poor people than in 2010. On this basis, we suggest two global targets: 1) reducing the proportion of the population living below \$1.25 to 8% by 2030, and 2) reducing the proportion of the population living below \$2 a day to 18% by 2030.

In a nutshell, the answer to research question two is: it depends. Chapter 5 shows that estimates of the average response of poverty to changes in incomes and distribution are not very informative. Both elasticities and semi-elasticities change systematically over the development process. Elasticities continually increase in absolute value, whereas semi-elasticities are first increasing and then decreasing in average income. Our framework can be used to predict the point (semi-)elasticity using only a country's average income level and Gini coefficient. Chapter 6 adds that this functional form also implies that a slowdown in the pace of poverty reduction is very likely. Although poverty reduction in the past has mostly been driven by income growth, there is nothing about this relationship that suggests redistribution could not play a larger role in the future. We also established

that institutions only affect poverty indirectly, through the levels of (and changes in) average incomes and inequality. Yet, our knowledge of the transmission channels is still very incomplete. The links between political institutions and inequality in particular are interesting avenues for future research.

Addendum on valorization

In accordance with article 23.5 of the “Regulation governing the attainment of doctoral degrees at Maastricht University” decreed by resolution of the Board of Deans, dated 3 July 2013, an addendum must be added about valorization.

Testament to the social relevance of this research is the fact that this doctoral dissertation has been funded by the Agence Française de Développement (AFD, the French development agency), within a larger multi-year research project on ‘Institutions and Long-term Growth’. The project originated as an attempt to understand, both theoretically and empirically, the conceptual and econometric relationships between institutions and long-term growth. The French development agency was particularly interested in better understanding the institutional determinants of long-term economic, social, and political trajectories.

The five studies presented in this doctoral dissertation contribute to advance our understanding of each of these three themes. Their social and economic implications are of considerable importance for the welfare of nations and the welfare of people. However, not all of these implications can be easily translated into policy prescriptions.

The first part of the thesis, on the political economics of crises, is concerned with the fundamental question of how political institutions affect economic crises. The relevance of such a question can hardly be understated at a time when much of the developed world is still dealing with the consequences of the Great Depression and parts of the developing world are experiencing economic uncertainty due to political upheaval, such as the Arab spring or other political crises. Development aid today, too, has shifted away from purely economic investments towards supporting institutions, improved governance, capability building and civil society. Aid investments in the productive sector (agriculture, industry, tourism, banking, etc.), for example, have decreased from 36.6% of total bilateral aid in 1967 to 13.4% 2012, while the relative share of aid supporting the social infrastructure (education, health, government, etc.) has increased by 23.3 percentage points over the same period (from 8.3% to 31.6%).¹

The overarching insight from the first part of the dissertation is that well-developed checks and balances (constraints) on the executive help to avert long and deep crises, even in the face of ethnic and political heterogeneity. Apart from this general finding, this part contains several insights for policy makers that are worth highlighting:

- Negative regime changes have the single largest effect on the probability of falling into a stagnation spell. Contrary to the widespread notion of the benefits of a “strong ruler”, institutional shocks involving an increase in executive power and reduction of political competition are not necessarily beneficial to growth, even in the short run.

¹Table 14.6 from Adam Szirmai, www.dynamicsofdevelopment.com, 2015.

- Open economies with an undervalued exchange rate seem to run a lower risk of experiencing a stagnation episode. This finding has direct implications for policy makers in central banks and those in the government engaged in trade policy.
- Severe downward volatility occurs often and has significant welfare implications for low income countries. Stronger constraints on the executive shorten the duration of declines, while greater ethnic divisions prolong the duration. It is important for the governments of developing countries to recognize where they are on this continuum, so as to appropriately manage economic crises.
- Ethnic diversity is not necessarily a problem. Political institutions can be designed to contain the adversarial element of ethnic heterogeneity, namely institutions that constrain the power of ruling elites to expropriate others. While the precise nature of these institutions requires further research, this finding is important for countries undergoing, or considering, constitutional reforms and complements a larger literature on the economic effects of political constitutions.

The second part of the thesis, on the speed of poverty reduction, is concerned with how poverty rates react to changes in average income and changes in income inequality. Even though past progress in the quest to reduce global poverty has been substantial – the proportion of the global population living on less than \$1.25 a day declined from about 52% in 1981 to about 20% in 2010 – there are substantial challenges ahead. Research in this field is of obvious social importance, as entire agencies such as the World Bank have been designed around the sole mandate of promoting long-term economic development and poverty reduction. Recently, the World Bank even ambitiously redefined its own mission as “ending extreme poverty within a generation and promoting shared prosperity”.

A principal result from the second part of the dissertation is that the past will not be a good guide for the future, at least when it comes to the pace of poverty reduction. This occurs for three reasons which are detailed below. Each of these three points has significant implications for policy makers and the larger development community.

- The geography of world poverty is changing rapidly. While China and India have contributed significantly to poverty reduction at the \$1.25 a day poverty line in the past, their contribution will rapidly decrease in the coming decades. The relative share of the global poor residing in Sub-Saharan Africa and South Asia will rise. Even though per capita consumption growth has accelerated in Sub-Saharan Africa since about 2000, it has lagged far behind that of other developing countries, trailing the growth rate of the rest by about 2.5 percentage points per annum. This highlights the importance of developing sustained sources of economic growth on the subcontinent, such as further improving agricultural productivity, so as to free up workers for the industrial or service sector.
- Within-country inequality (as measured by the Gini coefficient) has been more or less constant in the developing world. The research presented here shows that reductions in inequality can have potentially large effects on the poverty headcount ratio, especially when viewed in terms of percentage point changes in poverty rates. Nevertheless, we have seen few examples of countries that have reduced poverty and inequality simultaneously, while there are about as many examples of countries, where the distribution was becoming more unequal over time. This finding stands

in sharp contrast to the development community's rhetoric of 'pro-poor growth', or, more recently, 'shared prosperity'. The evidence presented here suggests that both have been the exception rather than the rule.

- Extreme poverty is not likely to end by 2030. The World Bank's target can only be reached if we make the unrealistic assumption of equally rapid growth in all developing countries. Instead, the projections show that the global rate of poverty reduction at \$1.25 a day will slow down markedly. Moreover, in absolute terms, the number of poor people in Sub-Saharan Africa is even likely to rise due to unabated population growth. On the other hand, at \$2 a day, remarkable gains are possible if the post-2000 growth trends continue. An optimistic estimate suggests that the \$2 a day poverty rate may fall below 20% by 2030 (down from about 40% in 2010), implying one billion fewer poor people than in 2010. Improvements in distribution are much needed to counteract unfavorable trends and ensure continued progress.

Chapter 6 of this dissertation then distilled these findings into a recommendation to the UN's high-level panel on the Sustainable Development Goals (SDGs). The chapter proposes two new goals, namely to "reduce the proportion of the population living below \$1.25 to 8% by 2030 and reduce the proportion of the population living below \$2 a day to 18% by 2030."

Last but not least, the work on this dissertation resulted in several ancillary outputs, in particular software components for estimating econometric models that have previously not been implemented. Some programs are add-ons to the commercial software Stata®, but the source code is publicly available at the author's website and subject to open source licensing. Other programs have been written for the open source software R. Additional programs and data sets will be made available once the results are published (either on the author's website or as part of the open data initiatives of academic journals). Making the code and data public serves to spread the use of these techniques and uphold the standards of reproducible research.

A core target group of this research consists of policy makers in international agencies (UN, World Bank, IMF, European Commission, regional investment banks, etc.). Other target groups of this research are national aid agencies (e.g. DFID, GIZ), central banks, national governments and statistical offices.

The French development agency played a key role in disseminating the work done within this project to other policy institutions. Several works have been reprinted as AFD working papers and/or have been summarized in policy briefs. Following a political economy workshop in Dresden, Chapter 3 has been released as CESifo working paper (one of the largest economic research networks in the world with a wide readership). All research outputs were presented at annual workshops at the AFD, several academic conferences, and policy institutions, such as the World Bank and the Center for Global Development.

The policy-oriented results from Chapter 6 are part of a joint research project funded by the Agence Française de Développement (AFD), the Institute for Development Studies (IDS) and the Japanese Development Cooperation (JICA). The chapter has been published in a book called "Growth is dead, long live growth: The quality of economic growth and why it matters" edited by Lawrence Haddad (IDS), Hiroshi Kato (JICA) and Nicolas Meisel (AFD). The results of this research project have been publicly presented to policy makers on January 26th 2015 in a high-level panel together with the respective

directors of these institutions. The project was subsequently covered in the international media, such as the website of the Guardian.

Furthermore, an extended version of the literature review that informed the discussion in the introduction was commissioned by the UNICEF Innocenti Research Centre in Florence (Italy) and presented there in 2011. A follow-up project to the second part of this dissertation on “Inequality in Sub-Saharan Africa” is being funded by DG EuropeAid of the European Commission.

The research also targets the wider academic community as well as the readership of journals and blogs (such as the UNU-MERIT blog). The output of two papers was discussed in a short interview, available on YouTube, as part of the dissemination efforts.

All research in this dissertation is to be published in peer-reviewed and internationally acclaimed journals. As a first step in this direction, Chapter 2 is forthcoming in *Macroeconomic Dynamics* and the remaining chapters are currently under review.

Given the nature of this dissertation, it contains no proposals for commercial products or commercially exploitable activities, such as patents or services, and no innovation regarding existing products, services, processes, activities and commercial activities. Therefore, the dissertation contains no proposals for exploiting market opportunities in the future.

Samenvatting

Het voorliggende proefschrift bestudeert verschillende aspecten van de ontwikkelingspuzzel, in het bijzonder de empirie en de politieke economie van groei en ongelijkheid. Het laat zich in twee delen opsplitsen. Deel I bevat drie essays over de theorie en empirie van economische crisissen en politieke instellingen. Deel II bestaat uit twee empirische analyses van groei, verdeling en armoede. In deze samenvatting worden de aanpak en bevindingen van elk hoofdstuk in het kort geschetst.

De overkoepelende onderzoeksvraag voor de hoofdstukken in deel I is: Welke invloed hebben politieke instellingen op het ontstaan en de duur van economische crisissen? Een eerste antwoord vindt men in Hoofdstuk 2, vertrekkende van de eenvoudigste definitie van economische stagnatie. Een episode van stagnatie begint met een krimp van het BBP per hoofd van de bevolking in een tijd waarin het hoger was dan ooit tevoren, en eindigt wanneer het weer het initiële niveau bereikt of overtreft. Wij zoeken dergelijke episodes van economische stagnatie op en analyseren ze in een wereldwijd panel van landen. We behandelen in het bijzonder twee vragen die betrekking hebben op de instabiliteit van de groei. Ten eerste, we onderzoeken of institutionele kenmerken en politieke schokken merkbare effecten hebben op de frequentie van stagnatie, en hoe dergelijke effecten zich verhouden tot standaard macro-economische verklaringen. Wat betreft politieke variabelen gaat het om regimewijzigingen, het plotse aftreden van leiders, oorlogen en burgeroorlogen. Wat betreft macro-economische variabelen gaat het voornamelijk om inflatie en de openheid van de grenzen voor financiële en handelsstromen. Ten tweede, we gaan na of de opgenomen variabelen dezelfde invloed hebben op de frequentie en op de duur van stagnatie. Met andere woorden, we onderzoeken of de variabelen die de kans op stagnatie verhogen of verlagen dezelfde zijn als de variabelen die de kans om te blijven stagneren verhogen of verlagen.

De antwoorden zijn meervoudig. Enerzijds vinden we geen aanwijzing dat de kwaliteit van politieke instellingen, zoals gemeten door de “Polity IV”-score, van invloed is op het ontstaan of de voortzetting van stagnatie. Anderzijds is het wel duidelijk dat negatieve regimewijzigingen de kans op stagnatie verhogen. Institutionele schokken kunnen dus wel degelijk in de weg staan van positieve groei. Verder zijn verschillende macrofactoren van belang. Economieën die heel open zijn voor financiële en handelsstromen of waarvan de munt ondergewaardeerd is lijken minder risico op stagnatie te lopen. Hoge inflatie is (niet verrassend) een signaal van macro-economisch wanbeleid en een voorbode van stagnatie. Openheid voor handel is de enige verklarende factor met duidelijk verschillende effecten op het optreden en de continuering van stagnatie: open grenzen verminderen de frequentie van stagnatie, maar versnellen het herstel niet.

Hoofdstuk 3 probeert een andere benadering. Daar worden economische crisissen gekarakteriseerd door een specifiek type trendbreuk of verschuiving in het groeiregime. De motivering is tweeledig. We willen cyclische periodes van laagconjunctuur uitsluiten en in plaats daarvan focussen op grote, onverwachte en negatieve afwijkingen van een

eerder positieve trend in het BBP per hoofd van de bevolking; alleen de bedoelde episodes noemen we “crisissen”. Bovendien willen we de aandacht nu ook beperken tot de neergaande fase van de crisis, laat ons zeggen de “inzinking”. Theoretisch zowel als empirisch kan de dynamiek van achteruitgang erg verschillen van de dynamiek van herstel, vandaar de noodzaak om de neergaande fase af te zonderen. We gaan nader in op drie specifieke onderzoeksvragen. Ten eerste, hoe kunnen we economische crisissen empirisch identificeren? Ten tweede, zijn er aanwijzingen van institutionele veranderingen in of rond de crisistijd? Ten derde, als het eenmaal crisis is, zal de duur van de inzinking dan afhangen van het type en de kwaliteit van de instellingen?

Een groot deel van het hoofdstuk wordt gewijd aan de econometrische identificatie van economische crisissen. We inventariseren de structurele trendbreuken die aan specifieke vereisten voldoen en laten zien dat deze aanpak in combinatie met een parametrische bootstrap effectief in staat is om crisissen te vinden en te dateren. We stellen vast dat neerwaartse volatiliteit vaak voorkomt en aanzienlijke gevolgen heeft voor het welzijn van lage-inkomenslanden. Vervolgens blijkt dat beperkingen opgelegd aan de uitvoerende macht de duur van inzinkingen verkort, terwijl etnische verdeeldheid hun duur verlengt. Daarenboven lijkt het gunstige effect van beperkingen opgelegd aan de uitvoerende macht sterk af te hangen van het niveau van de etnische verdeeldheid; de weerslag is het grootst in de meest heterogene samenlevingen. We tonen vervolgens aan dat machtsbeperking en etnische verdeeldheid een soortgelijke invloed hebben op de uiteindelijke diepte van inzinkingen, een invloed die uitgeoefend wordt via de duur en niet het tempo van de daling. Macro-economische factoren spelen eveneens een rol, net zoals in het vorige hoofdstuk, maar de effecten van beperkingen van de uitvoerende macht en etnische scheidingslijnen blijken onafhankelijk van macro-economische variabelen te werken. Bovendien vinden we dat bepaalde structurele kenmerken van moderne economieën enkel en alleen ontstaan in de aanwezigheid van sterke instellingen die (onder andere) de veiligheid van transacties kunnen garanderen. Een interessante vaststelling hierbij is dat we vaak brede institutionele hervormingen waarnemen kort na het begin van een inzinking. We interpreteren deze bevinding niet alleen als een mogelijk teken van de endogeniteit van politieke instellingen, maar ook als een aanwijzing dat zware crisissen kansen bieden voor diepgaande institutionele hervormingen.

Hoofdstuk 4 begint met een politiek-economische theorie van uitgesteld herstel dat bedoeld is om de empirische bevindingen van hoofdstuk 3 te verklaren, en vervolgt met een empirische toetsing van die theorie. De focus ligt nu op de machtsstrijd tussen politieke groepen omtrent het te voeren beleid tijdens een inzinking of crisis. We modelleren de situatie als een coöperatief spel met oneindige horizon, waarbij partijen onderhandelen over de mogelijke beleidsreactie op een exogene schok. We laten zien dat verdeeldheid in de samenleving en een uitvoerende macht zonder beperkingen wellicht leiden tot uitstel van samenwerking tijdens een economische neergangsfase. De verklaring is het gebrek aan een afdwingbare verbintenis tussen de uiteindelijke winnaars en verliezers van het verwachte herstel. Onzekerheid over de inkomensverdeling na het herstel en onbeperkte van de uitvoerende macht leiden tot een ‘winner-take-all’ effect. Afzonderlijke groepen anticiperen op deze onzekerheid en kunnen er voordeel uit halen om samenwerking lange tijd uit te stellen.

Uit deze analyse volgen vier theoretische resultaten. Ten eerste, uitstel van samenwerking komt vaak voor in een evenwicht. Ten tweede, betrouwbare beperkingen van de uitvoerende macht kunnen het herstel versnellen. Ten derde, hoe dunner de politieke macht gespreid is, hoe hoger de kans op uitstel van samenwerking. Ten vierde,

hoe groter het aantal politieke groepen waarin de maatschappij verdeeld is, hoe hoger de kans op uitstel. De empirische bevindingen zijn consistent met de voorgestelde theorie. We beschikken over gedetailleerde gegevens omtrent de etnische-linguïstische samenstelling van de bevolking en de machtsconfiguraties in de meeste landen van de wereld. Door deze te combineren met de crisisdata uit hoofdstuk 3 vinden we dat heterogeniteit inderdaad een probleem vormt in landen waar de uitvoerende macht weinig beperkingen kent. Het probleem lijkt voornamelijk voor te komen bij een relatief hoog aantal concurrerende machtsgroepen, eerder dan in politieke constellaties waar één of twee groepen dominant zijn, of waar er een sterke asymmetrie heerst tussen machtsgroepen. Hieruit volgt een belangrijke beleidsimplicatie: etnische diversiteit hoeft niet noodzakelijk problematisch te zijn. Politieke instellingen kunnen zodanig worden ontworpen dat spanningen tussen etnische groepen beheerst blijven en dat welzijnsaanwinsten niet meteen verloren gaan in een volgende crisis.

Samen genomen bieden de drie eerste essays een antwoord op de eerste onderzoeksvraag. Hoofdstuk 2 laat zien dat negatieve regimewijzigingen de kans op een crisis verhogen, maar voor het overige lijken politieke instellingen weinig invloed te hebben op de frequentie van crisissen. Hoofdstuk 3 laat echter zien dat de duur van economische crisissen wel afhangt van politieke instellingen, etnische heterogeniteit en machtsstructuren. In landen met zwakke instellingen en een grote verdeeldheid laat het herstel veel langer op zich wachten, met erg ingrijpende economische gevolgen. Hoofdstuk 4 suggereert dat langdurige inzinkingen het gevolg kunnen zijn van een gebrek aan verbintenis tussen toekomstige winnaars en verliezers; er zijn geen afdwingbare contracten waarbij machtsgroepen na het herstel moeten herverdelen wat afgenomen is van zwakkere groepen tijdens de crisis. Dat soort verbintenissen garanderen is precies een van de positieve rollen die degelijke politieke instellingen kunnen spelen. In het licht van onze theoretische en empirische resultaten kunnen wij besluiten dat politieke instellingen inderdaad een belangrijke rol te spelen hebben tijdens en rond economische crisissen.

Het tweede deel van dit proefschrift verschuift de aandacht van economische groei naar armoede. De centrale onderzoeksvraag in dit deel is: Hoeveel verandert het percentage mensen dat onder de armoedegrens leeft wanneer het gemiddelde inkomen of de inkomensverdeling wijzigen? We onderzoeken deze kwestie op twee complementaire manieren: i) welke invloed hebben inkomens- of consumptiegroei en herverdeling op de armoedecijfers gehad in het recente verleden (tussen 1981 en 2010); en ii) welke ontwikkelingen kan men verwachten of hopen in de nabije toekomst? In het bijzonder, wat zijn zinvolle armoededoelstellingen voor de komende decennia?

Hoofdstuk 5 heropent het debat over de doelmatigheid van armoedebestrijding door groei versus herverdeling. Eerdere bijdragen in de ontwikkelingseconomie hebben reeds belangrijke inzichten opgeleverd. Zo weten we bijvoorbeeld goed dat armoede gebonden is aan inkomensniveau en inkomensverdeling door een mathematische relatie, en dat de initiële niveaus van inkomen en van ongelijkheid parameters zijn in deze relatie. We stellen echter vast dat de literatuur tot nog toe de theoretische eigenschappen van armoederatio's onvolledig heeft uitgebaat. Het percentage mensen dat onder de armoedegrens leeft is een fractie met een beperkt bereik tussen 0 en 1, wat in de Engelse terminologie een "fractionele respons" heet. Het is goed mogelijk hoewel ongebruikelijk om deze inherente begrensdheid in een analytisch model in te bouwen, en daaruit elasticiteiten en semi-elasticiteiten van de armoede ten opzichte van het inkomensniveau en de inkomensongelijkheid af te leiden. Dit stelt ons dan weer in staat om de verdeling van inkomen of consumptie in de nabijheid van de armoedegrens zeer nauwkeurig te

benaderen, zonder dat specifieke aannames over de vorm van de inkomensverdeling nodig zijn. Wij verkrijgen dan ook schattingen van de elasticiteiten en semi-elasticiteiten met de wenselijke eigenschappen. Zo zijn semi-elasticiteiten laag bij initiële inkomens ver onder (dan wel boven) de armoedegrens, aangezien plausibele wijzigingen in het niveau of de verdeling van de inkomens niemand boven de armoedegrens zullen tillen (dan wel eronder zullen brengen). We breiden dit model uit om rekening te houden met verschillende storende omstandigheden, zoals onbenoemde verschillen in de onderliggende gezinsenquête, meetfouten in inkomens- of consumptiegegevens, en de onregelmatige timing van de enquêtes.

Het uitgebreide model leidt tot de volgende vaststellingen. De inkomenselasticiteit van armoede was van 1981 tot 2010 door de band genomen ongeveer twee, en de ongelijkheidselasticiteit ongeveer anderhalf. Hierbij horen twee kanttekeningen te worden gemaakt: i) elasticiteiten (hoewel veelvuldig gebruikt) zijn in dit verband niet erg informatief en onderschatten de bijdrage van ongelijkheid; en ii) historische gemiddelden leren ons erg weinig over de historische ontwikkeling en lokale omstandigheden. In plaats daarvan zouden zowel onderzoekers als beleidsmakers zich moeten baseren op geactualiseerde semi-elasticiteiten. Semi-elasticiteiten meten hoe een stijging van één procent in het gemiddeld inkomen of in de ongelijkheid zich vertaalt in procentpunten meer of minder van het armoedepercentage. We illustreren hoe onze methode de volledige evolutie van deze kwantiteiten kan traceren. Semi-elasticiteiten variëren aanzienlijk tijdens het ontwikkelingsproces. Men kan eraan zien dat het potentieel voor armoedebestrijding in een land toeneemt zolang de piek van de inkomensverdeling zich onder de armoedegrens bevindt. Eenmaal de piek de armoedegrens voorbij is neemt het potentieel voor armoedebestrijding geleidelijk af, omdat het aandeel armen steeds kleiner wordt.

Hoofdstuk 6 gebruikt hetzelfde empirische kader als hoofdstuk 5 maar is meer beleidsgericht. We gaan er de haalbaarheid na van de recent door de Wereldbank aangekondigde doelstelling om het armoedepercentage in de wereld terug te dringen, in termen van de armoedegrens van \$1,25 per dag, van ongeveer 20% in 2010 tot 3% in 2030. We herinneren eerst de oorsprong van de “één-dollar-per-dag” armoedegrens (tegenwoordig opgehoogd tot 1,25 dollar in verband met prijsinflatie). We voeren dan een gevoeligheidsanalyse van deze grens uit, en beschrijven de historische trends van armoede en ongelijkheid in de wereld. Dit dient twee doelen. Ten eerste willen we duidelijk maken dat de extreme armoedegrens van \$1,25 per dag niet absoluut vast staat, maar behept is met een aanzienlijke mate van onzekerheid. Veel andere grenzen zouden minstens even zinvol zijn. Ten tweede willen we benadrukken dat de landen die historisch gezien het meest hebben bijgedragen tot de vermindering van de wereldarmoedecijfers niet dezelfde zijn als de landen die dit in de toekomst zullen doen; zij zijn immers in de laatste twee decennia aan een buitengewoon snel tempo gegroeid, en hebben hun voorraad extreem armen grotendeels “uitgeput”. De geografie van de extreme armoede in de wereld is ingrijpend veranderd en dit zal zijn weerslag hebben op het tempo van de armoedebestrijding in de toekomst.

De lessen uit Hoofdstuk 5 laten een nieuw licht schijnen op de belangrijke beleidsvraag van Hoofdstuk 6. We hebben gezien dat de semi-elasticiteiten van armoede ten opzichte van inkomensniveau en ongelijkheid gaan dalen naarmate landen rijk(er) worden. Het inkomensniveau van arme landen *in de buurt van de armoedegrens* is het meest bepalend voor de snelheid waarmee de armoede verder teruggedrongen kan worden. Dit feit leidt ertoe dat het tempo waarmee de extreme armoede in de wereld daalt in de nabije

toekomst onvermijdelijk zal vertragen. De snel groeiende economieën van Oost-Azië zullen geleidelijk minder bijdragen tot de daling van de traditionele armoedemetingen, en een stijgend aandeel van de armen in de wereld zal geconcentreerd zijn in Zwart Afrika en Zuid-Azië. Het is waar dat verschillende landen op deze twee subcontinenten een aanzienlijk versnelde groei van het BBP per hoofd van de bevolking hebben gekend sinds de vroege jaren 2000. Toch stelt de nieuwe doelstelling van de Wereldbank, om tegen 2030 het aantal arme mensen onder de \$1,25-per-dag grens te drukken tot 3%, onrealistische vereisten aan de inkomens- en consumptiegroei. Uit alle scenario's voor de toekomst die wij onderzocht hebben blijkt dat het tempo van de armoededaling bij de gebruikelijke \$1,25 grens zal vertragen. Het wereldwijde aandeel van extreem armen zal onder de meest optimistische veronderstellingen nog steeds 8 à 9% bereiken in 2030. Echter, indien we de armoedegrens op \$2 per dag zouden plaatsen, dan krijgen we een ander beeld dat veel nauwer aansluit met het recente verleden. In Oost-Azië leven nog steeds heel veel mensen van inkomens tussen \$1,25 en \$2 per dag, zodat een daling met één procentpunt per jaar van het aantal mensen die van \$2 moeten leven nog decennialang binnen bereik ligt. Onze optimistische scenario's suggereren dat de armoede onder \$2 per dag terug zou kunnen vallen van ongeveer 40% in 2010 tot 20% tegen 2030. Dit betekent minstens een miljard minder "arme" mensen in 2030 dan in 2010. Op basis hiervan stellen we twee zinvolle globale doelstellingen voorop: 1) de vermindering van het aantal armen die leven van hoogstens \$1,25 per dag tot 8% tegen 2030; en 2) de vermindering van het aantal armen die leven van hoogstens \$2 per dag tot 18% tegen 2030.

Het antwoord op onderzoeksvraag 2 luidt dus: het hangt ervan af. Uit Hoofdstuk 5 blijkt dat schattingen van de gemiddelde respons van de armoedecijfers op stijgende inkomens en inkomensherverdelingen niet erg informatief zijn. Elasticiteiten en semi-elasticiteiten volgen een systematische evolutie gedurende het ontwikkelingsproces. Als functies van het gemiddelde inkomen stijgen elasticiteiten monotoon in absolute waarde, terwijl semi-elasticiteiten eerst stijgen en dan dalen. Ons model houdt rekening met de ingebouwde functionele beperkingen en laat toe beide evoluties precies te voorspellen voor ieder land, gegeven het niveau van het gemiddelde inkomen en de Gini-coëfficiënt. Uit Hoofdstuk 6 blijkt dat deze voorspellingen een vertraging in het tempo van de armoededaling impliceren in de nabije toekomst. Wellicht is het tijd om de traditionele armoedegrens en andere uitgangspunten van de armoedebestrijding te herzien. Hoewel armoedebestrijding in het verleden voornamelijk gedreven werd door inkomensgroei, is er niets in de relatie dat herverdeling belet een grotere rol te gaan spelen in de toekomst. We hebben tevens kunnen vaststellen dat de invloed van de instellingen op de armoede noodzakelijkerwijs loopt via gemiddelde inkomensniveaus en de ongelijkheid. Onze kennis van de betreffende transmissiekanalen blijft echter erg onvolledig. De verhouding tussen politieke instellingen en ongelijkheid in het bijzonder is een interessant thema voor toekomstig onderzoek.

About the author

Richard holds an MSc in Public Policy and Human Development (MPP) from Maastricht University, where he graduated *cum laude* (highest distinction) and received the university's 'top 3% award'. He has a bachelor's degree in International Business from HTW Berlin, Germany, and spent part of his undergraduate studies at the University of Technology in Sydney, Australia, and at the California State University Long Beach, USA (with a Fulbright scholarship). He has been consulting for the Agence Française de Développement (AFD), the European Commission, and UNICEF, as well as worked at the Limburg Institute of Business and Economic Research (LIBER), the International Labour Organization (ILO) and the Embassy of Canada in Berlin. At the graduate school, Richard lectured in the MPP program and taught several courses in econometrics. Most of the research in this dissertation was funded by the AFD as part of a multi-year research project on 'Institutions and Long-term Growth'. Richard joined the PhD program at Maastricht University in 2010 and is now an Assistant Professor (Akademischer Rat) at the Chair of Macroeconomics of Leibniz University Hannover.

MGSOG Dissertation Series

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 Elite: 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)

Agnieszka 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

Conflicting 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. Muñiz 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)