



---

# THE GEOGRAPHIC DIMENSIONS OF GROWTH AND DEVELOPMENT

---

SAMYUKTA BHUPATIRAJU

© Copyright Samyukta Bhupatiraju, Maastricht 2014

ISBN 978 94 6159 386 3

Datawyse / Universitaire Pers Maastricht

# **The Geographic dimensions of growth and development**

**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 Tuesday 25 November 2014, at 16.00 hours

by

Samyukta Bhupatiraju



**Supervisor**

Prof. Dr. Bart Verspagen

**Co-Supervisor**

Dr. Thomas Ziesemer

**Assessment Committee**

Prof. Dr. Eddy Szirmai (Chair)

Prof. Dr. Joan Muysken

Dr. Bart Los (Groningen University)

Dr. Alessandro Nuvolari (Sant'Anna School of Advanced Studies, Pisa, Italy)

# Contents

	Summary	i
	Acknowledgments	iii
Chapter 1	Introduction	1
Chapter 2	Summarizing Large Spatial Datasets: Spatial Principal Components and Spatial Canonical Correlation	15
Chapter 3	The Geographic Dimensions of Institutions	29
Chapter 4	Economic Growth, Institutions and Geography	77
Chapter 5	Multi-level Determinants of Inward FDI	107
Chapter 6	Conclusion	131
	Bibliography	137
	Valorization	143
	Samenvatting	147
	Biography	149



# Summary

One summary finding that has found much support in empirical growth and development literature is that a country's economic and social institutions are important in encouraging and promoting economic activity. The governance of these institutions result in either fostering or hindering the growth and development prospects of the country. Economic performance of countries is also often attributed to its geography. Since countries are not endowed in the same fashion, growth prospects are primarily dependent on these differences. According to the proponents of this literature the geography of a country has remained the same over time while countries have remained poor during and after the colonial era. This is because of the impact of geography on a country's on its growth prospects. Those who support the institutional hypothesis agree that geography is important but only indirectly and through its influence on institutions which in turn determine how a country performs. What this literature does not take into account are the spatial interactions between countries that lead to similar institutions in neighbouring countries. Neighbouring countries share common histories and are more likely to be influenced by each other than those further away. In this thesis, we propose that taking into account these spatial interactions leads to different conclusions

We start by introducing the literature and the motivation for this thesis in the first chapter. In the second chapter we introduce the methods of spatial principal component analysis (SPCA) and spatial canonical correlation analysis (SCCA). Using SPCA and SCCA methodologies we construct summary measures. The aim of these methods is to maximize spatial correlation underlying each dataset. While SPCA is used to summarize a single data set with a number of variables, SSCA is used to maximize spatial correlation between and within two data sets measuring two different phenomena. In this chapter we compare the two methods, providing examples for each. Further we compare our results to Wartenberg's method of spatial principal component analysis. In order to show the benefits of using these particular methods, we introduce an example application using geography data. We find that our methods give us results that are easier to interpret with a spatial narrative than a previously available method.

Using the methods developed and explained in Chapter 2, the aim of Chapter 3 is to use SPCA and SCCA to construct summary measures for various phenomena which are multi-dimensional in nature. We develop a new set of measures for institutions, absolute geography, economic and financial performance of countries. By doing so, we evaluate how the institutional quality measure is related to the latter three measures. This study thus contributes to the understanding of the interactions between the above mentioned phenomena using a spatial narrative. What we find is that when spatially weighted, the first components of the SPCA and SCCA conform to previous literature. However, the second components give us new stylized facts. We find that while economic development is highly correlated to high levels institutional quality, economic growth is strongly

correlated to higher levels of government intervention. Similarly, geography does not necessarily determine the economic fate of a country. Countries that share similar geography need not have the same or similar institutions. We conclude that taking account of space could change the way we understand and approach institutional change within a country.

In the next chapter, we use these spatially constructed measures of institutions and geography to see which of these along with trade best explain economic development. Using the Rodrik et al (2004) framework, which attempts to measure the contribution of each of these factors, we find that institutions are the most important explanatory variables only when we consider GDP per capita as our measure of development. When we use a broader definition of economic performance (which we construct using SPCA) we find that institutions lose their "primacy". Moreover, our results suggest that these same institutions could be detrimental to economic performance, suggesting that having high quality institutions are correlated to, but may not be the primal cause of economic performance.

Scale and spatial interaction need not be limited to cross country level. In Chapter 5, we shift focus to a multi-level analysis. We analyze the determinants of foreign ownership in domestic firms using a multi-level framework. Using data on firm structural and institutional variables as well as macroeconomic and institutional country characteristics, we estimate a multi-level logit model. This chapter contributes to existing literature by using a multi-level model which accounts for both micro and macro-level idiosyncrasies that lead to foreign investment. We find that firm structural characteristics and obstacles they face in day to day business are the most important determinants of foreign investment. At the country level, we find that macroeconomic variables that could influence profits of the firm are important and institutional variables are not influential in attracting foreign investment in host countries. Finally, Chapter 6 concludes the thesis by providing a summary of the results.

In summary, this thesis attempts to analyze the spatial interactions that influence institutional quality in countries. One contribution of this thesis was to develop a method by which we could maximize spatial correlation between variables in a data set. Additionally, using these methods, we attempt to understand the links between institutions, geography and economic performance. Further, the multi-level analysis, aims at disentangling country level perceptions of institutional quality and firm level perceptions to see which matter more for business enterprises.

# Acknowledgements

Writing this thesis has been a learning experience right from the start. It results from many experiences I have had over the last five years at UNU-MERIT and the many wonderful people I have met who I wish to acknowledge.

First and foremost, I would like to express my gratitude to my supervisors Professor Bart Verspagen and Professor Thomas Ziesemer. Bart Verspagen, thank you for supporting me with a research assistantship in the first year, for allowing me to teach your course at SBE and all the invaluable time and academic advise throughout these years. Thomas Ziesemer, thank you for showing me that all the difficulties and mistakes lie in the details. I would like to thank you both for the tremendous support and encouragement you have given me in writing this thesis. You have been a tremendous support without which I would have been lost, and this acknowledgement would remain unwritten.

It is difficult to overemphasize my thankfulness to my reading committee Adam Szirmai, Joan Muysken, Bart Los and Alessandro Nuvolari. Thank you for taking out time to evaluate my thesis and the valuable comments you have given me.

I have benefitted from working on the “Institutions, Governance and Long-term Growth” project with a team of researchers. Biniam Bedasso, Richard Bluhm, Luciana Cingolani, Denis de Crombrughe, Kristine Farla, Franziska Gassmann, Adam Szirmai, Kaj Thomsson and the AFD counterparts Nicolas Meisel and Jacques Ould Aoudia for the discussions and suggestions on my work. Adam Szirmai and Bart Verspagen, thank you for letting be a part of this project. Kristine Farla, a special thank you for being my econometrics guide and for all those discussions which helped me finish sooner than I would otherwise have.

I am grateful to the School of Business and Economics, Maastricht University for granting me the financial assistance and the teaching experience I gained from being a TA here. I would like to thank Silvana de Sanctis and Sylvia Beenen for their help with everything connected to SBE. UNU-MERIT would have been a much harder place to work in without the help of Eveline in de Braek, Sueli Brodin, Wilma Coenegrachts, Eric Engelen, Mourik Jan Heupink, Howard Hudson, Ad Notten, Herman Pijpers, Monique Raedts, Mittie Stenders, Marc Vleugels and the other staff that have helped in many ways. Eveline, thank you for being my 'go to' person with all the administrative questions.

No amount of thanking would show how grateful I am to have colleagues and friends who have helped me make Maastricht my home away from home. I have gained much not only from discussing my work with them, but also learned a lot from them outside academics. I would like to thank François Lafond, Alejandro Lavopa, Jocelyn Olivari, Daniel Opolot, Tatevik Poghosyan, and Giorgio Triulzi for being such a fabulous group of peers to have, I could not have asked for better. I would also like to thank Francesca Guadagno, Jennifer Taborda, Andrea Franco for always being there when I needed you,

no matter what. Nico Doranov, thank you for all the coffee breaks we had, and thank you and Asel for being the most amazing people I have met in Maastricht.

I am also grateful to Anthony Arundel, Théophile T. Azomahou, Irina Burlacu, Robin Cowan, Carlos Cadena Gaitan, Juan Carlos Castillo, Alison Cathles, Elisa Calza, Guney Celbis, Flavia Pereira de Carvalho, Nora Engel, Mulu Gebreeyesus, Charlotte Giullard, Jun Hou, Can Huang, Stefania Innocenti, Michiko Iizuka, Jojo Jacob, Ananth Kamath, Rodolfo Lauterbach, Hibret Maemir , Julieta Marotta , Gintare Mazeikaite, Bilal Mirza, Andres Mideros Mora, Pierre Mohnen, Dorcas Mbuvi, Paula Nagler , Cheng Boon Ong , Iman Rajabzadeh, Shyama V. Ramani, Shuan Sadre Ghazi, Simone Sasso, Gerry Silverberg, Luc Soete, Lina Sonne, Lilia Stubrin, Ajay Thutupalli, Raquel Tsukada, Eduardo Urias, Michael Verba, Daniel Vertesy, Adriaan van Zon, Jinjin Zhao, Nevena Zhelyazkova.

I will always be thankful to Anasuya Das, Hemali Shah, Nidhi Rao, Chirag Mehta, Janani Kedarnath, Keerthi Srilakshmi, Vaishnavi Suryanarayanan, Gowtham Varma, Shaik Rahil, Mayank Tomar, Rakesh Krishnan for always being only a phone call away. Kranthi Bhupatiraju for his constant company and encouragement everyday for the last five years.

Lastly and most importantly I would like to thank my family. My husband and best friend, Rahul Sirohi for being that one person I could count on to have an answer to all my doubts. I know it wasn't easy, but thank you for standing by me every step of the way. My brother, Sandeep whom I have turned and will continue to turn to for help all through life. My parents-in-law and brother-in-law for their continued love and support. My grandfathers whom I dearly miss. My parents Durga and Ramesh for being my greatest strength, my unconditional source of support and much more. This thesis is dedicated to them.

**CHAPTER 1**  
**Introduction**

## 1. Motivation

Among the most fundamental questions posed by economists since the establishment of the discipline itself, has been that of what causes economic growth and development. Why are some countries richer than others? Why have some of them managed to grow faster than others, some able to sustain growth while others stagnate? Why are some more productive than others? Traditional growth models starting from the Harrod-Domar model and Solow model emphasized the importance of saving, investment and human capital and improving efficiency in order to obtain more output per unit of input. The first endogenous growth theories relaxed the assumptions of diminishing marginal returns to scale, thus introducing positive externalities to explain economic growth. Endogenous growth models introduced by Lucas (1988), Romer (1990), Grossman and Helpman (1991) and Aghion and Howitt (1992) also endogenized technical progress and explained that those countries that invested more in technological innovations and knowledge experienced better growth prospects. While Lucas introduces endogenous growth via investment in human capital, Romer introduces it via investment in R&D.

Factor accumulation and technological change however were later believed to fall short in answering differences in economic growth countries. The above mentioned explanations of economic growth are described as *proximate* sources of growth (Maddison 1988; Rodrik 2003). They are considered proximate sources since the accumulation and assimilation conditions differ across countries and over time. These tend to be those growth causing variables that can be measured related to output. *Intermediate* sources of growth are those social, economic and technology policies as well as demand trends, which are in turn influenced by the *ultimate* sources of growth. Szirmai (2012) identifies policies that uphold macroeconomic stability, trade and financial openness that promote foreign investment and minimize protectionism among others as growth promoting intermediate sources. *Ultimate* sources of growth include socio-political and economic institutions prevailing in the country and their governance, physical geographic conditions of the country, culture and attitudes, prevalent social structures, historical shocks, etc. Szirmai (2012) further includes demographic and epidemiological trends, technological trends in the long run and distance from the technological frontier, absorptive capacities, class relations and political conflicts, international trade regimes.

Acemoglu (2009) classifies the ultimate or “fundamental causes” of growth into four categories of hypotheses: (i) the ‘luck’ hypothesis (ii) the geography hypothesis (iii) the culture hypothesis and (iv) the institutions hypothesis. The ‘luck’ hypothesis explains that the differences in growth among countries arose due to some uncertainty or heterogeneity that led to different choices that in turn led to different results (growth rates). This could also be due to different paths chosen among multiple equilibria. Chance or luck cannot be reasonably measured empirically (Acemoglu 2009). Accounting for all other sources of growth, it could be however be treated as a residual. Further, it is hard to decide on which factors are due to luck and which are not. The “geography” hypothesis refers to all the geographic factors that can lead to different rates of growth in the long run. Thus, factors such as soil quality, weather conditions, and natural resource availability affect productivity and attitudes directly. The “culture” hypothesis states that cultural

differences affect growth rates of an economy. Culture influences an individual's choice on such things as occupation, percentage of income saved/spent, trust towards other individuals, work ethics, political participation etc. Max Weber in his "The Protestant Ethic and 'The Spirit of Capitalism'" states that it was the protestant belief of Western Europe that allowed it to grow faster than other parts of the world. The fourth hypothesis is that of institutional development. According to this hypothesis, the institutional set up of an economy affects productivity.

Similarly, Rodrik (2003) in his framework of proximate and ultimate sources includes three key sources. These are geography, institutions and trade openness. In his view, geography is seen as the ultimate while trade openness and institutions are seen as intermediate sources that explain economic growth. The main difference between institutions hypothesis and the rest is that institutions are social choices that are historically persistent. They are endogenous in that they can change over time. Culture too is historically persistent (less so than institutional persistence) but is much harder to influence, and is outside the ambit of government directives and laws unlike institutions. Acemoglu (2009) puts it succinctly "Even though institutions might be hard to change in practice, culture is much harder to influence, and any advice to a society to change its culture is almost vacuous". Sen (2004) argues that culture important as it is, is dependent on the institutional structure that supports it, determining the incentives on which our actions are based. Although they take time to change, they can be altered to foster growth and hence development in the long run. Both the above mentioned frameworks focus on the importance of one factor as being more important than the others in influencing economic growth and development. It is however essential to recognize that these factors are not entirely independent of each other in their influence on growth and development. Abu-Lughod (1989) explains how chance played a role in the rise of European hegemony after the 13<sup>th</sup> century leading to growth prospects over the centuries that followed. She also points out how geographical and historical circumstances, as well as cultural influences, lead to different institutional features such as credit systems and the various financial instruments.

In this thesis, we will be focusing on the (1) institutions (and their governance) and (2) the geography hypothesis as explanations of economic growth and development which will be discussed in more detail below. Rather than focusing on the primacy of these factors, we study the interactions between them. We make two assertions in this study. One, institutional quality in a country cannot be studied within the confines of nation-state boundaries, given its own physical endowments, independent of influences from other nations. Influences from across the border have a sizeable impact on the economic performance of a country. Two, the analysis of institutions have to recognize that micro and macro level forces together form the institutional set up of a country, and need to be taken into account while studying their impact on economic performance. These will be discussed in the sections below. The reasons for concentrating on the geography and the institutions hypothesis is due their policy relevance. As mentioned above, if one country were to get lucky by chance, this cannot be replicated by other countries. Similarly, if one was to identify certain cultural practices that lead to improved economic performances of countries, it would be rather difficult to persuade others to change their practices from a

policy perspective. This is not to say that luck and culture do not play a role in a country's economic performance. We now proceed by outlining the literature on institutions and geography and then explain how geography and institutions are linked. We also explain how institutional quality should not be confined to the country level to understand how it impacts economic performance.

## 2. Institutions

There are three main views on why and how institutions cause growth. The first view is one in which institutions are seen as posing “constraints” on free markets, which reduce inefficiencies. According to North (2003) “institutions would not exist in a frictionless world where there is no uncertainty. Institutions exist to reduce uncertainty in the world.” He further notes that they constitute a set of constraints on behaviour in the form of rules and regulations which enable markets to function efficiently (North 1984, 1990). A second view sees institutions as “enabling” mechanisms which allow for smooth functioning of the economy, even if this is obtained through constraining certain actors, certain types of behaviour or activities (Chang and Evans 2000). Another view is one in which institutions are both constraining and enabling. Additionally, they are also “constitutive” (Chang 2002). This is to say, institutions have a cultural dimension which imposes a certain “world view” into people who live under a certain institutional set up. Hence, there is a certain degree of path dependency involved. This view (still largely ignored according to Chang) states that there is no clear boundary between institutions and culture in the long-run. According to this view rules practiced for a long time become embedded in the society, and eventually become a part of their “culture” (Srivastava 2004). Whether seen as constraining, enabling or constitutive, the consensus among institutionalists emerges in the importance of studying institutions in determining economic growth and development in countries across the world.

According to the current literature, “good institutions” provide an environment with the “right” incentives to agents in the economy. These incentives lead to learning and innovation, encourage investment in human and physical capital and hence increase capabilities. These factors in turn lead to higher growth rates in an economy. An improper or “bad” institutional set up on the other hand, does not incentivize such activities. Instead, it leads to corruption, rent-seeking and other non-productive or growth repressing activities (Chang 2001, Bennedsen et al. 2005).

To create these right incentives, the story goes (Bennedsen et al. 2005); states have to ensure that the private return (profits earned by risk taking individuals) and social return (returns to the economy) of an economic activity should be brought as close as possible since productive activities have a high social return, and are good for the growth of an economy. This would also include reducing uncertainty in the creation and expropriation of wealth, by ensuring return. Lessening uncertainty allows in building agents’ confidence in the economy. Consequently, if rules are followed by all the agents in the economy, it would lead to more productive economic activity, and would also reduce transaction costs.

‘Institutions’ cover many factors that affect the growth rates and development in an economy. Among these secure property rights is considered as one of the most important that foster growth. Well-defined property rights allow agents to invest in improvements in land or install state of the art technology, with long term gains in mind. This in turn improves productivity. In the case of no proper security available, agents would hesitate to plough back their profits to improve their productivity. Another important economic institution is the existence of markets and the quality if these markets. In most developing countries, one finds missing markets as well as low quality of existing markets. If there are no proper markets, the incentives to trade are reduced and this in turn implies improper allocation of resources. This results in an unequal distribution of wealth.

Among the political institutions, democracy is considered to be one of the most important institutions. In a democratic economy, one is ensured of a voice in the system, and is ensured both economic and social rights. However, empirical and historical evidence does not always support the need for democracy to ensure high growth rates. For example, South Korea under General Park saw very high growth rates (even though this could be attributed to the “catching up” that the economy achieved after its transition into a market economy).

Social institutions are as important as economic or political institutions when economic growth in terms of increased GDP, as well as economic development in terms of equality is to be achieved. Social institutions consist of proper health care markets, which ensure basic health needs of workers in an economy. When a proper health system is not in place, it will result in reduced productivity. In tropical countries, since the disease environments are much larger, it becomes more of a necessity to ensure proper health care for all. Most underdeveloped economies also suffer from a high population growth rate. Along with the lack of other institutional infrastructure this results in a higher rate of unemployment. Most of these countries also lack a proper unemployment support system. Labour market regulation through job search facilities can help an economy to produce more. Informal social institutions such as social norms present in a country also have an impact on productivity.

The mere presence of institutions themselves does not guarantee a well-functioning economy, and thus one that leads to growth and development. The governance of these institutions is equally important. Institutional development under the slogan of “good governance” has come to the forefront in recent years. Governance refers to the “essential parts of the broad cluster of institutions” (North et al. 2008). The interactions between governance and growth are “intimately linked” to interactions between institutions and economic growth (North et al. 2008). This will be discussed below.

## **2.1. Governance**

Governance allows an economy to reduce uncertainty and increase confidence of individuals in it, by providing assurance that rules will be followed by all individuals. This increases economic activity in the economy and thus leads to a better performance. Kaufmann et al. (2000) define governance as comprising “of the traditions and institutions by which authority in a country is exercised. This includes the process by which

governments are selected, monitored and replaced; the capacity of the government to effectively formulate and implement sound policies; and the respect of citizens and the state for the institutions that govern economic and social interactions among them.” The World Bank measures governance using six indicators (World Governance Indicators). These are: Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Regulatory Quality, Rule of Law, and Control of Corruption<sup>1</sup>.

In his book “Kicking Away the Ladder: Development Strategy in Historical Perspective” Chang includes the following institutions as constituting the “good governance package” that is currently being promoted as that which will lead to economic growth and development in developing countries. The list consists of (i) Democracy, (ii) Bureaucracy and the Judiciary, (iii) Property rights regimes and their quality, (iv) Corporate Governance which includes limited liability, bankruptcy law, financial reporting and information disclosure to public, and a competition law, (v) Financial institutions which include banking regulation, central banking, securities regulation and public finance institutions, (vi) Social welfare and labour institutions which include social welfare institutions which work as ‘safety nets’ for the economy, institutions to eradicate child labour, and institutions regulating working hours for the adults.

The goals of anti-corruption and democracy can hardly be questioned as ends in themselves. The post-war period gave rise to arguments that democratic institutions were expensive for countries to maintain. The present institutional view however is that democracy is a pre-condition for economic development. On the other hand, historical evidence shows that the current advanced countries did not have democratic institutions before the time of take-off. Empirical evidence is also ambiguous. Butkiewicz and Yanikkaya (2004) finds that democratization of institutions lead to higher growth rates. Persson and Tabellini (2006) finds that economic development is dependent on the type of democratic reforms. Barro (1996, 1999) shows that while a minimal level of democracy is necessary for productive economic activity to occur, higher levels of democracy have a negative impact on growth. Papaioannou and Siourounis (2008) find that democratization has long run impacts, but does not show any significant increase in growth rates in the short run. Alesina et al. (1996) finds an insignificant relationship between democracy and growth while Dollar and Kraay (2003) find a negative and significant relationship between the ‘Voice’ variable of the World Bank governance indicators. Recent historical evidence has shown that those developing countries that have successfully transformed themselves into fast growing countries have not had the prescribed “investment climate” of a ‘service delivery state’ and instead their governments have played a critical role in assuring structural change (Khan 2005). We can also state that there are no historical examples wherein improved governance has been undertaken first, and growth rates have risen after these changes (Chang 2002).

The existence of corruption in an economy leads to inefficiencies no doubt. Khan (2006) identifies a circular relationship between economic stagnation, high transaction costs,

---

<sup>1</sup> <http://info.worldbank.org/governance/wgi/index.aspx#home>

weak property rights, corruption and rent-seeking and unaccountable government. Empirical evidence is abundant to show that there is a strong negative relationship between the two (Mauro 1995, Welsch 2003). There are counter examples such as Indonesia under Suharto, where corruption was at its highest levels when the country achieved very high growth rates. India and China, which have been growing at rates over 5% also, rank very low on the transparency index, at around 3.5 out of 10.

The implicit advantages of improved governance, such as lack of corruption or the presence of democracy are not unknown to us. However, major questions arise: can there be a universal solution to all institutional problems of economic growth? Can the institutions that have worked for the developed economies be transferred to developing economies and will this improved governance and increased economic growth in these countries, irrespective of the historical, geographical and socio-economic situation of the economy? The question thus becomes one of “can the standard analysis be applied to developing countries?”

## **2.2. Governance in Developing Countries**

The current global political scenario is different from what it was when the now developed economies were still in the developing stage. The requirements to be fulfilled by all countries are more sophisticated now. Globalization and integration of world markets have led to greater sophistication of institutions now than those of developing countries earlier. These global requirements in themselves lead to some degree of universal standardization of institutions. It is nonetheless important to realize the importance of different social set ups of different countries.

Critics argue that international financial institutions intervening in ‘governance issues’ modeled on those of developed countries are too demanding both financially and on human resource requirements for developing countries. It should also be noted that some of these reforms go against the social and cultural values of some countries. Institutional transplantation has proven to be difficult and quite unsuccessful in countries with different conditions than those in which the prescribed institutions have been fostering growth. Given that social, political and cultural factors also contribute to growth, imposing a common institutional standard may not be wise. Nevertheless, to avoid the justification of all existing institutions in the world, it is necessary to find some institutions that foster growth in given situations. Transplanting institutions has proven to fail in more than one case. The Structural adjustment programmes undertaken have failed in Latin America, and transition from a controlled economy to a market economy has resulted in greater poverty in erstwhile USSR. China, on the other hand, seems to have found a dual approach which ensures elimination of missing markets over a period. Learning from historical examples is another approach (Chang 2002).

Thus, even though there is a need to understand the relationship between various governance and institutional factors and growth, it need not necessarily be done through universally applicable standards. In reality, we find countries at various stages of development, with a wide variety of institutional arrangements and yet have similar growth rates. Thus, it is important to be able to understand and answer this phenomenon.

### 3. The Geography Hypothesis

The geography hypothesis is based on the fact that not all the areas in the world are endowed in the same fashion. According to this hypothesis, the prosperity of an economy depends on its geographical endowments. It determines that characteristics such as soil quality which directly affects agricultural productivity, natural resources which contribute to wealth as well as provide the raw material base for industrial activity, topography which contributes to costs of transportation, climate that affects the productivity of workers, and the disease environment which also directly affects the productivity of workers (Acemoglu 2009, Diamond 1997, Gallup et al. 2001, Sachs 2001, Landes 1998). Similarly, lack of access to rivers as well as land-locked geography affects the growth of the country negatively. The geography hypothesis gained popularity over the last decade with some empirical work done on it as well.

The proponents of this view argue that during and after colonization the poor countries have remained poor. There has been no reversal of fortunes. However, what has remained the same is the geographical position of these countries (Sachs (2001) and Bloom and Sachs (1998)). Thus, it is primarily geography which determines the economic situation of a country.

#### 3.1. How does geography affect growth?

There are three different versions of the geography hypothesis. The first version of this theory comes from Montesquieu in 'The Spirit of the Laws' which states that the climate in a country affects physical strength and leads to laziness and passivity. Other proponents of this view include Marshall. He believed that racial qualities could be explained by climate.

A more sophisticated second view on how geography affects growth was developed by Gunnar Myrdal. In his *Asian Drama: An Inquiry into the Poverty of Nations* he writes "serious study of the problems of underdevelopment...should take into account the climate and its impact on soil, vegetation, animals, humans and physical assets- in short, on living conditions in economic development." (Myrdal 1968). This view states that the productivity of technology of the countries in the temperate zone were far higher than those of technologies developed in the tropics. This argument refers mainly to technologies related to agriculture. However, when actual data are compared, it is during and after the industrial revolution that differences in incomes and growth rates started to become wider.

The third view states that the disease environment is a burden to the country, and therefore, is directly linked to the poverty that is prevailing in it. This view has become popular over the past decade and has been treated in empirical studies such as Sachs (2000), Bloom and Sachs (1998), Gallup and Sachs (2001), etc. While this view seems more plausible than the previous two versions, the causality of poverty in the wake of disease

environment and economic development is not very clear<sup>2</sup>. McArthur and Sachs (2001), concludes that institutions and geographically-related variables affect the GNP per capita of countries.

#### **4. Institutions, geography and space**

The underlying narrative of the geography hypothesis is that the natural endowments of the country determine the development path of a country, and this becomes the destiny of the country. This geographical determinism has been championed by economists like Jared Diamond, Jeffery Sachs, Andrew Mellinger, John Gallup, etc. They assert that these natural endowments encourage or hinder a country's ability to trade with others, its agricultural productivity, etc., thus making absolute geography the ultimate determinant of economic growth and development. The reason this hypothesis attracts attention in the empirical literature is because geographical variables such as the length of the coastline, number of frost days in a year, the presence of specific types of parasites, are all "exogenous." This makes absolute geographic variables "as exogenous as an economist could hope for (Rodrik et al. 2004).

This hypothesis is contested by the institutional hypothesis, which states that institutions are more important than geography in explaining economic growth and development. While agreeing that absolute geography does have an impact on growth and development, it is relegated to a secondary role, one in which it influences the type of institutions that develop in these countries. The story goes as follows: the colonizer (European) countries had a set of 'good institutions' conducive for growth, which in turn were developed due to the exogenously given good absolute geographic conditions. These good institutions were taken along with them to their "settler colonies." In countries where the disease ecology and harsh living conditions made it difficult for them to settle, they built extractive institutions were set up. These historical, colonial institutions were path dependent and have since persisted, detrimental to current economic prospects of these countries while settler colonies prospered.

While both these hypotheses use absolute geography as the backcloth to explain past and present economic growth and development, they have been pitched against each other as contesting explanations of what causes growth and development.

The geography hypothesis assumes that absolute geography of a country distorts an otherwise level playing field in which any country could have achieved development before the others. This hypothesis has been criticized by many, given the fact that countries in the tropical regions were once the thriving countries in the world, before the rise of Western Europe (Pomeranz 2000, Blaut 1999). Geographical location was, therefore, not the main reason for the West to get ahead in the last two hundred years. Also, many

---

<sup>2</sup> Western Europe was plagued by many diseases in the previous two centuries. It was economic development which led to control of diseases and also providing healthier living conditions. (Acemoglu 2009). Diseases such as malaria in the tropics claim a lot of lives each year. However, this could as much be a consequence of the inability of provision of healthy living conditions.

of the tropical countries do not conform to this pattern. For example, Chile, Peru, Argentina and Brazil have overcome their monetary instability, and over the last few years, displayed very respectable rates of growth. Similarly, China and India have had high growth rates, joining the list of emerging countries in the recent past. The dynamics of growth shown by these countries against the East Asian tigers which were the emerging countries in the previous decade is not answered by the geography hypothesis.

The institutional hypothesis, on the other hand, assumes that the evolution of institutions is more or less internally driven and for the most part these institutions are not influenced by the regional or international context in which they exist. Conversely, historical evidence shows that both these assumptions are incorrect. There are ample examples that show that geography is not destiny, the spatial interactions between countries, i.e., the relative geography of a country as an influence on the local mechanisms is not difficult to see. Peet (2011) points out that public health investment has helped developed countries counter the environmental threat, arguing that it is economic development that results in the social reconstruction of better environments. He further notes that since Western European exploitation of the world was mainly via sea transportation (as opposed to land-transportation), ports and navigable rivers had seen a higher GDP than other parts of the world. This, he says, was a historically-socially constructed advantage, rather than an inherent advantage that water had. Sheppard (2011) points out that the Gaza strip has the right set of absolute geographic conditions, but is extremely influenced by the regional socio-political situation. Similarly, while landlocked Switzerland has “bad geography” it is influenced by the institutions of its surrounding countries. The influences need not necessarily be regional. For example, development over the past two centuries has led to the overuse of the world’s natural resources, and has resulted in the depletion of the ozone layer. The economic consequences of which are faced by Australia. More recently, Lin et al. (2014) suggests that trans-boundary pollution caused by Chinese production and transportation of goods to the West coast of the USA has increased, even as pollution caused by production within the US has reduced due to outsourcing production. Giving the example of the impact of the Indian IT education on the Silicon Valley labour market, Harvey (2011) notes that space need not be continuous. Hence, it becomes important to study these spatial interactions, which could have significant impacts on the evolution of institutions within countries. We can thus say that analyzing institutions as determinants of growth by assuming them to be independent can lead to wrong conclusions. In this thesis, we want to take into account the socio-spatial influence that countries have on each other’s institutions. In order to do so, we take account of the relative geography/ spatial interactions of countries.

Most growth literature that takes into account relative geography/ spatial interactions have been carried out using country data or regional data, most often studying regional convergence of incomes, growth, technology diffusion, etc. Regional income convergence on intra-national scale as well as regional scale (EU) has resulted in strong consensus on income convergence among these geographical units. The theoretical mechanisms of technology diffusion, factor mobility, and transfer payments that drive regional convergence have explicit relative geographical components (Rey and Montouri 1998, Quah 1997, Cheshire and Magrini 2009, López-Bazo et al. 2004, Fingleton, and López-

Bazo 2006, Verspagen 2010 among others). Gleditsch and Ward (2000), and Cho (2003), Beck, Gleditsch and Beardsley (2006) study the influence of spatial interactions on democracy.

When analyzing diffusion of any socio-economic or political phenomenon, one needs to specify “proximity” very carefully. While there is no specific rule regarding which distance matrices should be used, and which should not be, some common ones are Euclidean distances, trade dyadic distances, shared political borders, etc. In this thesis, we use bilateral distances between countries to understand the diffusion process of institutional quality.

## **5. Institutions, FDI and scale**

While studying the socio-spatial aspects of institutions, however, we need not limit our understanding of them to the scale of nation-states. Scale effects can be extended to firm-level and can in turn be linked to country level, to understand the impact of institutions on economic growth. In the last chapter of this thesis, we turn to the influence of institutional quality at the firm level, within the country level. In order to do so, we choose to study the impact of institutional quality at both levels on the choice of foreign firms to invest in host countries. Capital mobility across countries has a strong spatial structure. These capital flows, are usually dependent on various characteristics of host countries, including their macroeconomic performance and institutional performance.

Endogenous growth theory and related empirical growth studies show that by taking relevant policy measures a country can influence its long run growth rates. One such channel is by attracting foreign direct investment (FDI) into the country. International organizations such as the World Bank and the IMF have stressed that importance of opening up and liberalizing trade in order to generate growth in developing and less developed countries. Although there is no consensus on the influence of FDI on growth, there are numerous studies that advocate opening up of trade in these countries in order to allow FDI, which in turn would boost growth in these economies. FDI flows across countries have increased ten-fold over the last twenty years. Host country economic policies make efforts to attract FDI by making their markets conducive for MNE's to enter. Along with the economic environment, a strong institutional credibility plays a role in attracting FDI into countries (Naude and Krugell 2007, Sharma and Bandara 2010). Rule of Law, democracy, low levels of corruption and flexible labour market institutions are all said to important country characteristics that lead to increased inward FDI flows (Daude and Stein 2007, Delbecque et al. 2008, Pajunen 2008, etc.).

The scale at which a study is done is important the conclusions and policy suggestions made. There is a tendency of extrapolating results in such a way that results from macro-level studies are used to make micro-level policy conclusions (Meentemeyer 1989). Busse and Hefeker (2007) perform a cross-country analysis and find that government stability, religious tensions, and democratic accountability, as well as corruption and bureaucratic quality impact the volume of inward FDI flows in countries. They go on to conclude “these political risk and institutional indicators matter the most when multinational corporations confront decisions about where to invest in developing countries.”

Aggregate FDI flows are a macro-level measure where as a firm's entry decision is a micro-level phenomenon. In order to see if conclusions such as the one above can be made, we study the impact of institutional quality and other macroeconomic variables along with firm structural variables on inward FDI. We aim to understand with in and between country institutions and other economic determinants of FDI in developing countries.

## 6. Thesis Outline

The main objectives of the thesis are to understand the socio-spatial interactions of institutions between countries, and how this influences the level of economic performance in them.

In the first part of the thesis, we focus on the debate on the fundamental causes of economic performance which is centered on the institutional and geography hypothesis. In the second part of the thesis, we move shift focus to the inter-country differences in institutional performance and their impact on attracting capital into the host countries, which in turn are believed to influence economic performances.

Institutions, economic performance and geography are all multi-dimensional in nature. Single variables such as democracy, GDP per capita or latitude do not fully capture these phenomena. Using such single variables, as has been done in most literature, only leads to partial understanding of economic performance. Given this premise, we introduce alternate methods of spatial principal component analysis and spatial canonical correlation analysis. Both these methods are dimensionality reduction methods that allow us to summarize large data sets into smaller latent dimensions that are uncorrelated to each other. The second chapter introduces spatial principal component analysis and spatial canonical correlation analysis. These methods result in summary measures of any given data set that has maximum spatial correlation. The spatial principal correlation analysis is used when we want to summarize a single data set, and obtain the spatial correlation within the data set. Spatial canonical correlation is used when we have two data sets, each measuring a different phenomenon, that need to be summarized, while obtaining the spatial correlation between the two sets of data. Chapter 2 introduces these two methods with examples for each of them. We compare our method and the results, with previously available spatial principal component analysis by Wartenberg (1985). Further, we show how the two methods proposed produce different results.

Using these methods elaborated in Chapter 2, we construct summary measures of institutions, economic performance, absolute geography and financial performance of countries. Using these new measures (variables), in Chapter 3, we examine the role of institutions relative to a country's economic performance, absolute geography and its financial performance, in the context of its relative geography. The intention of this chapter is to identify if relative geography is an important aspect that needs to be taken into account in order to better understand economic performances of countries in the context of institutional quality.

In chapter 4, we test the Rodrik et al. (2004) framework to explain the differences in development levels across countries using the new measures of institutions, economic

performance, and absolute geography, constructed in Chapter 3. Rodrik et al. (2004) use a cross-sectional instrumental variable regression to measure “the respective contributions of institutions, geography and trade in determining cross-country income levels” using instruments for institutions and trade. They conclude that geography has “at best” an indirect effect on the income of a country and that “the quality of institutions ‘trumps’ everything else.” We test if this is indeed true, using the spatially weighted measures of each of these variables, using the same instruments as in their paper.

After having analyzed the importance of institutions from a “territorial” perspective, we move to analyzing their “scale” aspects. In Chapter 5, we analyze the determinants of foreign direct investment (FDI) using a multi-level regression framework. We use micro-level and macro-level measures of institutions and other economic determinants of FDI, and see which of these determinants are the most important in attracting inward FDI flows into developing countries.

Chapter 6 concludes the thesis. This chapter provides a summary of the results and discusses the limitations of the study.



## CHAPTER 2

# Summarizing Large Spatial Datasets: Spatial Principal Components and Spatial Canonical Correlation<sup>3</sup>

---

<sup>3</sup> This chapter has also been published as a UNU-MERIT working paper 2013-011 (Bhupatiraju, Verspagen and Zieseimer 2013a)

## 1. Introduction

Principal components analysis is an often used technique used to summarize data. It is based on correlation analysis, and can, roughly, be seen as a way to summarize a dataset of many variables into only a few dimensions. Spatial analysis has a counterpart of correlation analysis, e.g., in the form of the Moran coefficient. The Moran coefficient measures the extent to which a phenomenon, measured by an indicator, is spatially concentrated (a positive spatial correlation), or spread-out (a negative correlation).

As spatial datasets often consist of a large number of variables, it was only natural that Wartenberg (1985) turned to the principal components method in order to summarize spatial correlations. He proposed a method that was closely analogous to principal components. As “regular” principal components essentially means to undertake a spectral decomposition (obtaining eigenvalues and eigenvectors) of the correlation matrix of a dataset, Wartenberg (1985) simply proposed to spectrally decompose the matrix of Moran coefficients of a spatially organized dataset.

While the spectral decomposition has a clear interpretation in the case of regular principal components, it does not have the same, nor an alternative, clear-cut interpretation in the case of Wartenberg’s proposed method. His justification of the method was purely at the intuitive level, both by the analogy of taken eigenvalues and eigenvectors, and by the results that it produced for a number of constructed datasets that show different kinds of spatial dependence.

Below, we will first propose an alternative method for undertaking spatial principal components. Like Wartenberg’s method, it is based on spectral decomposition. However, rather than taking eigenvalues and eigenvectors of the raw matrix of Moran coefficients in a dataset, we propose to spectrally decompose a slightly different matrix. Our method results from a clear objective, i.e., that the Moran coefficient of the resulting component is maximized. Thus, our summary measure is aimed at showing maximal spatial correlation itself.

The spatial principal components method that we propose can also be extended in a direction that is similar to canonical correlation analysis. Canonical correlation analysis is a way of summarizing two separate datasets, each into one or a few components, in such a way that the correlation between these components, across the datasets, is maximized. Since our spatial principal components method already maximizes spatial correlation within a single dataset, it can easily be extended to maximize spatial correlation between the summary measures of two datasets.

The rest of our paper is organized as follows. The new method for spatial principal components analysis is explained in the next Section (2). Section 3 presents the spatial canonical correlation method. Section 4 provides a few applications of the two methods, illustrating their use, and comparing the weighting schemes that they produce to Wartenberg’s method, and regular principal components.

## 2. Spatial Principal Components: A New Method

Like Wartenberg (1985), our method starts from the Moran coefficient (e.g., Cliff and Ord, 1981). We will denote our data matrix of  $n$  observations (spatial units) and  $k$  variables by the symbol  $X$  (i.e.,  $X$  is an  $n \times k$  matrix of observations). Throughout the paper, we will assume that the  $k$  variables are  $z$ -scores, i.e., that for each variable the mean is zero and the standard deviation is one (this simplifies the notation). The  $n \times n$  matrix of spatial weights is denoted by  $W$ , and we assume that the sum of elements of this matrix is equal to  $n$  (i.e., the weights matrix is standardized). Then, the  $n \times n$  matrix of Moran coefficients between the  $k$  variables is calculated as

$$M = X^T W X. \quad (1)$$

The superscript  $T$  denotes the matrix transpose.

In case of no spatial weighting ( $W$  is the unity matrix), equation (1) would yield a matrix of normal (Pearson) correlation coefficients. In that case, an eigenvector of the correlation matrix, denoted by  $u$ , would yield a set of weights that could be used to calculate a composite measure  $Xu$ . This composite measure is actually a projection of the data matrix  $X$  onto the vector  $u$ , and it can be shown that the eigenvector that belongs to the largest eigenvalue is associated with the projection that minimizes the residual variance between  $X$  itself and the projection. In other words, that eigenvector maximizes the “fit” between the data and the lower-dimensional projection. The eigenvector associated with every next largest eigenvalue correspondingly maximizes the fit of the remaining residual variance. This procedure is called principal components.

Note that the length of the vector  $u$  is constrained to one, so that the projection is identified, and hence the procedure of minimizing residual variance (or maximizing fit) is a problem of constrained optimization. In particular, the principal components procedure maximizes  $u^T X^T X u$  subject to  $u^T u = 1$ , which yields the first-order condition and eigenvalue problem  $X^T X u = \lambda u$ .

Wartenberg’s spatial principal components analysis takes eigenvalues of the matrix  $M$ , which can be seen as the outcome of maximizing  $u^T X^T W X u$  subject to  $u^T u = 1$ , which yields the first-order condition and eigenvalue problem  $X^T W X u = \lambda u$ . Note that this can also be written as  $M u = \lambda u$ , which brings out clearly that what we are, in fact, doing in this case is taking the eigenvalues of the matrix of Moran coefficients  $M$ . Thus, Wartenberg maximizes  $u^T X^T W X u$ , which can be seen as the covariance of the non-spatially weighted factor ( $Xu$ ) and the spatially weighted factor ( $W Xu$ ). While this is mathematically sound, we argue that this maximization does not have a clear, intuitive interpretation, and hence the procedure is in need of an objective that is more directly related to the basics of spatial analysis.

Our proposal is to calculate a set of weights  $v$  in such a way that the Moran coefficient of the weighted summary variable  $Xv$  (which, as in the case of principal components, is the projection of the original data on the vector  $v$ ) is maximized. This idea is the intuition that we want to put behind the spatial principal components procedure, which then be-

comes aimed at finding a summary measure that yields maximum spatial correlation. Mathematically, it can be implemented as follows. As long as the vector  $Xv$  is a  $z$ -score, the Moran coefficient that we want to maximize can be written as

$$(Xv)^T W(Xv) \tag{2}$$

The mean of  $Xv$  is zero because  $X$  is already expressed as  $z$ -scores. However, in order for  $Xv$  to be a  $z$ -score, we also need to choose the weights  $v$  such that the variance of  $Xv$  is one. Hence, given our objective of maximizing the Moran coefficient of  $Xv$ , we have to maximize equation (2) subject to the condition  $(Xv)^T(Xv) = 1$ . Note that in comparison to Wartenberg's procedure (as outlined above), we only propose to change the constraint, not the objective function. Wartenberg (implicitly) uses the constraint  $v^T v = 1$ , whereas we use  $(Xv)^T(Xv) = 1$ .

It can relatively easily be shown that this constrained maximization problem leads to the following first-order condition:

$$(X^T X)^{-1}(X^T W X)v = \lambda v, \tag{3}$$

where  $\lambda$  is the Lagrange multiplier. This is clearly an eigenvalue problem. On the left hand side of (3), we find the matrix of Moran coefficients (1), pre-multiplied with the inverse of the covariance matrix of the non-spatially weighted data. The solution of our constrained maximization problem is an eigenvector of this left hand side matrix. The eigenvector that belongs to the largest eigenvalue will maximize the Moran coefficient of the projection  $Xv$ , and the corresponding eigenvalue will be equal to the largest possible value of the Moran coefficient, given the constraint of unit variance.<sup>4</sup>

Summarizing, our modification of Wartenberg's procedure is to pre-multiply the Moran matrix  $M$  of the variables in the dataset by the inverse of its (non-spatially weighted) covariance matrix. Doing so provides a clear interpretation of the eigenvectors of the matrix: this eigenvector is the set of weights that maximizes the spatial correlation of the projection of the raw data on to this eigenvector.

### 3. Spatial Canonical Correlation

The method of spatial principal components as outlined in the previous section has an intuitive extension into the direction of canonical correlation analysis. Suppose that instead of just one dataset, we have two, i.e., a matrix  $X$  and a matrix  $Y$ . Both  $X$  and  $Y$  have  $n$  rows (observations), but their number of variables (columns, denoted by  $k_X$  and  $k_Y$ ) may differ. As before, we assume that  $X$  and  $Y$  are  $z$ -scores.

---

<sup>4</sup> Note that we can also obtain (3) through maximizing  $v^T(X^T X)^{-1}(X^T W X)v$  subject to  $v^T v = 1$ , which would be a more general form that does not require the data to be standardized.

The aim of our spatial canonical correlation analysis will be to find two sets of weights, vectors  $v_X$  and  $v_Y$ , such that the weighted summary variables  $Xv_X$  and  $Yv_Y$  are maximally spatially correlated to each other. This can be achieved by maximizing the Moran coefficient

$$(Xv_X)^T W (Yv_Y) \quad (4)$$

subject to two conditions:  $(Xv_X)^T (Xv_X) = 1$  and  $(Yv_Y)^T (Yv_Y) = 1$  (the rationale for these conditions is the same as in the previous section). After re-arranging, this problem yields two first-order conditions:

$$(X^T X)^{-1} X^T W Y (Y^T Y)^{-1} (W Y)^T X v_X = \lambda_1 \lambda_2 v_X \quad (5)$$

$$(Y^T Y)^{-1} Y^T W^T X (X^T X)^{-1} X^T W Y v_Y = \lambda_1 \lambda_2 v_Y \quad (6)$$

where  $\lambda_1$  and  $\lambda_2$  are Lagrange multipliers. Each of these conditions is an eigenvalue problem. In fact, the first  $m$  (where  $m$  is the minimum of by  $k_X$  and  $k_Y$ ) eigenvalues of the two problems will be identical, and they will be equal to the square root of the (maximized) Moran coefficient that is associated to the eigenvectors in each of the two conditions. Defining  $A = (X^T X)^{-1} X^T W Y (Y^T Y)^{-1}$  and  $M = X^T W Y$ , we get

$$A M^T v_X = \lambda_1 \lambda_2 v_X$$

$$A^T M v_Y = \lambda_1 \lambda_2 v_Y$$

The solution then is

$$\lambda_1 \lambda_2 = v_X^T A M^T v_X = v_Y^T A^T M v_Y$$

Thus, the two eigenvalue problems (5) and (6) provide a different perspective on summarizing information in a spatial way: for each of the two datasets  $X$  and  $Y$  they provide a summary measure ( $Xv_X$  or  $Yv_Y$ ) that maximizes spatial correlation to the other dataset. The eigenvectors belonging to the largest eigenvalue in either (5) or (6) provide this summary measure, while the square of the largest eigenvalue is equal to the Moran coefficient between the two summary measures.

## 4. Applications

We now proceed to illustrate the proposed procedures to some real-world data. The datasets that we use are taken from Gallup et al. (1999), who provide a wealth of information on geography related phenomena on a global, country-level scale. We use data from two of their subsets of data: the physical characteristics of countries (mainly access to waterways and climate), and agriculture-related indicators (soil condition and climate zone). The data are available for 162 countries, and we use data from the CEPII dataset

on distances between countries (Mayer and Zignano, 2011) in order to construct our spatial weights (we use the distance between the largest cities in a pair of countries). These spatial weights are constructed as binary weights, where each (row) country has a weight of one for each of its closest 10 neighbour countries, and zero for other countries. This matrix is symmetrized by taking the average of cells  $(i,j)$  and  $(j,i)$  and assigning this average to both cells. We use such a symmetric spatial weights matrix because it makes the interpretation of the various Moran coefficients somewhat easier (e.g., it produces a symmetric Moran matrix  $M$ ), but we have also applied the method to cases with a non-symmetric matrix  $W$ .<sup>5</sup> Finally, the spatial weights matrix is standardized by dividing each cell by the matrix sum (which obviously is  $162*10$ ).

#### 4.1. Spatial Principal Components

We start by applying spatial principal components analysis to the dataset on physical characteristics of countries. This dataset has 13 variables, which are presented in Table 1. Nine of the 13 variables relate to access to waterways, either navigable rivers, or sea. The emphasis on this characteristic comes from the attention to landlocked countries in the debate on economic development (e.g., Gallup et al., 1999). The other variables relate to elevation, climate, and population density.

We apply regular (non-spatially weighted) principal components analysis, Wartenberg's spatial principal components analysis, and our own spatial principal components analysis to this dataset.<sup>6</sup> We start by looking at what would be called loadings in regular principal components analysis, i.e., the eigenvectors that were extracted in each of the three procedures. These are documented for the first two components (two largest eigenvalues) in graphical format in Figure 1a and 1b.

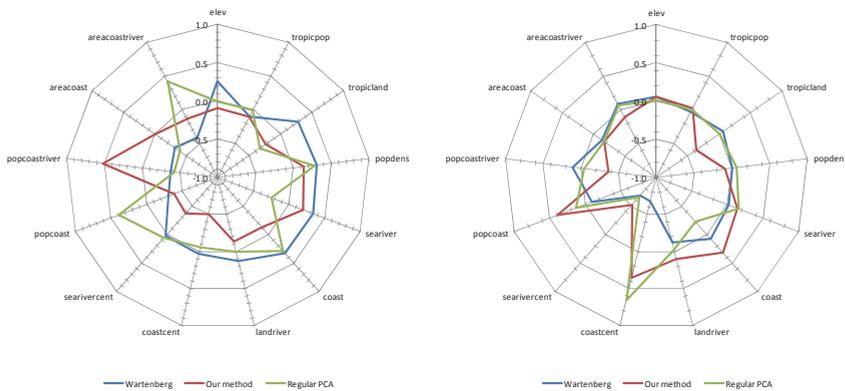
---

<sup>5</sup> If the spatial weights matrix ( $W$ ) is non-symmetric, the matrix of Moran coefficients between the 13 variables will also be non-symmetric. This means that that the eigenvalues in the two spatial variants of principal components (and spatial canonical correlation analysis) potentially have imaginary parts. When we worked with non-symmetric  $W$  matrices (i.e., not in this paper), we always obtained large purely real eigenvalues.

<sup>6</sup> All our calculations are done in Matlab. The functions that were written to perform the analysis are available on request.

**Table 1. Variables in the physical characteristics dataset<sup>7</sup>**

Variable description	Variable code
mean elevation (meters above sea level)	<i>elev</i>
mean distance to nearest coastline or sea-navigable river (km)	<i>seariver</i>
mean distance to nearest coastline (km)	<i>coast</i>
mean distance to nearest inland navigable river (km)	<i>landriver</i>
distance from centroid of country to nearest coast (km)	<i>coastcent</i>
distance from centroid of country to nearest coast or sea-navigable river (km)	<i>searivercent</i>
Ratio of population within 100 km of ice-free coast to total population	<i>popcoast</i>
Ratio of population within 100 km of ice-free coast/navigable river to total population	<i>popcoastriver</i>
population in the geographical tropics (%)	<i>tropicpop</i>
% Land area within 100 km of ice-free coast	<i>areacoast</i>
% Land area within 100 km of ice-free coast/navigable river	<i>areacoastriver</i>
Typical population density experienced by an individual (persons/km <sup>2</sup> )	<i>popdens</i>
% Land area in geographical tropics	<i>tropicland</i>



**Figures 1a (left) and 1b (right):** Eigenvectors for physical characteristics dataset, largest eigenvalue (left) and second-largest eigenvalue (right)

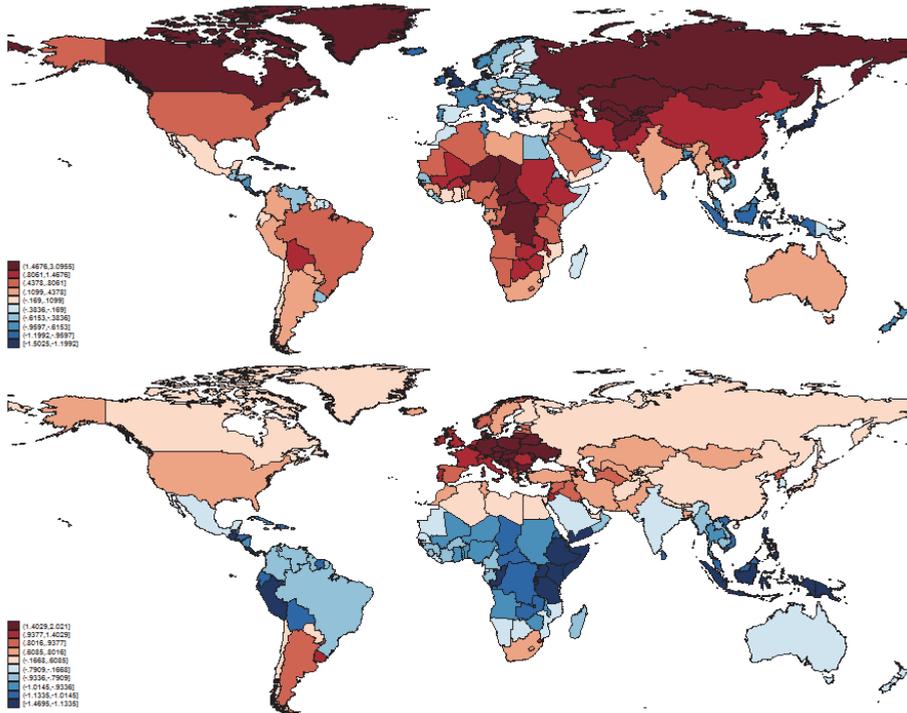
The average spatial correlation between the 13 variables in the dataset is modest to low (mean Moran coefficient of 0.04, with a standard deviation of 0.22). The maximum Moran coefficient observed between two variables in the dataset is 0.77. In terms of the Moran coefficients between the resulting summary variables (what would be called “scores” in regular principal components analysis), the three methods differ by rather much. The regular principal components analysis produces no spatial correlation at all (Moran coefficients -0.01 and 0.01). Wartenberg’s method produces Moran coefficients

<sup>7</sup> The variable names have been retained as in the original datasets (<http://www.cid.harvard.edu/-ciddata/geographydata.htm>)

of 0.38 and 0.78, respectively for the first and second largest eigenvalues. Interestingly, the Moran coefficient associated with the second-largest eigenvalue is larger than that of the largest eigenvalue. In other words, if one would be looking for a high degree of spatial correlation, using the second eigenvector would be better than using the first eigenvector. For our version of spatial principal components, the Moran coefficients are 0.83 and 0.50, respectively for the largest and second largest eigenvalue.

The profiles of the eigenvectors in Figures 1a and 1b show clear differences between the three methods. For the largest eigenvalue (left figure), regular principal components analysis puts most emphasis (highest “loadings”) on *popdens*, *coast*, *popcoast* and *area-coastriver* (positive loadings) and *popcoastriver*, *areacoast*, *tropicland* and *seariver* (negative loadings). Wartenberg’s method stresses *elev*, *tropicland*, *popdens*, *seariver* and *coast* (positive loadings) and *popcoast*, *popcoastriver*, *areacoast* and *areacoastriver* (negative loadings). Our method emphasizes *popdens*, *seariver*, *popcoast* and *popcoastriver* (positive loadings) and *tropicland*, *coastcent*, *searivercent* and *areacoastriver* (negative loadings). All in all, it is clear that the three methods are far from equivalent in terms of which variables should be weighted strongest.

This leads to a rather different picture of which countries are behind the spatial correlation that is observed in the dataset as a whole. Maps 1a and 1b provide an overview of these differences. The colours in the maps are an indication for the “component scores” (data vectors multiplied by the eigenvectors) for Wartenberg’s method and our method for spatial principal components. We do not document maps for any second largest eigenvalues, nor for regular principal components (which produces no spatial correlation). As could already be suspected from the loadings in Figure 1, the two maps produce rather different pictures. Wartenberg’s method shows high scores around the globe, but especially so in the two large countries that border on the North polar area (Canada and Russia). Large parts of Africa, the Americas and Asia also show high scores. Europe is the exception: here Wartenberg’s method finds mostly lower values. This is exactly opposite for our method, which produces very high values for a cluster of West- and Central-European countries, stretching into the Caspian Sea area. This method produces also consistently lower values around the equator (on all continents). As can be seen from the above maps, Map 1b is easier to categorize into sub groups of countries. For example, we see a clear tropical region. This gives us the advantage of a better spatial interpretation than is possible with results in Map 1a. Thus, we find it more advantageous to use our method.



**Map 1a (top) and 1b (bottom):** Component scores for Wartenberg's method (top) and our method (bottom), only largest eigenvalue.

## 4.2. Spatial Canonical Correlation Analysis

In order to apply the spatial canonical correlation analysis, we introduce a second dataset from Gallup et al. (1999). This refers to agriculture, and consists of two main categories of variables: suitability of land (for irrigation and general soil quality) and the percentage of cultivated land in a particular climate zone. The variables are listed in Table 2.

We undertake the spatial canonical correlation analysis for the physical characteristics dataset and the agriculture dataset. This means that we will be looking, for each of the two datasets, for weights that summarize the dataset in such a way that the spatial correlation (Moran coefficient) with the summary variable from the other dataset is maximized. We start by looking at the loadings (eigenvectors), and concentrate only on the eigenvectors that belong to the highest eigenvalue. The Moran coefficient that belongs to this is 0.79 (i.e., the eigenvalue is the square of this value, 0.62).

The loadings are in Figures 2a and 2b. We compare the loadings from the spatial canonical correlation analysis with those of the spatial principal components analysis (our method), to see whether any notable differences arise between the cases where we aim for maximizing spatial correlation within the dataset, or between the two datasets. For the spatial principal components analysis for the agriculture dataset, we do not document

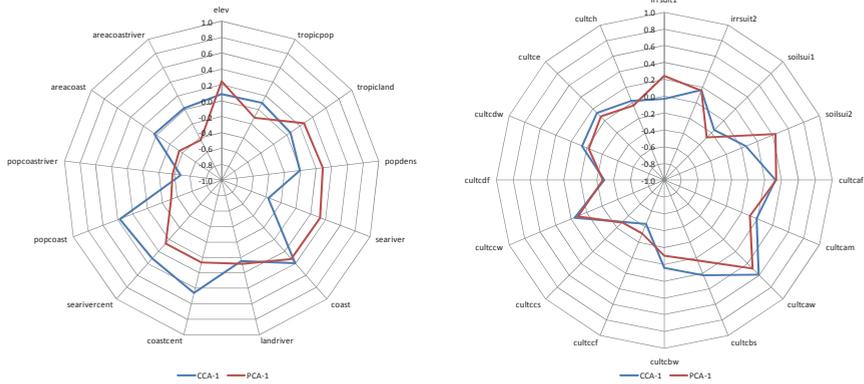
further details of this procedure, while for the physical characteristics dataset, these are the results from the previous section.

**Table 2. Variables in the agriculture dataset**

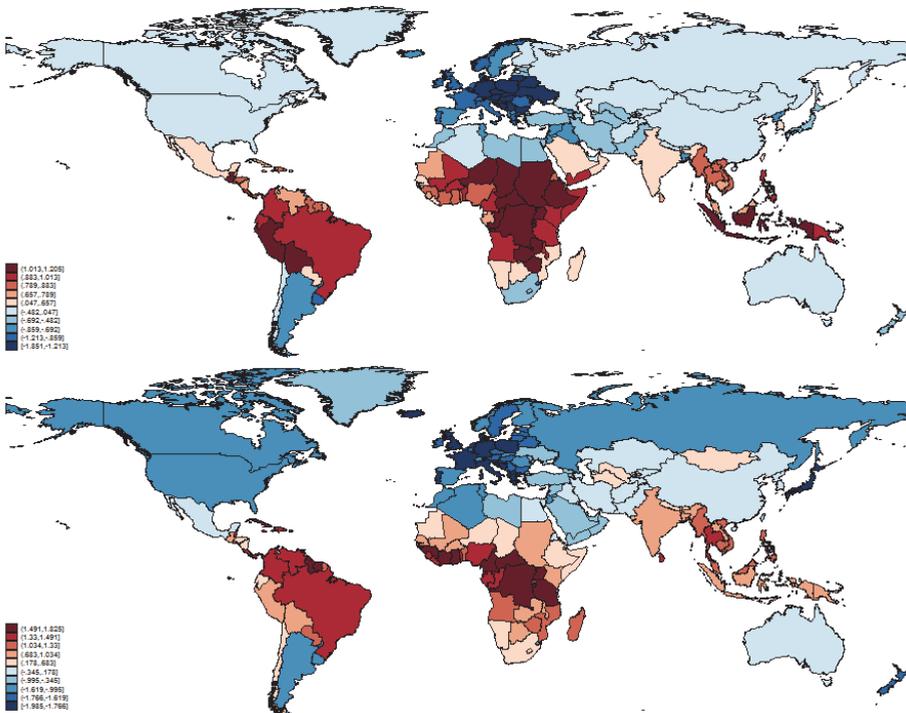
Description	Variable code
Mean irrigation suitability, very suitable (%)	<i>irrsuit1</i>
Mean irrigation suitability, moderately suitable (%)	<i>irrsuit2</i>
mean soil suitability 1, very suitable (%)	<i>soilsui1</i>
mean soil suitability 2, moderately suitable (%)	<i>soilsui2</i>
% cultivated land in Köppen-Geiger Af zone	<i>cultcaf</i>
% cultivated land in Köppen-Geiger Am zone	<i>cultcam</i>
% cultivated land in Köppen-Geiger Aw zone	<i>cultcaw</i>
% cultivated land in Köppen-Geiger Bs zone	<i>cultcbs</i>
% cultivated land in Köppen-Geiger Bw zone	<i>cultcbw</i>
% cultivated land in Köppen-Geiger Cf zone	<i>cultccf</i>
% cultivated land in Köppen-Geiger Cs zone	<i>cultccs</i>
% cultivated land in Köppen-Geiger Cw zone	<i>cultccw</i>
% cultivated land in Köppen-Geiger Df zone	<i>cultcdf</i>
% cultivated land in Köppen-Geiger Dw zone	<i>cultcdw</i>
% cultivated land in Köppen-Geiger E zone	<i>cultce</i>
% cultivated land in Köppen-Geiger H zone	<i>cultch</i>

For the physical characteristics dataset, we clearly see changes compared to the previous section. Compared to the principal components analysis, the canonical correlation analysis loads much higher on *coastcent*, *searivercent*, *popcoast*, *areacoast* and *areacoastriver*, and much lower on *seariver*. Note that these are all variables related to waterways. The differences for the other variables are less strong, but still there are differences for these as well (canonical correlation analysis is lower on *elev*, *tropicland* and *popdens*, but higher on *tropicpop*).

The differences between spatial canonical correlation analysis and spatial principal components seem to be somewhat less for the agriculture dataset. Here, many variables are remarkably close in Figure 1b, with *soilsui2* as a main exception, and *irrsuit1*, *cultcbs* and *cultcbw* as more minor exceptions.



**Figures 2a (left) and 2b (right):** Eigenvectors for physical characteristics dataset (left) and agriculture dataset (right), largest eigenvalue only.



**Map 2a (top) and 2b (bottom):** Component scores for spatial canonical correlation analysis between physical characteristics dataset (top) and agriculture dataset (bottom), only largest eigenvalue

Next, we look at the maps for the component scores for the two datasets, again only for the scores belonging to the largest eigenvalue. These maps are displayed in Maps 2a and 2b. The map at the top (2a) is for physical characteristics, and can therefore be compared to Map 1b above. The first thing that catches the eye in this comparison is that the cur-

rent map (2a) is almost a mirror image of the previous one (1b). Here (before), we have high (low) values around the equator, and low (high) values in West- and Central-Europe. There are small deviations from this comparison (e.g., Kazakhstan, the USA), but overall, this is a rather strong similarity, despite the differences in loadings in Figure 2a.

Naturally (because of the strong spatial correlation as indicated by the eigenvalue), the map for the agriculture dataset looks similar. Here, however, the areas with high (low) values are somewhat more concentrated within the larger area around the equator (e.g., the Congo and Uganda area rather than all of sub-Saharan Africa), or within Europe.

## 5. Conclusions

We have presented a method for spatial principal components analysis that has two important advantages over the method that Wartenberg (1985) proposed, and which has found its place in the toolbox of computational methods for spatial analysis (e.g., the *ade4* R package, see Jombart, undated). The first advantage is that, contrary to Wartenberg's method, our method has a clear and exact interpretation: it produces a summary measure of a dataset that itself has maximum spatial correlation. Thus, rather than working at an intuitive level, our method specifies exactly what is the goal of the weighting procedure that is derived using the eigenvalue decomposition of the Moran matrix. Further, this weighting procedure helps us see clear clusters of countries, motivated by a spatial narrative, one that makes empirical conclusions more exact than that of the Wartenberg method.

Second, by this goal of the analysis, an easy and intuitive link can be made to canonical correlation analysis. Our spatial canonical correlation analysis produces summary measures of two datasets (e.g., each measuring a different phenomenon), and these summary measures produce the maximum spatial correlation between them. This provides an alternative weighting scheme as compared to spatial principal components analysis.

The methods that we propose are computationally easy (with modern computers), and can easily be implemented in a variety of software packages. We have Matlab routines available on request.

We provided example applications of the methods and showed that the spatial principal components analysis may produce rather different results than Wartenberg's method. We also illustrated how spatial canonical correlation analysis may produce different results than spatial canonical correlation analysis. We hope that practitioners in the field of spatial statistical analysis will apply our methods to different problems, and thus show their usefulness.

## Appendix

### Our approach of a spatial canonical correlation analysis compared to standard CCA

We want to show in this appendix how our broad approach of maximizing equation (4) in the main text subject to the unitary variance of  $Xu$  and  $Yv$  relates to standard CCA. There is one main difference between standard CCA and our spatial variant: we introduce the spatial weights matrix  $W$ , (in the maximand, not in the constraints). In order to show the equivalence, we will first drop the spatial weights  $W$ , and show that in this case our approach and CCA are identical.

The starting point is our maximization problem, which can be written as the Lagrangean

$$L = (Xu)^T W (Yv) - \lambda_1 ((Xu)^T (Xu) - 1)/2 - \lambda_2 ((Yv)^T (Yv) - 1)/2. \quad (A1)$$

### The non-spatial version of our approach

Dropping  $W$  from the Lagrange function above yields the following alternative Lagrangean:

$$L = (Xu)^T (Yv) - \lambda_1 (u^T X^T Xu - 1) - \lambda_2 (v^T Y^T Yv - 1) \quad (A2)$$

The maximand is  $\text{Cov}(Xu, Yv)$ , where  $X$  and  $Y$  have the same number of rows but potentially a different number of columns. The constraints are identical to the Lagrangean (A1).

The first-order conditions are

$$\begin{aligned} L_u &= X^T Yv - \lambda_1 X^T Xu = 0, \\ L_v &= Y^T Xu - \lambda_2 Y^T Yv = 0. \end{aligned}$$

From the first-order conditions we get  $u = (X^T X)^{-1} X^T Yv / \lambda_1$  and  $v = (Y^T Y)^{-1} Y^T Xu / \lambda_2$ . Insertion into the first-order conditions yields

$$\begin{aligned} L_u &= X^T Y (Y^T Y)^{-1} Y^T Xu / \lambda_2 - \lambda_1 X^T Xu = 0, \\ L_v &= Y^T X (X^T X)^{-1} X^T Yv / \lambda_1 - \lambda_2 Y^T Yv = 0. \end{aligned}$$

Multiplying through by the Lagrange multiplier and  $(X^T X)^{-1}$  and  $(Y^T Y)^{-1}$  respectively from the left yields

$$\begin{aligned} L_u &= (X^T X)^{-1} X^T Y (Y^T Y)^{-1} Y^T Xu - \lambda_2 \lambda_1 u = 0, \\ L_v &= (Y^T Y)^{-1} Y^T X (X^T X)^{-1} X^T Yv - \lambda_1 \lambda_2 v = 0. \end{aligned}$$

For the eigenvectors belonging to the highest eigenvalues then the solution is  $\lambda_2 \lambda_1 u = (X^T X)^{-1} X^T Y (Y^T Y)^{-1} Y^T Xu = v^T (Y^T Y)^{-1} Y^T X (X^T X)^{-1} X^T Yv$ .

## Equivalence with CCA

Going back to the first set of first-order conditions above, we can show that the result is equivalent to CCA, but the above version is much simpler because it goes more directly to the eigenvectors.

Define  $e = (X^T X)^{1/2} u$  and  $f = (Y^T Y)^{1/2} v$ . Then,  $u = (X^T X)^{-1/2} e$  and  $v = (Y^T Y)^{-1/2} f$ .

Replacing  $u$  and  $v$  in the first set of first-order conditions yields

$$\begin{aligned} L_u &= X^T Y (Y^T Y)^{-1/2} f - \lambda_1 (X^T X)^{1/2} e = 0, \\ L_v &= Y^T X (X^T X)^{-1/2} e - \lambda_2 (Y^T Y)^{1/2} f = 0. \end{aligned}$$

From these first-order conditions we get  $e = (X^T X)^{-1/2} X^T Y (Y^T Y)^{-1/2} f / \lambda_1$  and  $f = (Y^T Y)^{-1/2} Y^T X (X^T X)^{-1/2} e / \lambda_2$ . Using these latter equations to replace  $f$  in the first and  $e$  in the second equation yields

$$\begin{aligned} L_u &= X^T Y (Y^T Y)^{-1/2} (Y^T Y)^{-1/2} Y^T X (X^T X)^{-1/2} e / \lambda_2 - \lambda_1 (X^T X)^{1/2} e = 0, \\ L_v &= Y^T X (X^T X)^{-1/2} (X^T X)^{-1/2} X^T Y (Y^T Y)^{-1/2} f / \lambda_1 - \lambda_2 (Y^T Y)^{1/2} f = 0. \end{aligned}$$

Multiplying through by the Lagrange multiplier and  $(X^T X)^{-1/2}$  and  $(Y^T Y)^{-1/2}$  respectively from the left yields

$$\begin{aligned} L_u &= (X^T X)^{-1/2} X^T Y (Y^T Y)^{-1/2} (Y^T Y)^{-1/2} Y^T X (X^T X)^{-1/2} e - \lambda_2 \lambda_1 e = 0 \\ L_v &= (Y^T Y)^{-1/2} Y^T X (X^T X)^{-1/2} (X^T X)^{-1/2} X^T Y (Y^T Y)^{-1/2} f - \lambda_1 \lambda_2 f = 0 \end{aligned}$$

Multiplying with  $e^T$  and  $f^T$  from the left yields and solving for  $\lambda_2 \lambda_1$  yields:

$$\begin{aligned} \lambda_2 \lambda_1 &= e^T (X^T X)^{-1/2} X^T Y (Y^T Y)^{-1/2} (Y^T Y)^{-1/2} Y^T X (X^T X)^{-1/2} e = \\ & f^T (Y^T Y)^{-1/2} Y^T X (X^T X)^{-1/2} (X^T X)^{-1/2} X^T Y (Y^T Y)^{-1/2} f \end{aligned}$$

Taking the eigenvectors and values and calculating  $u$  and  $v$  is exactly what CCA does (see Johnson and Wichern 2007, chap. 10), but here it has been obtained from our method. By implication, our method is equivalent to CCA if  $W$  is dropped.

## CHAPTER 3

# The Geographic Dimensions of Institutions

## Introduction

Explaining uneven economic development amongst countries has been a primary concern of empirical growth literature. Growth performance of countries is often explained in the context of convergence or divergence, given their wide-ranging ‘initial conditions.’ One such strand of literature that has taken centre stage in explaining these differences is new institutional economics (NIE). According to NIE, differences in institutional quality and governance structures<sup>8</sup>, explain growth differentials among countries<sup>9</sup> (Acemoglu et al.(2001, 2002, 2004), Engerman and Sokoloff (2005), Rodrik et al.(2004)).

Institutions refer to *formal rules* (constitutions, laws and regulations, political systems etc...) and *informal rules* (value systems, beliefs, social norms etc...) that humans use when interacting within a wide variety of repetitive and structured situations at multiple levels of analysis (North 2005, Ostrom 2005). Institutions are categorized into economic (property rights, presence of markets, regulatory structures etc...), political (constitutional arrangement) and social institutions (presence of health care markets, unemployment support systems, pension provision systems, and other informal social institutions such as social norms, religious norms etc...) (Bennedsen et al.2005).

Although there is no absolute definition for “good governance” most include the depth of democracy and public accountability, bureaucracy and judiciary, stability of property rights, corporate governance, financial institutions, social welfare and labour institutions (Chang 2002, Khan 2004). In other words, it is how the social and economic environment of a country is managed (World Bank 1992:1)<sup>10</sup>.

Studies in this area establish that there is a definite causal link between institutions and economic development directly and a link between geography and economic development via their influence on institutions (Rodrik et al.2004, Easterly and Levine 2003, etc.). These studies have helped establish a stylized fact that better institutions are strongly related to higher economic development. Most of these studies tend to use GDP per capita as the single most important measure of economic development. GDP per capita is no doubt an important indicator of economic development. However, most economists would agree that it does not help paint a complete picture of a country’s level of development. In this paper, we would like to use a broader set of indicators to measure the economic performance of a country, in order to re-examine the relationship between development, institutions and geography.

---

<sup>8</sup> Although, institutional quality, institutions and governance do not mean the same, they are often used interchangeably. In this paper too, we use them interchangeably.

<sup>9</sup> It is understood in NIE that the ‘quality’ of institutions is determined by how strongly property rights are protected and how “free” markets are, i.e., stronger property rights laws and “freer markets” imply higher quality of institutions (Chang 2011).

<sup>10</sup> See Section 2.1 for the formal definition.

Geographic determinism and international trade as the key drivers of growth in a country are the other two strands of literature that stand out. The former states that geographic differences (environmental differences) determine development differences in a country. It states that geography (climate, natural resources, disease ecology, etc.) determines not only the capacity of food production, but also labour productivity<sup>11</sup>, and trade policies of a country<sup>12</sup>. The latter states that international trade is the key driver of productivity and thus has a direct impact on the growth of a country<sup>13</sup>. Some studies have tried to evaluate which of these competing hypotheses gains ‘primacy’ over the others in explaining the ultimate driver of growth and development<sup>14</sup>.

Geography is often considered as the most exogenous factor that can be used in econometric studies. Geographic determinism primarily enters literature in the form of its ‘absolute location’. By this we mean the physical geographic characteristics of a country, like its climatic conditions, physical endowments etc., which determine the ability of a country to develop. Geography in this form is discussed as an ultimate determinant of economic development. In recent years however, with the emergence of literature on regional convergence, geography in the form of its ‘relative location’ has been gaining importance. ‘Relative location’ refers to the relative spatial characteristics of a country, namely, its neighboring countries and their influence on it (Abreu et al. 2004)<sup>15</sup>. In this paper, we would like to use both these concepts. In particular, we would like to see the impact of physical geography and institutions on economic growth, given the relative location of countries and their influence on each other. We do this with the help of a canonical correlation analysis, which allows us to correlate one set of variables against another set of variables.

The aim of this paper is to understand the interactions between institutions, economic performance, financial performance and the geography of a country from a spatial perspective. The analysis aims at understanding the underlying spatial patterns of institutions when summarized in combination with geography, economic performance or financial performance of countries. We do so by using a new method of spatial canonical correlation analysis, which is explained below. We describe these spatial patterns at a cross-country level, looking specifically at what factors in the two sets of data commonly result in spatial clusters. Rather than entering the debate on the primacy of absolute geography or institutions as determinants of economic performance, we want to see how

---

<sup>11</sup> The prevalence and incidence of diseases (like malaria, HIV) are much higher in tropical countries and thus burdensome not only in terms of costs incurred for treatment but also in its debilitating effect on labour, often reducing labour productivity (Bloom and Sachs 1998, Sachs 2001, Gallup and Sachs 2001 etc)

<sup>12</sup> Landlocked economies tend to have lesser trade openness and higher transport costs. Countries that have a coast line tend to have lower transport costs as well as policies that encourage international trade, thus allowing them more growth opportunities (Mellinger, Sachs and Gallup 1999)

<sup>13</sup> Frankel and Romer (1999), Dollar and Kraay (2003) etc.

<sup>14</sup> These include Rodrik et al. (2004), Easterly and Levine (2003), Rodrik (2003) etc.

<sup>15</sup> ‘Absolute location’ and ‘relative location’ are often referred to as ‘first nature’ geography and ‘second nature’ geography of a country.

each of these phenomena interacts with the other in a broader context. This paper does not try to address the causality between institutions and economic development but rather provide a descriptive overview of the spatial patterns.

The rest of the paper is organized as follows. Section 2 gives a brief overview of related literature. Section 3 describes the empirical methodology and the data used. Section 4 explains the results and the last section concludes the paper.

## 1. Related Literature

### 2.1. Institutions and Economic Development

Literature on institutions and economic growth and development has regained popularity over the last two decades. The predominant discourse in this literature states that institutions that ensure freer working of markets (read as the least amount of government intervention) as well as strong property rights protection provide the best environment for economic development (Chang 2011). According to the current literature, “good institutions” create an environment for providing the “right” incentives to agents in the economy. These incentives lead to learning and innovation, investment in human and physical capital as well as capabilities. These factors lead to higher growth rates in an economy. An improper or a bad institutional set up, on the other hand, does not incentivize such activities. Instead, it leads to corruption, rent-seeking and other non-productive or growth repressing activities.

There are numerous studies based on the institutions hypothesis. “Core” papers in this area of research include Acemoglu et al. (2002, 2005), Rodrik et al. (2004) Easterly and Levine (2003), Hall and Jones (1999) etc., among others. All these studies examine how various institutions effect economic growth and development outcomes, concluding that *institutions matter*.

Acemoglu, Johnson and Robinson (2002), use settler mortality rates of the colonizers in the colonies as an instrument for institutional quality during the colonial period. They hypothesize that Europeans took their institutions to ‘settler colonies’, while they set up different ‘extractive’ institutions in colonies where the prime intention was extraction of wealth. Thus, colonies where they settled got better institutions and those where they had an extraction policy ended up with bad institutions. These historical conditions further influenced the quality and type of current institutions. They find a strong and significant influence of historical institutions from the colonial era on current institutions as well as of current institutions on per capita income. Hall and Jones (1999), use “distance from equator” and the “extent of to which Western languages are spoken in the country” as instruments to find a strong correlation between output per worker with respect to social infrastructure (which is an average of an institutions and openness of economy index). Kauffman et al. (1999) also use “extent of to which Western languages are spoken in the country” to find a positive and significant impact of the quality of institutions on growth. Other cross-country studies like Mauro (1995), Easterly and Levine (2003), Dollar and

Kraay (2003) also support the hypothesis that institutions generate development. In fact, there are but a few studies that do not find any significant effect<sup>16</sup>.

The proponents of the institutional hypothesis agree that institutions have evolved endogenously, depending on the country's geographic conditions, among other factors. Examples of such studies include Hall and Jones (1999), Acemoglu et al. (2001), Rodrik et al. (2004). These studies have made empirical efforts to explain the possible linkage between geography and institutions. Hall and Jones (1999) hypothesize that Western Europeans preferred to live in places that resembled their own, and thus indirectly use geography as an explanatory variable. Engerman and Sokoloff (1997) suggest that the climatic conditions either 'lent themselves' to plantation agriculture or small farm agriculture. This in turn meant that the plantation economies led to less democratic power relations and small farms, by empowering a larger section of society led to more democratic institutions and thus to more development. Rodrik et al. (2004) and Easterly and Levine (2003) find that geography influences economic growth development only via its impact on institutions.

Much of this literature has not been sensitive towards the individual development experiences of countries. Most of these theories "have been rooted in the historical and social experiences of a few Western industrialized countries" (Brohman 1995). Policy prescriptions of growth have thus been a set of generalized recommendations, irrespective of their applicability to culture or the varying historical conditions of each country (that are most often very different from the Western European conditions). While it is important to understand growth at a more general level, policies made towards encouraging growth and development need to be made at the country level, catering to the specific conditions of the economy. The failure of the Washington Consensus has further proved that we need to look beyond the "one size fits all" approach.

## 2.2 Geography and Economic Development

The geography hypothesis is based on the fact that not all the areas in the world are endowed in the same fashion. The geography of a country determines characteristics such as soil quality which directly affects agricultural productivity, natural resources which contribute to wealth as well as provide the raw material base for industrial activity, topography which contributes to costs of transportation, climate that affects the productivity of workers, and the disease environment which also directly affects the productivity of workers (Acemoglu 2009). No access to rivers, as well as land-locked geography, affects the growth of the country negatively. The proponents of this view argue that during and after colonization the poor countries have remained poor. There has been no reversal of fortunes and what has remained the same is the geographical position of these countries (Sachs (2001) and Bloom and Sachs (1998)). Thus, it is primarily geography, which determines the economic situation of the country.

---

<sup>16</sup> For example, Dollar and Kraay (2003) find a positive and insignificant effect between "rule of law" and growth, while Sala-i-Martin (1997 a) finds ethno-linguistic fractionalization on growth as insignificant.

There have been many empirical studies carried out to determine the importance of geography as a determinant of growth. The latitude of the country is a popular variable in testing the geography-development hypothesis. Sala-i-Martin (1997 a, b), Bloom and Sachs (1998), Easterly and Levine (2001) are among some of those studies that find a positive and significant impact of the latitudinal position of a country on its growth levels. The numbers of frost days in a year, availability of arable land, the length of the coastline, rainfall, and temperature have been other variables used to study the effect of geographical position of the country on its growth rates. Some of the papers that test these variables empirically are Bloom and Sachs (2003), Masters and Sachs (2001) and Bloom et al. (2003). All these find a positive and significant effect of the first three variables on growth, while the last two variables are found to have a negative and significant effect. That is to say, lesser and more unpredictable rainfalls as well as high temperatures directly affect the productivity as well as output negatively. Landlocked economies have been tested to find a negative impact on growth (Easterly and Levine (2001), Sachs and Gallup, (1999)). Easterly and Levine (2003), and Sachs (2003) also use measures such as the proportion of the population at the risk of malaria, life expectancy at birth, infant mortality rate etc., to find a significant and negative impact of disease ecology on growth. McArthur and Sachs (2001), concludes that institutions and geographically-related variables affect the GNP per capita of countries.

The results of the empirical studies mentioned above find that the geography does indeed have an impact on the country's growth statistics. As mentioned earlier the geography-development hypothesis is mainly based on the 'absolute location' of a country. That is, they consider those characteristics of a country's geography which are exogenous, time-invariant (except for some disease variables) and are not influenced by its 'relative location' vis-à-vis its neighbours and their socio-political and economic conditions.

### **2.3. Spatial econometrics and Economic Development**

The impact of the 'relative location' of a country on the quality of institutions and thus economic growth has been given relatively less attention. How does the institutional quality of a country get affected by the institutional quality of its neighbours? A large literature on spatial dependence in the field of economic growth can be found in the regional convergence literature. Regional income convergence on intra-national scale, as well as regional scale, has resulted in strong consensus on income convergence among these geographical units. The theories of technology diffusion and factor mobility have strong spatial components. For example, Verspagen (2010), López-Bazo et al. (1998, 2004) study the technology diffusion and growth amongst European regions. Similarly Rey and Montouri (1998) study spatial patterns in income across the states of USA. Other studies in this literature include Rey and Montouri (1998), Quah (1997), Cheshire and Magrini (2009) etc.

Spatial dependence with respect to institutions and governance has been addressed in a couple of studies. Ward and Gleditsch (2008) study the spatial spillover effect of democracy at a cross-country level. They find southern, and western Africa, as well as India, are unusual situations where in democracy in their country is not explained by the situa-

tion in their neighboring countries. Examining if the governance structures in an economy are influenced by their neighboring countries could help unbundle institutions in a more effective way. They find a positive impact of democracy of neighbouring countries. Leeson and Dean (2009), Sobel and Leeson (2007), etc., also study the effects of democracy on neighbouring countries, with a spatial or a temporal lag (Hosseini and Kaneko 2012).

### 3. Data and Methodology

Institutional quality, geography, economic performance and the financial structure of a country are all multi-dimensional in nature. Using single variables to describe these phenomena can bias understanding of how they influence each other. In this paper, we want to better understand how the geography of a country (its ‘absolute location’), its economic performance and its financial structure, each in turn, are related to the institutional set up of a country (given its relative location). In order to understand the role of ‘relative geography’ of countries, we introduce an alternate method of *spatial principal component analysis* and *spatial canonical correlation analysis*. We would like to study the dependence of one latent variable on the other, given the relative location of the observations, i.e., we would like to account for the spatial dependence of observations in one country on its neighbours<sup>17</sup>.

Similar to the Principal Component Analysis (PCA), canonical correlation analysis (CCA) is a dimensionality reduction method. PCA transforms a set of probably correlated variables into a smaller set of uncorrelated variables. This is done by constructing a new set of latent variables that are orthogonal (uncorrelated) to each other. In other words, the new variables are weighted linear combinations of the original variables, which maximize the variance of the new latent variables in the data set, called principal components. While the PCA does this for one set of variables, the CCA finds the relationship between two sets of multiple variables X and Y, for which we obtain two sets of latent vectors. These are called canonical components. The weights are obtained in such a way that they not only reduce the number of variables in each data set, but also maximize the correlation between the corresponding new components in each dataset (i.e., the first component in data set 1 is maximally correlated to the first component in data set 2 and so on. Each pair of components is together called a canonical function)<sup>18</sup>. The residual variance is explained by each successive canonical function that is, the first function explains the most amount of variance in the data and consecutive ones explain the maximum of residual variance. The highest number of components you can have is determined by the number of original variables in the smaller data set.

The analysis results in following results for PCA and CCA respectively: (i) *factor loadings* (or *canonical loadings*) for both sets of data (X and Y), for each common pattern

---

<sup>17</sup> For a detailed explanation on the methodology used see (Bhupatiraju et al.2013)

<sup>18</sup> This is different from the PCA in that the correlations are found within one set of data and not the inter-correlations between two sets of data.

that exists in the two sets (for each canonical function). These weights correspond to each of the variables in the data sets, determining how they weigh in on the identified common pattern. With the factor loadings we can identify the structural pattern underlying the data (Clark 1975). These loadings are obtained from the correlation matrix, by the means of Eigen value decomposition. (ii) *Factor scores* (or *canonical scores*) for each of the data sets (X and Y). These scores explain how each country fares in the common pattern. This, is therefore, a good indication of the spatial patterns underlying the data. The scores are obtained by pre-multiplying the loadings with the standardized raw data. (iii) *Canonical correlation coefficient (Moran coefficient)* is the correlation coefficient between the latent variables of both sets of data (for each canonical function). It can be understood as the Pearson r coefficient between the two latent variables. The spatial correlation between the two factor score vectors is equal to the square root of the corresponding Eigen values.

We want to maximize the *correlation coefficient* obtained from the PCA and CCA to explain a maximized overall spatial pattern. We call this the *spatial correlation coefficient*. In order to do this, we introduce spatial weights (a *connectivity matrix*) that allow us to define the ‘relative location’ of the countries in the data sets<sup>19</sup>. We first calculate the *spatial lag* of one set of variables. A *spatial lag* of observation  $i$  of variable  $y$  ( $y_i$ ) is the linear combination of values of its neighboring countries and is given by  $\sum_j^n w_{ij}y_j$ , where  $w_{ij}$  is the spatial weight matrix. This is the weighted average of the observation  $i$ 's neighbouring countries ( $j$ ). In other words, a spatial lag introduces a diffusion process such that the variables in the data set are influenced by their neighbours. This method decomposes a spatially weighted matrix, in order to give it an explicit spatial perspective. Bhupatiraju et al. (2013) explains this methodology in detail. We perform a spatial PCA on the institutional data in order to reduce the number of institutional variables, which we further use for a spatial CCA along with the other data sets.

In order to obtain the spatial lag, we choose a set of geographic distance weights for the connectivity matrix. We use an exponential decay to obtain the weights matrix, given by the formula,  $w_{ij} = e^{-0.0015d_{ij}}$  where  $w_{ij}$  is the spatial weight between countries  $i$  and  $j$ , and  $d_{ij}$  is the bilateral distance between the centroids of the two countries. The data of bilateral distances are taken from the CEPII gravity dataset<sup>20</sup>. This specific distance decay has been chosen after having experimented with a few other weights, the details of which are presented in Appendix A. It is an arbitrary choice, but reflects a rapid decay of weights given the distance between the countries. This ( $n \times n$ ) matrix is row-standardized so as to take into account that all countries are -at some distance- neighbors of each other and, therefore, is not symmetric. It is also important to note that the weights are non-stochastic and are exogenous to the model.

---

<sup>19</sup> For a detailed explanation refer to Wartenberg (1985)

<sup>20</sup> Bilateral distances and common (official) language come from the CEPII distance database (<http://www.cepii.fr/anglaisgraph/bdd/distances.htm>).

The analysis is based on four datasets, which we will briefly present. The first dataset Institutional Profiles Database, 2009 version (IPD)<sup>21</sup> provides us with the data on institutional characteristics of 123 countries. The IPD data is divided into four institutional sectors and nine institutional functions. The categories will be presented in Appendix B. For our analysis we use the two-digit aggregation level which has 93 variables.

The data on physical geographic factors of countries are taken from the Gallup, Mellinger and Sachs geography dataset<sup>22</sup>. The dataset contains information on infectious diseases, general measures of geography and agricultural measures. This particular data set has been divided into two different sets for our analysis, one describes the physical geographic characteristics of countries (GEOAREA) and the second describes the soil quality related variables (GEOSOIL), mainly based on the Köppen-Geiger climate zones. The information on infectious diseases has been dropped out of this analysis (since they are time-varying and are affected by policy and is not necessarily exogenous as noted in the previous section).

The data on economic variables are taken from the Penn World Table<sup>23</sup> (PWT). We use a broader set of economic indicators to measure economic performance. These take into account not only the size of the economy (the initial level of GDP) and how fast it is growing (average annual growth rate of GDP per capita over the period of 2000-2007), but also some indicators of its expenditures which include investment, government spending, and consumption. We also include an indicator of the economy's trade openness as well as the log of population all using the definitions as in PWT. While GDP per capita is generally used as the measure of economic development/performance, we feel that these additional variables allow us to measure different dimensions of a country's economic situation. By including other aspects of the economy that leads to higher levels of growth in a country, the aim is to go beyond the simple measure of GDP per capita as a measure of economic development/performance.

The financial data is obtained from the Financial Structure database distributed by the World Bank<sup>24</sup> (FINANCE). This is a database of 31 indicators that describe the financial development and structure of countries. We use data for the year 2009. The variables in the data measure the "size, activity and efficiency of financial intermediaries and markets" (Beck et al. 2000). Variables that measure the size and activity of non-bank financial institutions (insurance, private bond market and stock markets), international debt and remittances, as well as liquid liabilities of countries, have missing data for our set of countries. Due to this lack of data, we have dropped these 12 variables from the dataset<sup>25</sup>. We use 19 variables that give details of the banking sector across countries. The details of these variables are given in Appendix C.

---

<sup>21</sup> <http://www.cepii.fr/anglaisgraph/bdd/institutions.htm>

<sup>22</sup> <http://www.cid.harvard.edu/ciddata/geographydata.htm>

<sup>23</sup> We use PWT (7) of the table.

<sup>24</sup> Thorsten Beck, Asli Demirgüç-Kunt and Ross Levine, (2000), "A New Database on Financial Development and Structure," World Bank Economic Review 14, 597-605.

<sup>25</sup> These are: other financial institutions assets / gdp (ofagdp), liquid liabilities (in mil. 2000 usd) (ll\_usd), life insurance premium volume / gdp (inslife), non-life insurance premium volume / gdp (insnonlife), stock market

## 4. Empirical Results

### 4.1. First stage results

The IPD database contains 367 “elementary items” that are aggregated to 133 “three digit level” variables and further aggregated to 93 “two digit level” variables. For our analysis, we use the two-digit aggregation level which has 93 variables. Because there are very few degrees of freedom when we use the 93 variables for 118 countries, we would like to further aggregate these before our second step of the analysis. We use a pre-defined categorization (of the 4 sectors) to obtain new latent variables for each of the sectors. Sector A which constitutes institutional variables related to public institutions, and civil society contains 36 two-digit level variables. Sector B which constitutes variables related to the goods and services markets contains 28 variables. Sector C constitutes 13 variables describing the capital market, and Sector D constitutes 16 variables describing institutional variables related to the labour market. We first do a spatial PCA on each of sectors of IPD. We select the new latent variables by looking at the scree plot of the Eigen values. Out of the components we obtain; we choose to work with 12 components (four from Sector A, three from Sector B, two from Sector C and three from Sector D). We discuss the results for some of these components before proceeding to the next step of the analysis.<sup>26</sup>

We call the first PC of Sector A, A1. This component gives the highest positive weights to ‘Social Inclusion’ (A93), ‘democracy, legality & freedom (A10)’, and ‘political proximity to an EU country’. ‘Change and innovation’ and ‘political proximity to Japan’ have strong negative loadings. This component in general can be interpreted as representing a ‘*socially inclusive democracy*’ with political proximity to an EU country. In other words, we can see that countries that function similar to the non-Anglo Saxon Western European economies obtain a higher score than others. It is interesting to note the Algeria and Libya also get a high score along with Canada. These are the only countries, not in the geographical region of Western Europe, which have a high score.

The second PC of Sector A is called A2 here. The variables that load high and positively in the second component are ‘Change and Innovation’, ‘Security of transactions & contracts’, and ‘political proximity to Japan’. ‘Institutional capacity’, ‘domestic public security & control of violence’ and ‘emulation of neighboring countries’ are some other variables that have relatively high weights. ‘Governance of public administration & the justice system’<sup>27</sup>, ‘government capacity to reform’, ‘autonomy in operation & creation of

---

capitalization / gdp (stmktcap), stock market total value traded / gdp (stvaltraded), stock market turnover ratio (stturnover), no. of listed companies per 10k population (listco\_pc), private bond market capitalization/gdp(prbond), public bond market capitalization / gdp (pubbond), international debt issues/gdp (intldebt), loans from non-resident banks (net) / gdp (intldebtnet), remittance inflows / gdp, (remit).

<sup>26</sup> The loadings for the first-stage are given in the Appendix D.

<sup>27</sup> Defined as the efficient public administration (which includes transparency in public accounts, economic policy and public procurement contracts, effective control of corruption) and an independent justice system

organizations' and 'decentralization' have strong negative loadings. The eastern part of the world has the highest scores in this component, along with Northern European countries. Africa and Latin America have the lowest scores along with Mediterranean Europe. This component can be interpreted as one that stresses on '*society's responsiveness to change and innovation*'.

The third PC of Sector A (A3) has high and positive loadings on 'change & innovation', 'governance of public administration & the justice system', 'domestic public security & control of violence', 'traditional solidarity', 'political proximity to China', 'social inclusion', and 'cooperative behavior in society'. It has negative loadings on 'government capacity to reform' (which has the strongest loading), 'national cohesion', 'decentralization', 'subsidies on commodities' and 'democracy, legality & freedom'. This component could be interpreted as stressing on the "*strong presence of the government*".

The last component that we consider from Sector A is A4. This constitutes of 'security of transactions & contract', 'capacity of state to co-ordinate stakeholders', 'institutional capacity', 'autonomy in creation and operation of organizations', and 'decentralization' which have positive loadings. Those variables that load negatively include 'social inclusion', 'political proximity to an EU country', 'governance of public administration & the justice system', 'control of state violence by NGO's', 'strategic capacities' as well as 'subsidies on commodities'. While 'government capacity to reform is the strongest negative loading on A3, 'political proximity to an EU country' along with 'social inclusion' is the strongest negative loadings in A4 and 'security of transactions' has the highest positive loading in A4. Although it is hard to give a straight forward meaning to the fourth component, it gives the highest scores to economies that guarantee a security of transactions but are not necessarily socially inclusive. We notice a strong Mediterranean cluster. This component could be interpreted as stressing on the '*security of contracts*'.

From Sector B, that has variables representing the goods and services market, we take three components. The first one, B1 loads positive on variables such as 'land tenure-security of ownership', 'joint ventures', 'technical environment', 'privatizations', 'nationalizations' and 'land tenure- demand for land'. It loads strongly negative on 'non-national access to land', 'rural land tenure- traditional property rights', 'consideration of public interest in government- business relationships' as well as 'government recognition of various land tenure rights systems'. From this, it is obvious that the most important aspect of the goods & services market pertains to '*security and enforcement of property rights laws*' in a country. From the scores, we find that the developed countries have the highest scores. Kazakhstan and Turkey also obtain a high score and countries in Sub-Saharan Africa get the lowest scores.

In the second component of Sector B, B2, the 'importance of economic zones', 'technical environment', 'land tenure- security of ownership', 'considerations of public interest in government-business relations', 'shareholder weights – weight of the government' and 'openness to business' have positive weights. 'Non-national access to land', 'infor-

---

(this includes the effectiveness of the fiscal system, justice system and of urban governance.) (de Crombrughe, Farla, Meisel, de Neubourg, Ould Aoudia, Szirmai (2009)

mation on shareholders', 'ease of starting a business', 'density of sub-contracting relations', 'public aid for R&D', 'rural land tenure- traditional property' and 'land tenure-development policies' load negatively. From the variables that load high -either positively or negatively- we see that the most important functions in this component are those pertaining to '*strong presence of government regulation and active economic policies (in the goods and services market)*'. This component could be understood as representing strong regulation by the government and having active economic policies pertaining to economic zones. Countries that score high on this factor include the Middle East, Central and South Asia as well a few African countries.

The last component of Sector B, B3, loads high and positive on 'privatizations', 'information on share-holders', 'land tenure- demand for land', 'free movement of people and information', 'shareholders- weight of government' and 'land tenure- security of ownership'. Variables that load negatively on the component include 'competition on the good & services market', 'importance of economic zones', 'governance of natural resources', 'rural land tenure- public property' as well as 'consideration of public interest in government-business relations'. Although it is hard to label this component based on the loadings, we find that the Americas along with Southern Africa and China get the lowest scores. Thus we label this component 'strong competition in the goods market'

We choose to retain only two components from Sector C. C1 constitutes variables like 'regulation of competition in banking', 'importance of venture capital', 'monitoring and auditing in banking', 'financial openness', 'nationalization of the financial sector' as well as 'freedom in allocation of loans' which load positively. 'Micro lending', 'financial information', 'sovereign wealth fund policy' and 'competition within the banking system' have a negative weight on the component. This component stresses on a '*well developed free and open financial sector*'.

C2, the second component of Sector C, is almost a mirror image of the first component<sup>28</sup>. Of those variables that load positively in C1, 'regulation of competition in banking', 'importance of venture capital', 'financial openness' and 'freedom of allocation of loans' load negatively in C2. Of those variables that load negatively in C1, we find that 'micro lending', 'sovereign wealth fund policy' and 'competition within the banking system' show up positively in C2, along with 'competence of bank executives'. This component stresses on a financial sector that is more or less controlled and regulated by the government, yet has a well-developed structure. We refer to this component as '*government regulated financial sector*'. While the first factor emphasizes openness of the financial sector, the second one emphasizes strong government control over the financial sector. It is interesting to note that in both factors most of Africa gets low scores while Australia and Canada get high scores. While 'monitoring and auditing in banking' influences the first component, competence of bank executives influence the second.

From Sector D, we retain three components. The first one, D1, loads positively high on 'freedom of association & trade union pluralism', and 'distribution of income' and negatively on 'low incidence of child labour', 'social mobility', 'management of labour',

---

<sup>28</sup> Since the two latent variables have to be orthogonal.

‘weak employment contract rigidity’ as well as ‘retraining and re-skilling measures’. Lack of social mobility and incidence of child labour clearly dominate this factor. This component stresses on a ‘*non-meritocratic weak labour market*’ where there is discrimination based on gender, ethnicity etc., and one in which social mobility is based on connections rather than on merit. We find that most African countries, along with Peru, Bolivia, Paraguay, India and Nepal score the highest.

The second component D2 loads positively on ‘low incidence of child labour’, ‘weak employment contract rigidity’, ‘social mobility: young higher education graduates’, ‘quality of the supply of public goods’, ‘strikes and wage bargaining at the individual level’. It loads negatively on ‘freedom of association & trade union pluralism’, ‘flexibility of the labour market’, ‘adaptive education system’ and ‘management of labour’. This component stresses on a ‘*flexible and adaptive labour market*’.

The last component from this sector, D3 loads high and positive on ‘low incidence of child labour’, ‘strikes’, ‘management of labour’, ‘weak segmentation of the labour market’, and ‘freedom of association & trade union pluralism’ and negatively on ‘quality of supply of public goods’, ‘social mobility’, ‘distribution of income’ and ‘social mobility: young higher education graduates’. This component stresses on a ‘*weakly segmented labour market*’.

Looking at these factor loadings and factor scores, we find that among the first components, A1, B1 and C1 move in the same direction with very similar scores for each country while D1 moves in the opposite direction. Using these 12 components from the IPD data set which we use as indicators of institutional quality, we now perform the spatial canonical correlation analysis along with a set of 7 economic variables (PWT). In the analysis that follows these 12 components will be called the IPD variables.

**Table 1: Labels for the first stage IPD components<sup>29</sup>**

Description	Variable code
‘socially inclusive democracy’ (with political proximity to an EU country)	A1
‘society’s responsiveness to change and innovation’ (with political proximity to Japan)	A2
‘strong presence of the government’	A3
‘security of contracts’	A4
‘security and enforcement of property rights laws’	B1
‘strong government regulation and active economic policies’ (in the goods and services market)	B2
‘strong competition in the good market’	B3
‘well developed free and open financial sector’	C1
‘government regulated financial sector’	C2
‘non-meritocratic weak labour market’	D1
‘flexible and adaptive labour market’	D2
‘weakly segmented labour market’	D3

<sup>29</sup> These labels are used to refer to each of the components. It should be noted however that these labels are not all encompassing in describing the details of the components. The reader might want to look at the details of the loadings provided in the Appendix.

## 4.2. Second stage results

Below we look the resulting correlations of IPD with the other data sets of economic variables (PWT), financial variables (FINANCE), and geographic variables of physical factors (GEOAREA) and soil quality (GEOSOIL). We start by looking at the correlations between institutional variables and the economic variables (IPD-PWT).

### A. Institutions and Economic Development (IPD-PWT):

We extract three components and analyze them here. The scores of the first canonical component from IPD and PWT are plotted in Figure 1a. The spatial correlation is 0.77. There are clear clusters of African countries, most of which are in the third quadrant, scoring low on both institutional and economic development variables. We also find most of the Asian countries around the origin, along with the Middle Eastern countries. In the first quadrant and close to the origin we find the Latin American countries and further up, scoring high on both institutional and economic variables we find the developed countries.

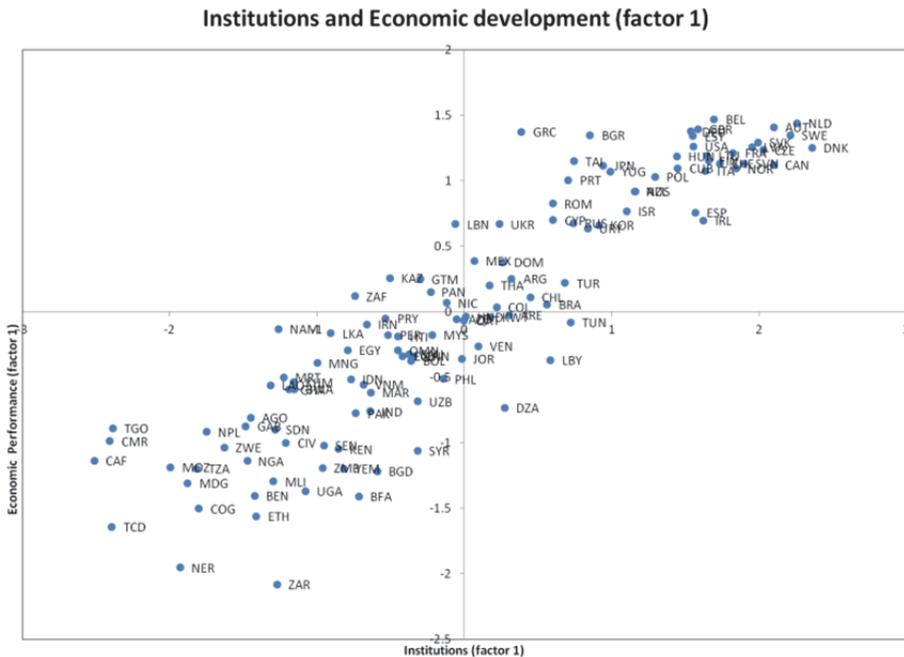


Figure 1a: Spatial correlation between institutions and economic performance (factor 1)

We can use the loadings to understand characteristics underlying the factors. On the IPD side, we find that '*socially inclusive democracy*' (A1), '*Society's responsiveness to change and innovation*' (A2) and '*enforcement of property rights laws*' (B1) load positively while '*strong presence of the government*' (A3) and '*non-meritocratic weak labour market*' (D1) load strongly negative. For the PWT data the first component loads positively on all variables except population growth rate which loads strongly negative. It loads strongly positive on the initial level of GDP, as well as on consumption. Growth rates, openness of the economy and investment load positive but low. This component can be interpreted as representing the level of economic development of countries. It captures rich countries which have a high level of initial GDP, high level of consumption and positive but not high growth rates.

The first canonical function thus emphasizes the relationship between proper functioning of political institutions, relatively high level of freedom and public rights, legality and legitimacy, the State's capacity to bring about a convergence of interests, authorities' strategic vision and high capacity to absorb new technology and high initial level of GDP, as well as high consumption in a country. The first components seem to capture the well-known and widely accepted correlation between institutions and level of GDP per capita. The relative position of the country clearly has an impact on its score.

Figure 1b shows the factor scores of institutional quality for the 118 countries. In general, there is a high level of clustering. The cluster with the highest scores is the developed country cluster. The cluster with the second highest set of scores is Latin America, Eastern Europe and Russia as well as some countries in North Africa (which could be due to their loadings on the variable 'political proximity to the EU'). The next cluster constitutes the Middle East, South and South East Asia and four Latin American countries. The least institutional scores are given to Sub-Saharan Africa. From a strictly spatial perspective, Australia and New Zealand are the outliers. Figure 1c shows the scores of economic performance which more or less corresponds to the institutional variable scores. The maps show that countries like Australia, New Zealand, the Indian sub-continent and Northern Africa have better institutions relative to their economic performance. Countries like Japan and South Africa have a better economic performance given relative to their institutional set up. In other words, the institutional scores capture the western European model of institutions for this function.

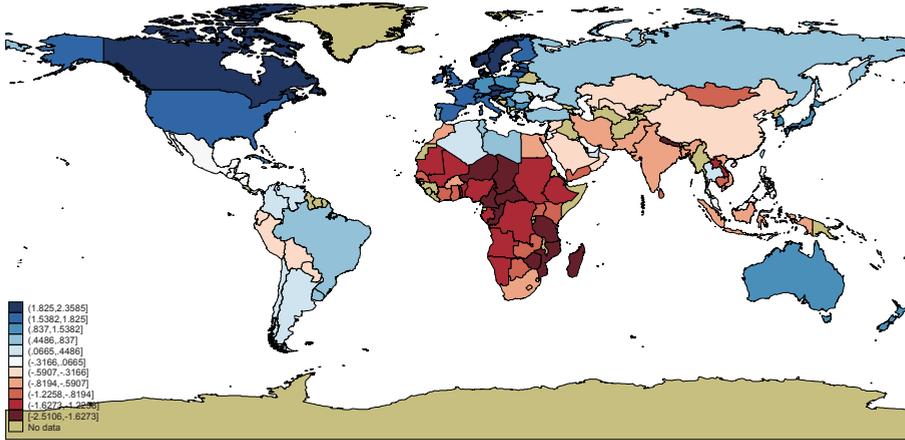


Figure 1b: IPD scores Factor 1 (IPD-PWT)

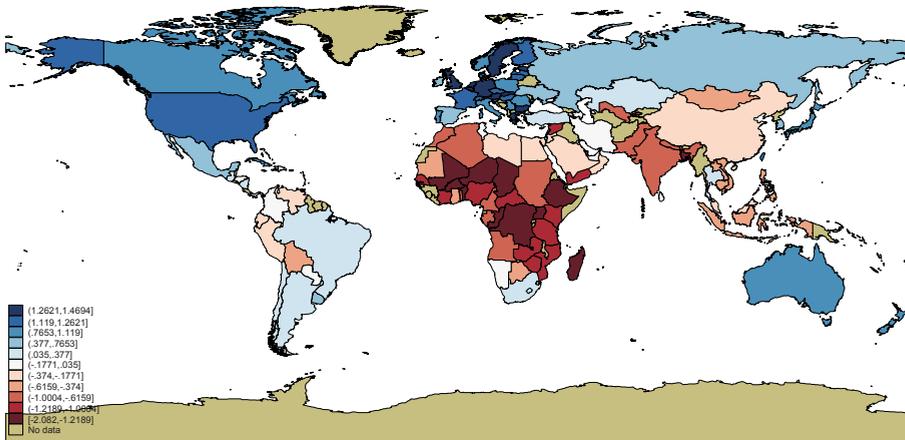
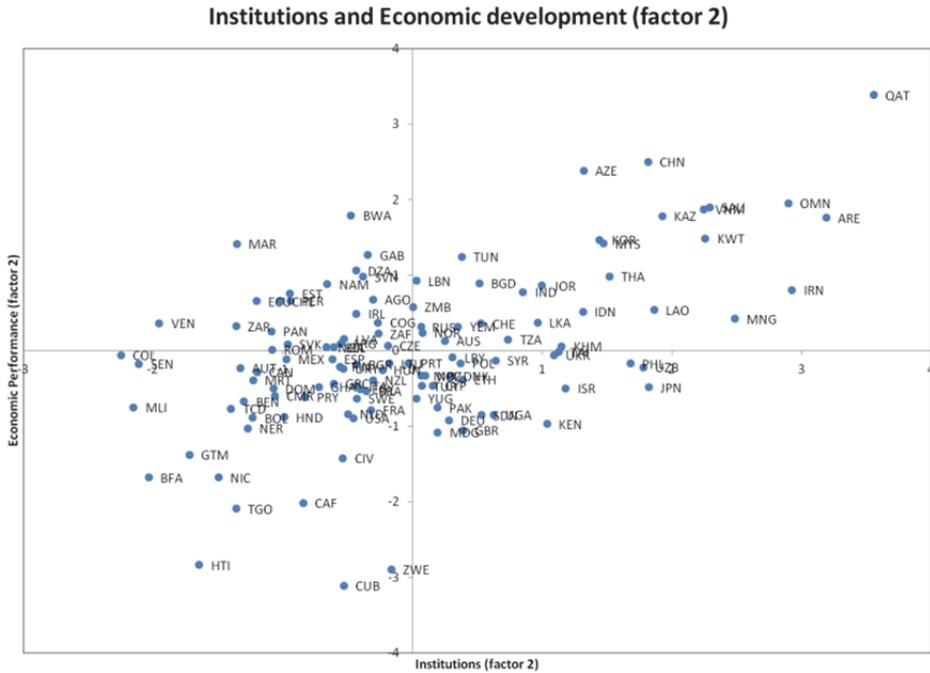


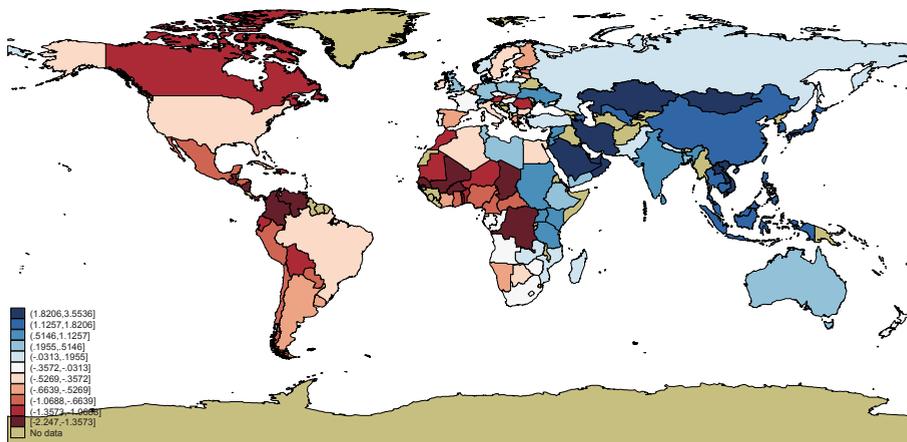
Figure 1c: PWT scores Factor 1 (IPD\_PWT)

The spatial correlation for the second canonical function (0.48) is significantly smaller than the first. On the IPD side, *'security and enforcement of property rights laws'* (B1), *'strong government regulation and active economic policies'* (B2) and *'flexible and adaptive labour market'* (D2) are among those that load positively and *'socially inclusive democracy'* (A1) loads strongly negative. The PWT data loads strongly positive on growth rate and investment. It loads negatively on the initial level of GDP, consumption and government spending. Openness is positive but with a relatively low weight. Initial level of GDP has a negative weight along with government spending. We can say that this component identifies "catching-up" countries, with a market steering government. Figure 2a shows the scatter plot of the scores for 2<sup>nd</sup> function. We see that there are spatial clusters but not as strong and clear as in the first factor. We find countries like Qatar, China, Oman, etc., at the higher end of the plot and most African countries, Haiti and Cuba at the lower end.

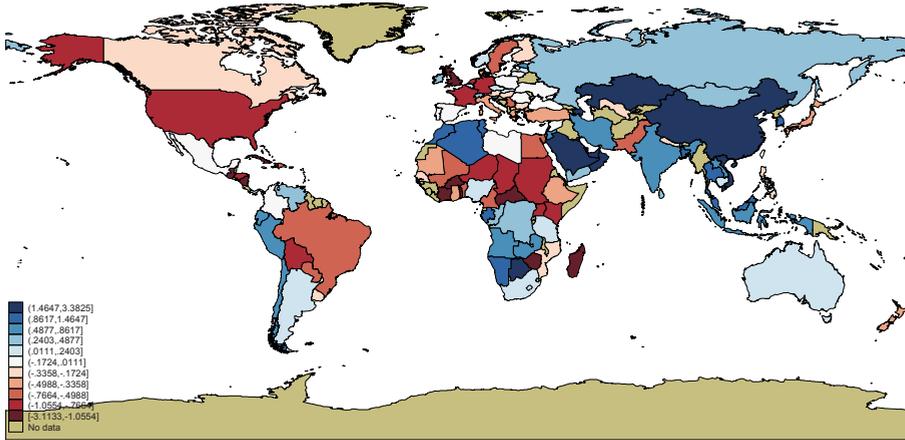


**Figure 2a:** Spatial correlation between institutions and economic performance (factor 2)

Figure 2b maps the institutional scores and Figure 2c maps the scores of economic performance. From the two maps, we can see that there is very close spatial correspondence between the two sets of scores in the Eastern part of the world. This is not the case among African, Western European and Latin American countries between the two sets of scores.



**Figure 2b:** IPD scores Factor 2 (IPD-PWT)



**Figure 2c:** PWT scores Factor 2 (IPD-PWT)

The variables that have high positive loadings in the third loadings vector of IPD are ‘security of contracts’ (A4), ‘strong presence of the government’ (A3) and ‘weakly segmented labour market’ (D3). The variables that have high negative loadings include ‘strong competition in the goods market’ (B3), ‘well developed free and open financial sector’ (C1) and ‘government regulated financial sector’ (C2). It is the political institutions and the financial institutions that are the most important ones in this factor. The only PWT variables that load positively but low are openness and investment. The highest negative loadings are on the initial level of GDP and population growth rate. The spatial correlation for the third component is 0.33. Zimbabwe is clearly an outlier here.

### Institutions and Economic development (factor 3)

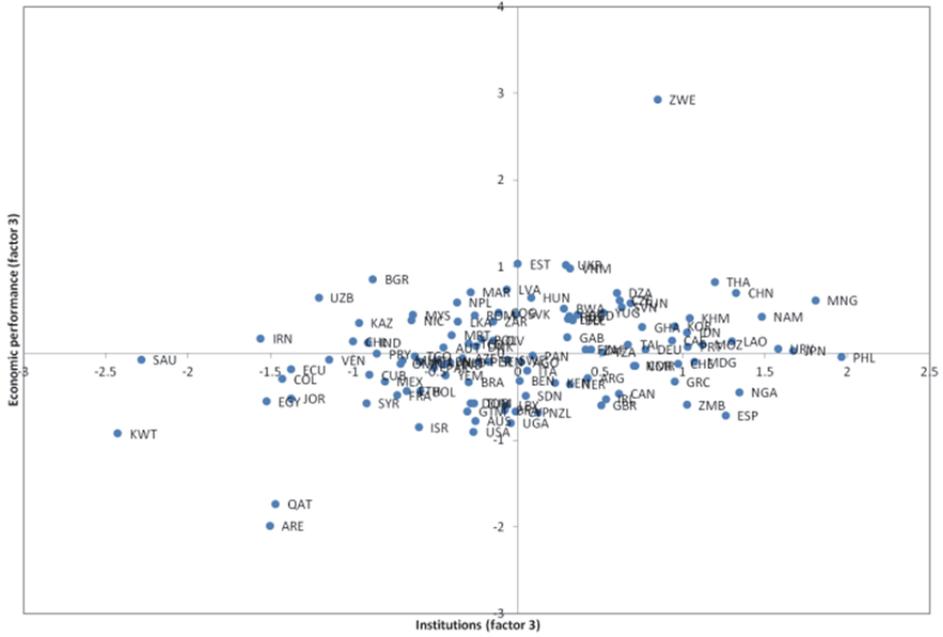


Figure 3a: Spatial correlation between institutions and economic performance (factor 3)

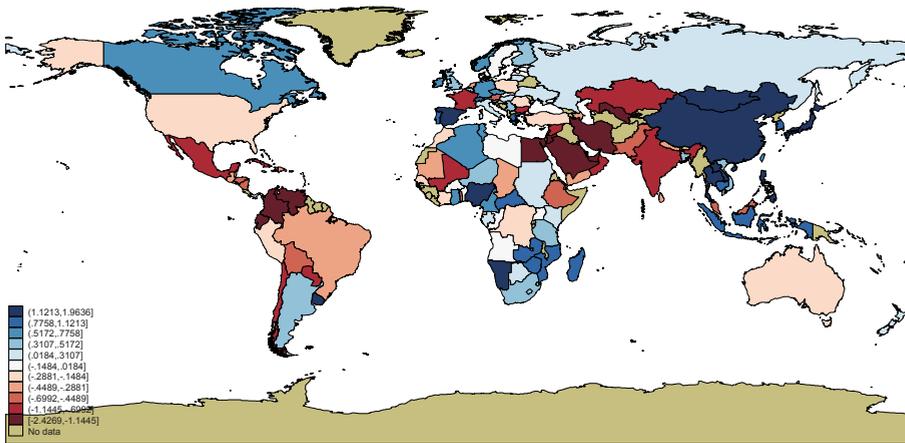


Figure 3b: IPD scores Factor 3 (IPD-PWT)

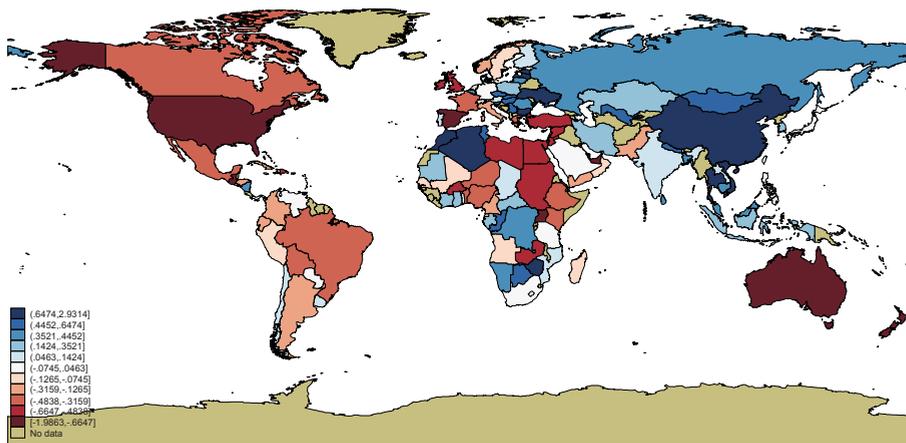


Figure 3c: PWT scores Factor 3 (IPD-PWT)

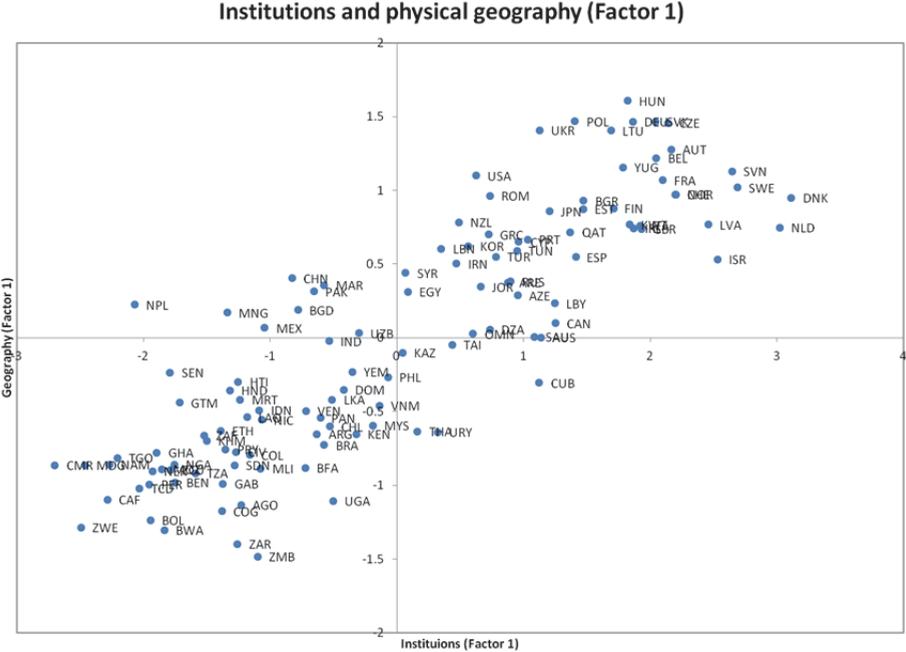
## B. Institutions and Geography

The absolute geographic location and the endowments that countries have, are often used in studies as an ultimate source of growth. While some studies incorporate variables such as the latitude of the country as having a direct impact on growth, others use variables such as the number of frost days in a year (which are essential for creating a rich top soil, thus making it more agriculturally productive) as an indirect source. The absolute geography variables that are used for both types of studies are a large and varied set. Rather than picking out one single variable, we divide the Gallup et al. (1999) data set into three different groups namely, (i) soil quality and climate zones, which contains – indicators referred to as GEOSOIL (ii) attributes of physical geography, which contains – indicators referred to as GEOAREA and (iii) the disease ecology of countries. In this study, we use only variables of GEOSOIL and GEOAREA since these are the invariable over time.

### IPD-GEOAREA

The IPD variables that load high in the first factor are ‘*socially inclusive democracy*’ (with political proximity to an EU country) (A1) ‘*security and enforcement of property rights laws*’ (B1), and ‘*society’s responsiveness to change and innovation*’ (with political proximity to Japan) (A2). It is again the political institutions and a positive market attitude that are important here. The geography variables that load high and positive are the latitude of country centroid, mean distance to the nearest coastline (km), and percentage of land area within 100 km of ice-free coast, while mean distance to nearest coastline or sea-navigable river (km), percentage of land area within 100 km of ice-free coast and percentage of land area in geographical tropics load negatively on the factor. The correlation we obtain here is 0.84. From the loadings, we can see that the most important features are the latitude of the country, easy access to ice-free coasts and easy

access to coastline or sea-navigable river (km). The scores of the factor are plotted in Figure 4a. Here again, we find most African countries have the lowest scores and European countries along with Canada and Australia score among the highest. Figure 4b maps the institutional scores and Figure 4c maps the GEOAREA variables. The GEOAREA clusters are tighter than clusters of institutional scores. Central and South Africa has the largest cluster, scoring the lowest, while Northwestern and Eastern Europe have the highest scores. USA and New Zealand are outliers. The institutional clusters are much smaller and more scattered.



**Figure 4a:** Spatial correlation between institutions and physical geography (factor 1)

This canonical function captures the geography hypothesis that institutional development is highly correlated to a certain set of geographic conditions (often proxied by latitude and access to waterways and coastline) which are considered favorable for economic development. What we see is that the geography scores clearly set apart Sub-Saharan Africa whereas the institutional scores are more varied. Similarly, we see that Western European institutional scores are slightly more varying than the geography clusters for the same region. Spain, for example, has an institutional score closer to the other European countries, but has a geography score closer to North Africa. Similarly the Neo-European countries have institutional score closer to the European ones than their geographic scores.

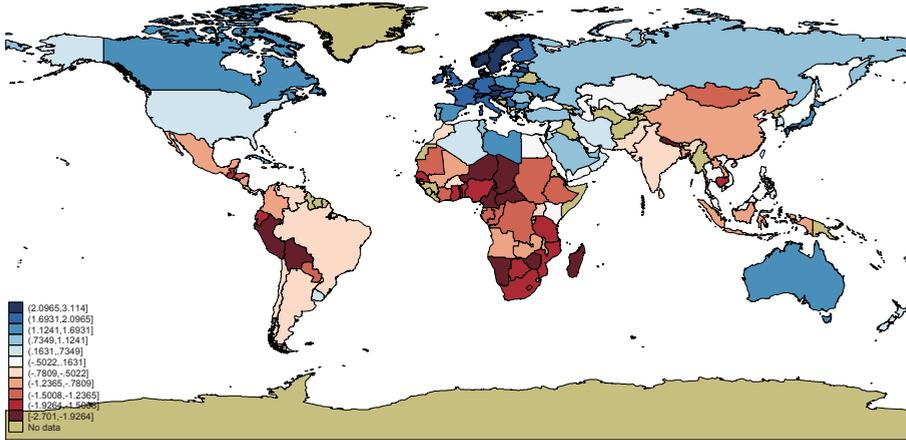


Figure 4b: IPD scores Factor 1 (IPD-GEOAREA)

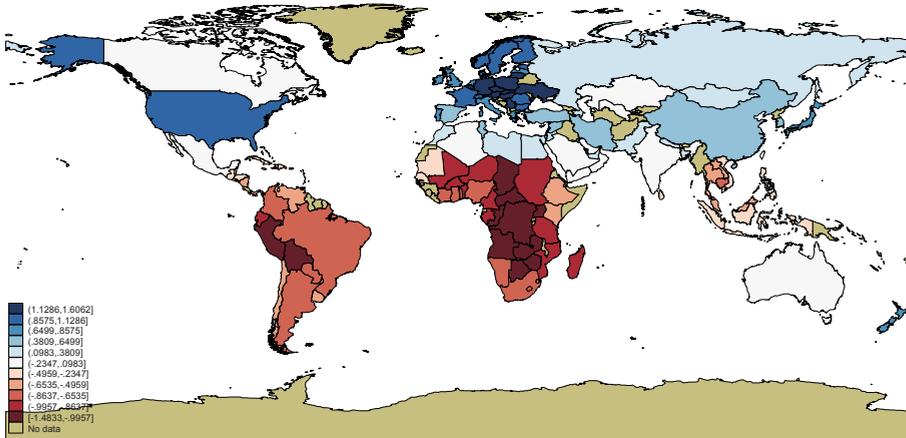


Figure 4c: GEOAREA scores Factor 1 (IPD-GEOAREA)

The second canonical component has a spatial correlation is 0.60. The second factor loads positively on ‘*strong competition in the good market*’ (B3), ‘*government regulated financial sector*’ (C1), and ‘*weakly segmented labour market*’ (D3) and strongly negatively on ‘*security and enforcement of property rights laws*’ (B1), and ‘*security of contracts*’ (A4). This factor stresses on the functioning of the Goods & Services market sector. The geographic variables that load strongly positive are mean distance to nearest coastline or sea-navigable river (km) and latitude of country centroid, and negatively on mean distance to nearest coastline (km) and percentage of land area within 100 km of ice-free coast. This shows that availability of water transport matter the most, as in the first factor. Figure 5a shows the factor scores of this canonical function.

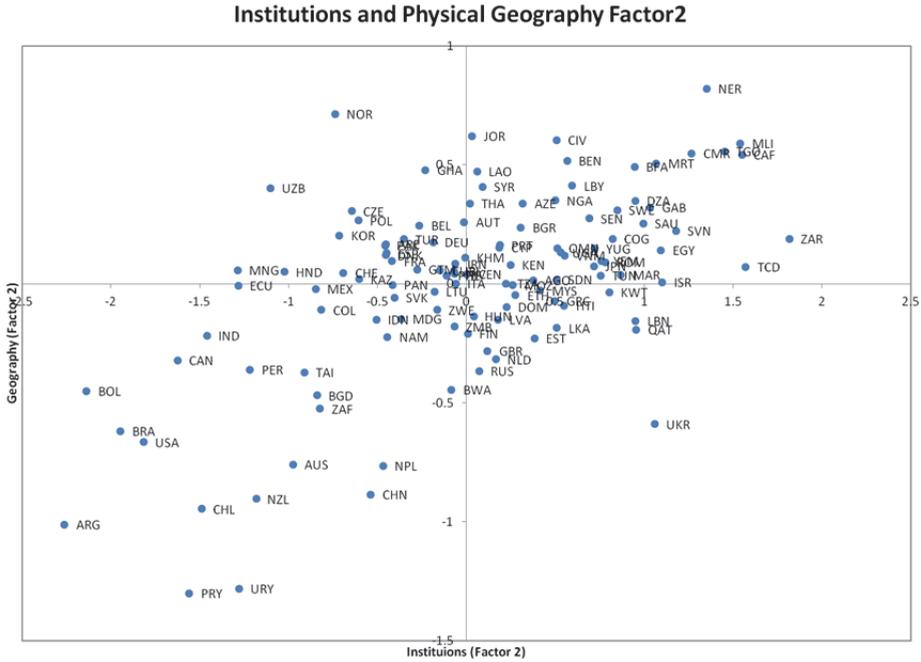


Figure 5a: Spatial correlation between institutions and physical geography (factor 2)

Figure 5b and 5c map the scores for institutions and geography respectively, for the second canonical function. Here we find that the institutional clusters are larger than the geography clusters. Here we find no clear differentiation between developed and developing countries. Rather we find that large countries have the least institutional as well as geography scores, excepting those in the African continent and Russia.

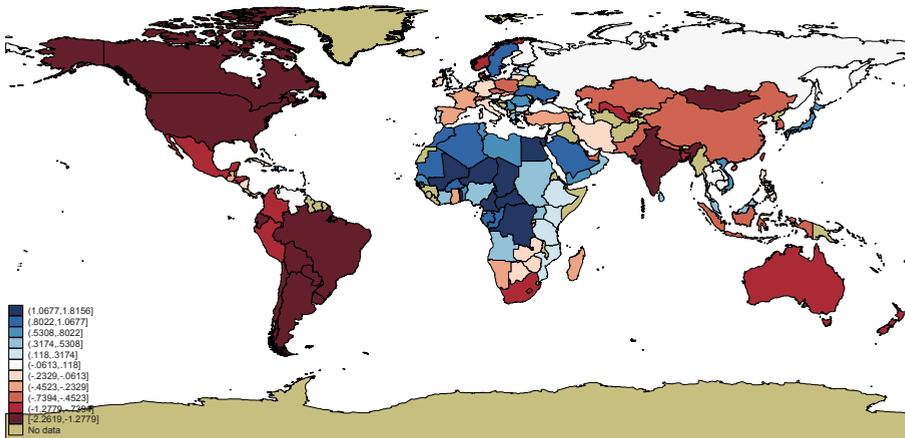
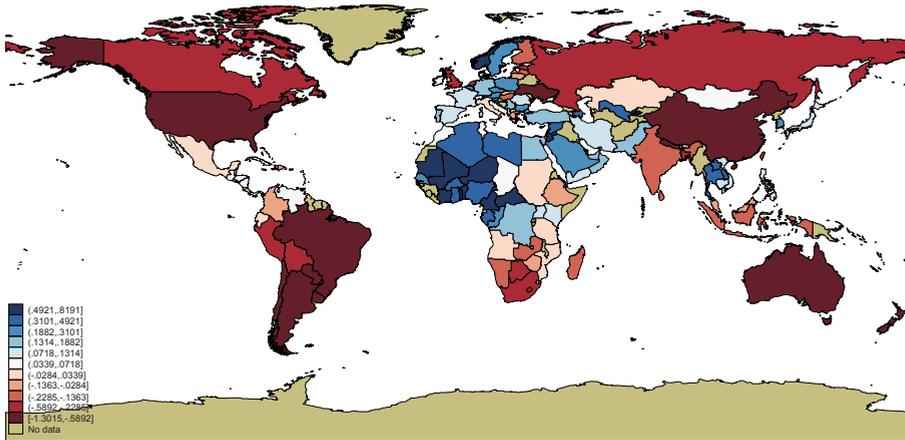


Figure 5b: IPD scores Factor 2 (IPD-GEOAREA)



**Figure 5c:** GEOAREA scores Factor 2 (IPD-GEOAREA)

The third canonical component has a spatial correlation is 0.37. This factor loads positively on ‘*socially inclusive democracy*’ (with *political proximity to an EU country*) (A1), and ‘*society’s responsiveness to change and innovation*’ (with *political proximity to Japan*) (A2). It loads strongly negative on ‘*strong government regulation and active economic policies*’ (B2), ‘*security and enforcement of property rights laws*’ (B1), and ‘*strong presence of government*’ (A3). The geographic variables that load strongly positive are distance from centroid of country to nearest coast (km) and to the nearest sea-navigable river and percentage of land area within 100 km of ice-free coast. It loads negatively on mean distance to nearest coastline or sea-navigable river (km). On the institutional side, we see that countries that have a strong presence of government have the lowest scores. Here again, we see that the institutional clusters are quite different from the geographical clusters.

### Institutions and physical geography (Factor 3)

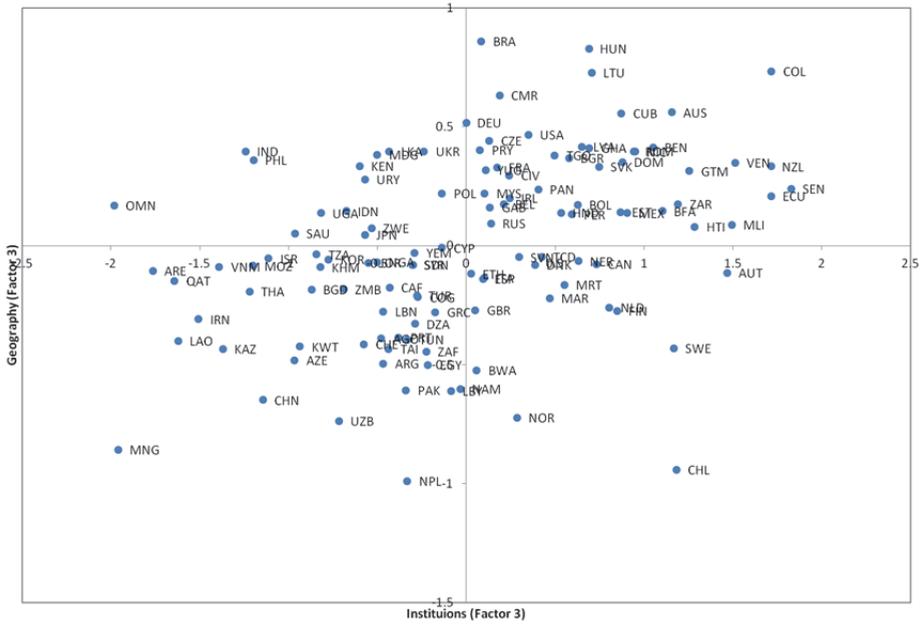


Figure 6a: Spatial correlation between institutions and physical geography (factor 3)

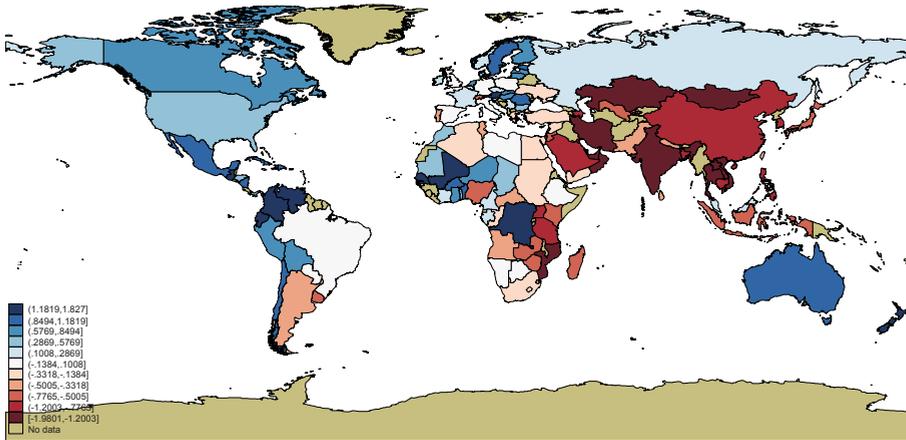


Figure 6b: IPD scores on factor 3 (IPD-GEOAREA)

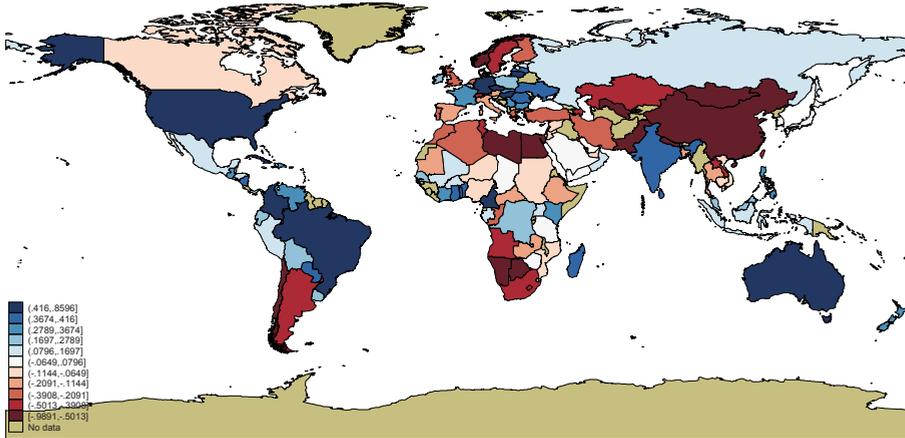
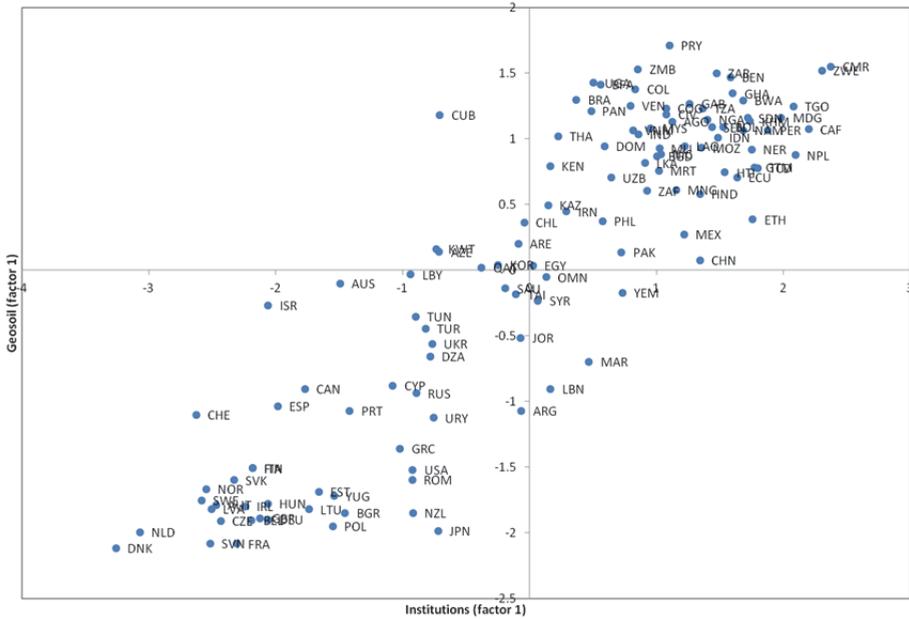


Figure 6c: GEOAREA scores on factor 3 (IPD-GEOAREA)

## IPD-GEOSOIL

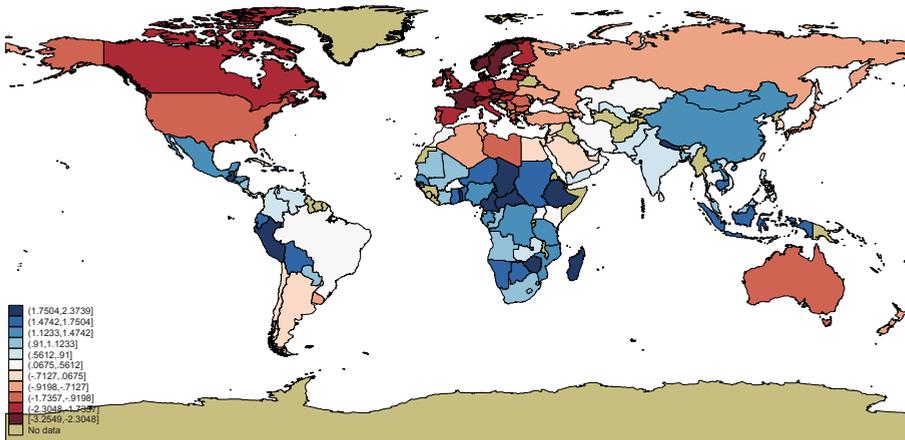
The first canonical function of IPD-GEOSOIL has a correlation of 0.83. The most significant loadings are given to ‘*socially inclusive democracy*’ (with *political proximity to an EU country*) (A1) and ‘*security and enforcement of property rights laws*’ (B1) both of which are strongly negative while ‘*non-meritocratic weak labour market*’ (D1) loads moderately positive. On the geography side, we find that percentage of land cultivated in the Köppen Geiger zones of humid temperate, cold climate with moist winter and warm climate with dry summers load strongly negative. Very suitable-mean soil suitability (soilsui1) also loads moderately negative. Steppe climate, savanna climate, and moderately suitable-mean irrigation suitability (irrsuit2) load moderately positively. Figure 7a shows clear clusters of countries, with Africa and South & South East Asia on the one hand and Western Europe, North America, North Africa and Australia on the other. We see that this component again captures the emerging consensus that institutional development has been shaped by geography. This stylized fact has been popularized by economists such as Jeffery Sachs, Jared Diamond, etc.

**Institutions and GEOSOL Factor 1**

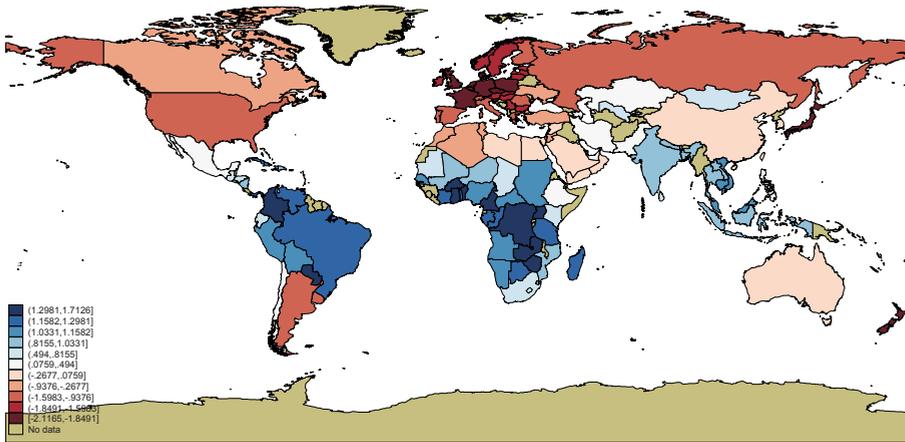


**Figure 7a:** Spatial correlation between institutions and soil (factor 1)

From the spatial maps in Figure 7b and 7c we can see that those climatic conditions found in Western Europe that are considered good for overall economic development, which is in turn related to institutional development are also the ones that are most emphasized in the literature. Japan and New Zealand are the only other countries that score similar to Western Europe in Figure 7c. What is interesting to note here is that in Figure 7b we see that China scores low on institutions while Australia scores high. However, in Figure 7c, we see that both countries have similar geographic conditions.



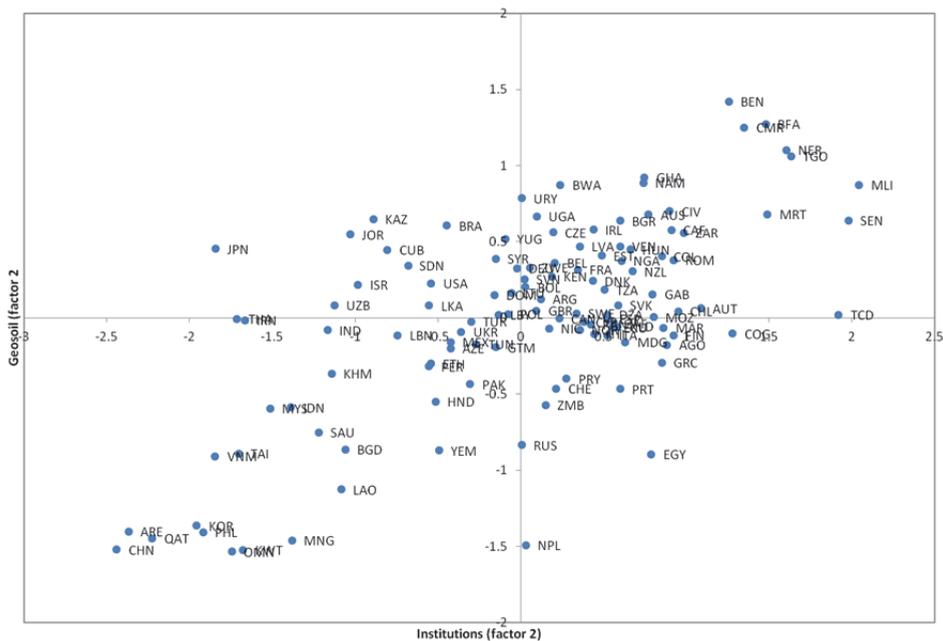
**Figure 7b:** IPD scores Factor 1 (IPD-GEOSOL)



**Figure 7c:** GEOSOL scores Factor 1 (IPD-GEOSOL)

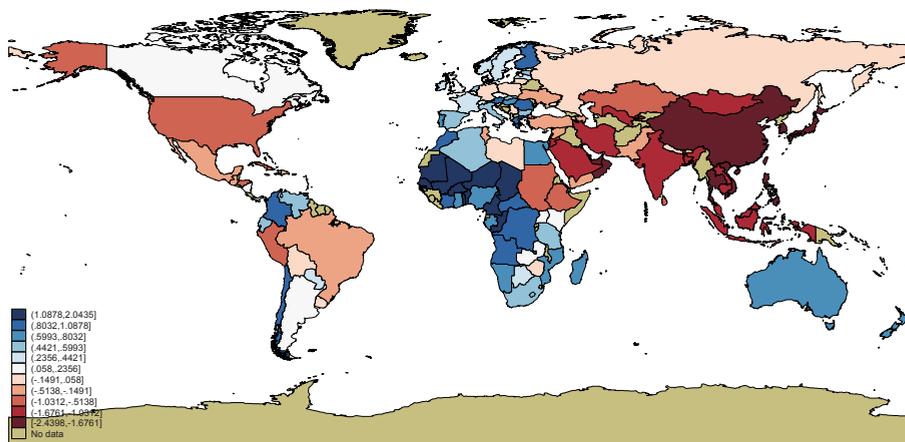
In the second factor, ‘*society’s responsiveness to change and innovation*’ (with political proximity to Japan) (A2) has the largest loading which is strongly negative. ‘*Strong presence of government in the goods and services market*’ (B2) and ‘*security and enforcement of property rights laws*’ (B1) also load negative, while ‘*strong presence of the government*’ (A3) loads positively. The geography variables that load high and positive include the percentage of land cultivated in the Köppen- Geiger zones of warm temperate climate with a dry summer, steppe climate, tropical savannah with dry winter and snow climate-fully humid and moderately suitable-mean soil solubility (soilsui2). The spatial correlation for this canonical function is 0.63. Figure 8a shows the scatter plot of the factor scores.

### Institutions and GEOSOIL Factor 2

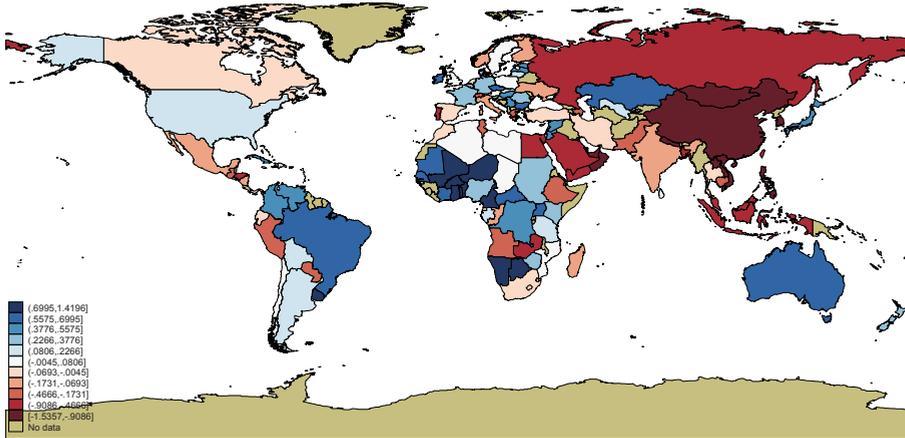


**Figure 8a:** Spatial correlation between institutions and soil (factor 2)

Here again, like in the second Figures 6b and 6c, we find that the institutional clusters are larger than the geography clusters, especially for Europe and Africa. These second factors show that geography is not necessarily destiny.



**Figure 8b:** IPD scores Factor 2 (IPD-GEOSOIL)



**Figure 8c:** GEOSoil scores Factor 2 (IPD-GEOSoil)

In case of the third factor, we find that ‘*security of contracts*’ (A4) has the strongest positive loading followed by ‘*society’s responsiveness to change and innovation*’ (with political proximity to Japan) (A2) and ‘*socially inclusive democracy*’ (with political proximity to an EU country) (A1). The geography variables that load positive and high are the percentage of cultivated lands in warm temperate fully humid climate, tropical savannah, warm temperate climate with dry winter, steppe climate, snow fully humid climate and tropical rain forest. This factor has a spatial correlation of 0.6. This canonical function stresses on the political institutions only. East Asia has the highest institutional scores, as seen in figure 9b. The rest of Asia has the lowest scores along with North Africa and Central America. The geographical clusters are larger here. Like in the first function, here we find larger geographical clusters than institutional clusters.

### Institutions and GEOSOIL Factor 3

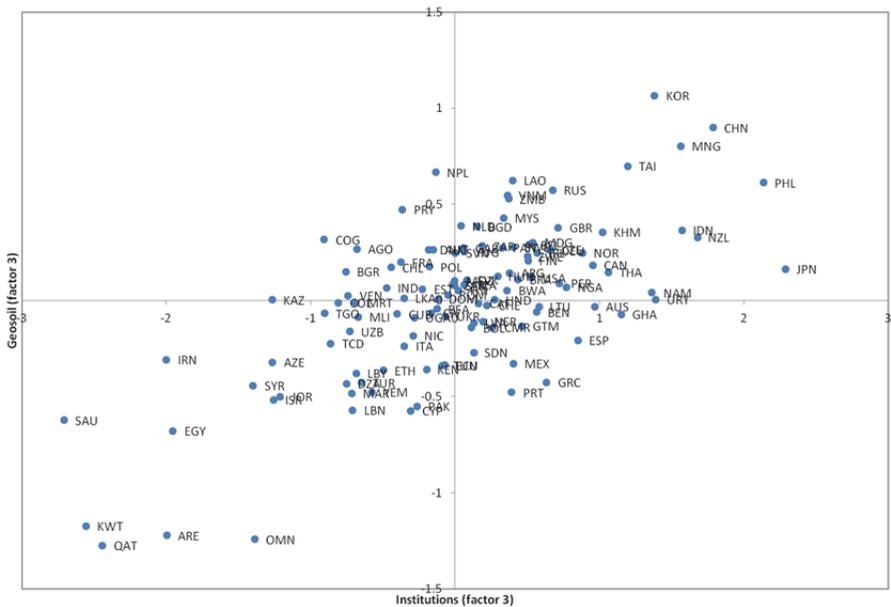


Figure 9a: Spatial correlation between institutions and soil (factor 3)

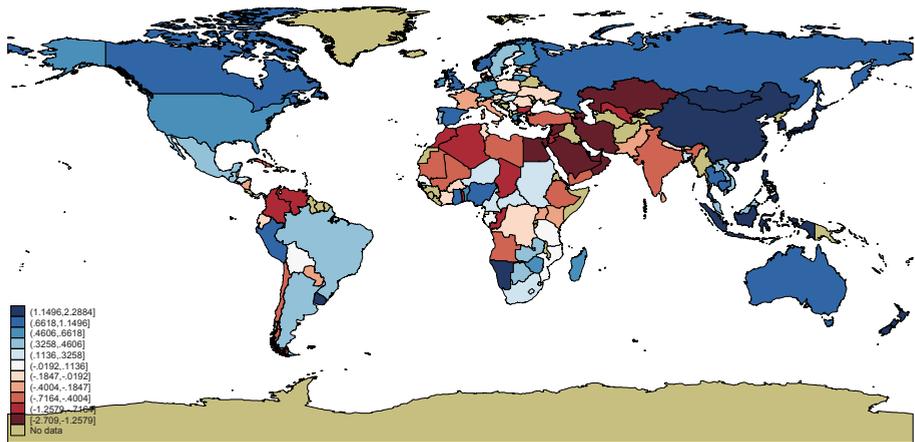


Figure 9b: IPD scores Factor 3 (IPD-GEOSOIL)

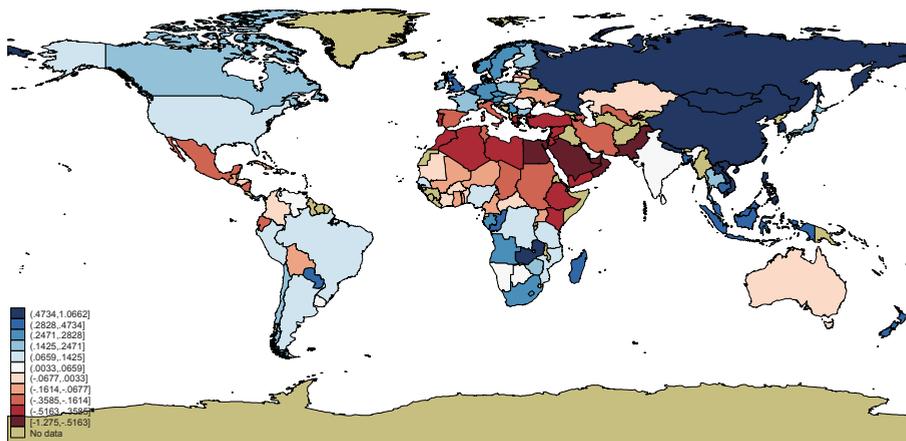


Figure 9c: GEOSOL scores on factor 3 (IPD-GEOSOL)

### C. Institutions and Financial Development (IPD-FINANCE):

Finally, we have a look at the interactions between institutional variables and finance variables. Figure 10a shows the scatter plot of the first set of scores from IPD and FINANCE plotted against each other. The spatial correlation here is 0.70. For the IPD data, we find that the first component loads positively high on ‘government controlled financial sector’ (C2) and ‘non-meritocratic weak labour market’ (D1) and negatively on ‘society’s responsiveness to change and innovation’ (A2), ‘strong presence of the government’ (A3), ‘enforcement of property rights laws’ (B1) and ‘competition in the good market’ (B3). It is interesting to note that the political institutional environment that was very important in determining the economic environment of the country is not as important a determinant for the first factor.

The FINANCE data loads positively high on the ratio of bank deposits to GDP (bdgdp) and private credit by deposit money banks / gdp (pcrdbgdp). The ratio of deposit money bank assets (dbagdp) and ratio of financial system deposits to GDP (fdgdp) load strongly negative. Fdgdp and bdgdp measure the absolute size of the financial sector (Beck and Demirgüç-Kunt 2009). While fdgdp measures the lending activities of the entire financial sector, bdgdp measures deposits of deposit monetary institutions. Since they are perfectly correlated, we find that if one variable loads positively, the other has to and does load negatively. Dbagdp and pcrdbgdp measure financial depth in the economy. King and Levine (1993) and Beck, Levine and Loayza (2000) find that countries where deposit money banks or commercial banks play a bigger role in financial intermediation than central banks, there is a higher level of financial development. This is captured by the first component. Fdgdp is stock indicator of resources available to the financial sector for its lending activities, which is one indication of the size of the financial system. From the first component we see that those countries with a higher level of income have a higher fdgdp ratio. Those countries which have a low ratio of private credit from both deposit money banks and other financial institutions to GDP, lower ratio of bank depos-

its to GDP and ratio of overhead costs to GDP are mainly low income countries. From the first component we see that the size of the banking and financial sector in the economy as well as the magnitude of private financial sector is what matters for development. The other variables have relatively negligible weights in all the factors.

Figure 10b maps the IPD scores of the first factor. From the figure we see clear clusters in the Americas and Asia (except for the Indian subcontinent). In Europe and Africa we find more variation. Figure 10c maps the Finance scores, from which we find smaller clusters, more varied across the map.

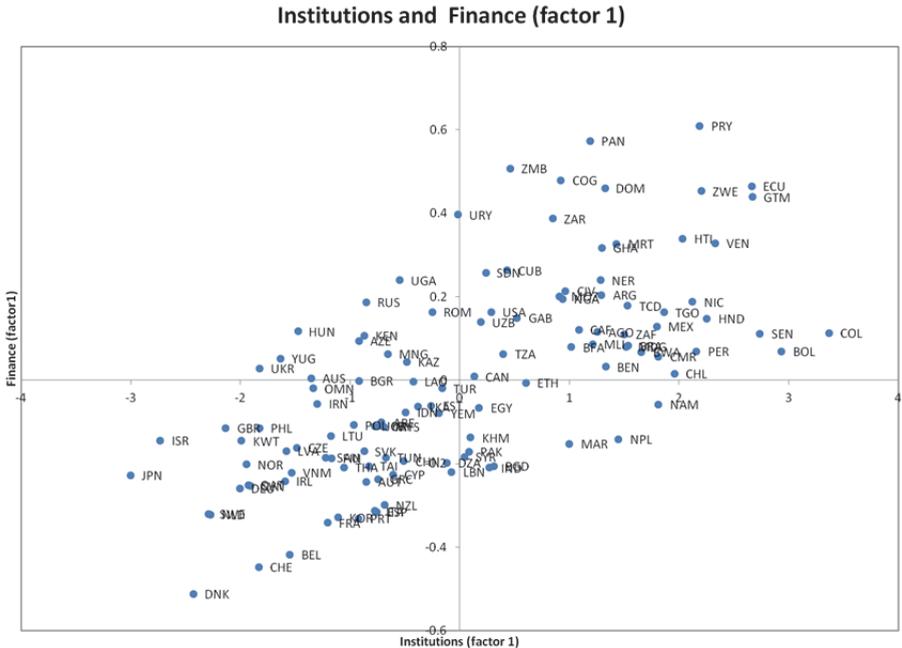


Figure 10a: Spatial correlation between institutions and financial development (factor 1)

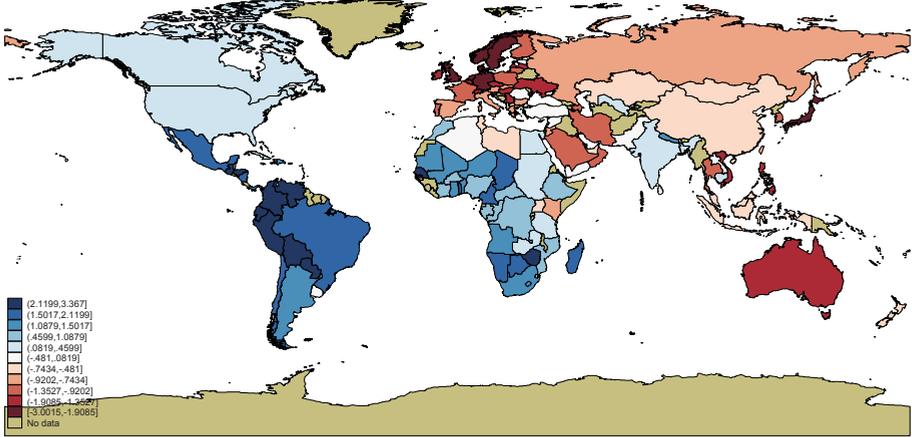
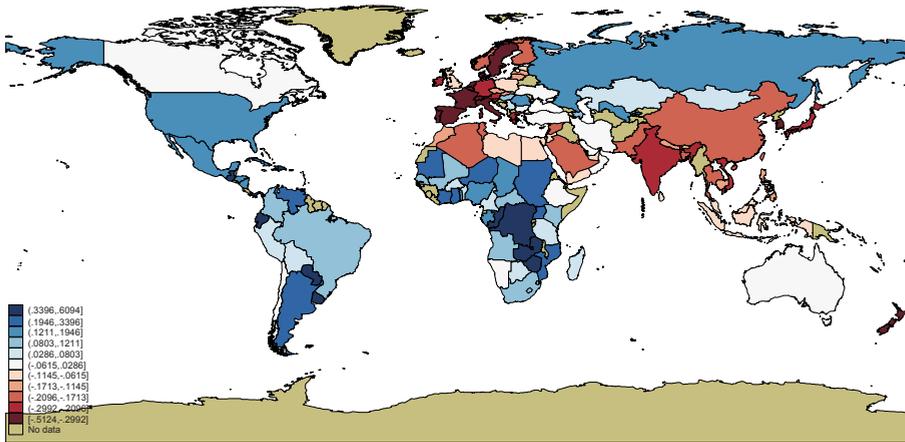
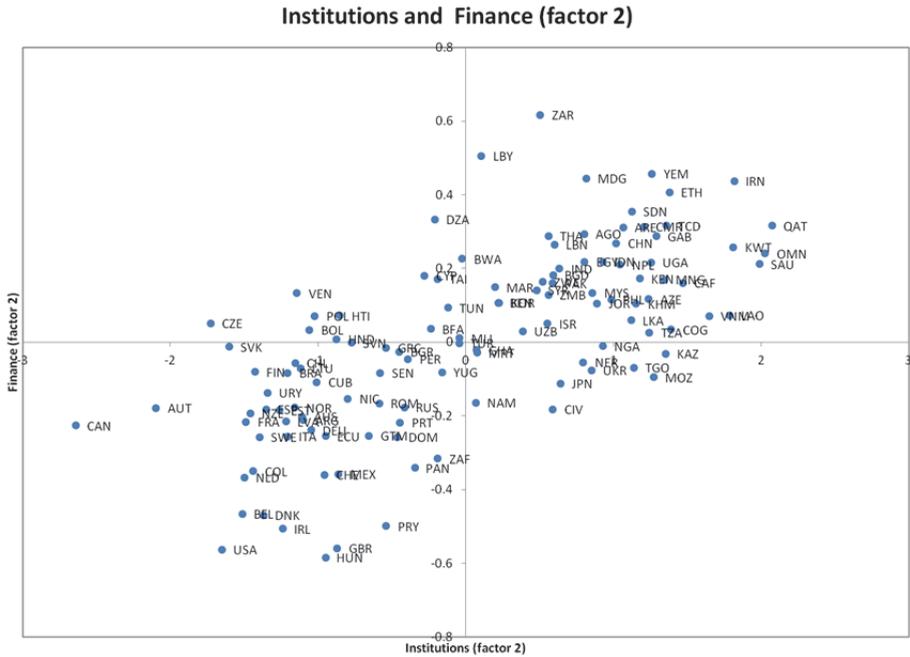


Figure 10b: IPD scores Factor 1 (IPD-FINANCE)



**Figure 10c:** FINANCE scores Factor 1 (IPD-FINANCE)

The spatial correlation of the second set of factors for the IPD-FINANCE data is 0.64. From the IPD data we see that *'strong presence of the government'* (A3) and *'non-meritocratic weak labour market'* (D1) load positively while *'socially inclusive democracy'* (A1) highly negative (-0.8402) and the fourth factor of sector A (A4) has a weight of -0.2284. We see that in this factor *'Public institutions and civil society'* is the most important sector of institutions. However the variable *'socially inclusive democracy'* has the largest loading. The loadings on the financial structure variables show that the ratio of financial system deposits to gdp (fdgdp) and the ratio of private credit by deposit money banks to gdp (pcrdbgdp) are positive and the ratio of bank deposits to GDP (bdgdp) are negative. No financial variables from the IPD are captured in this component, unlike in the case of the first factor.



**Figure 11a:** Spatial correlation between institutions and financial development (factor 2)

The map of the IPD scores, Figure 11b shows a cluster of North and South America along with Western Europe. Australia and New Zealand also belong to this cluster and are geographically outliers. We find small clusters in Africa and Asia as well. The map of finance scores shows more variety in South America as against the IPD scores. Africa and Asia however have larger clusters. With respect to the IPD-Finance analysis, we find that a larger financial sector goes hand-in-hand with a higher level of GDP in a country and that a more open financial sector leads to higher growth. We also find that countries with a lower level of democracy are also those which have a lower financial depth.

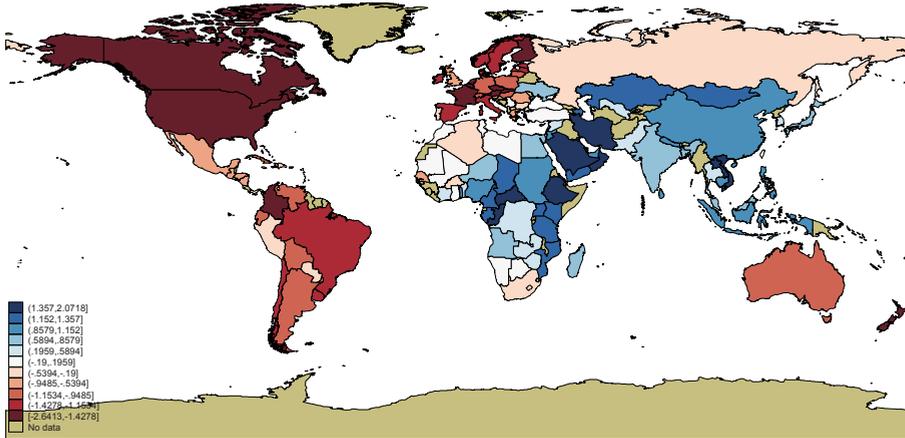


Figure 11b: IPD scores Factor 2 (IPD-FINANCE)

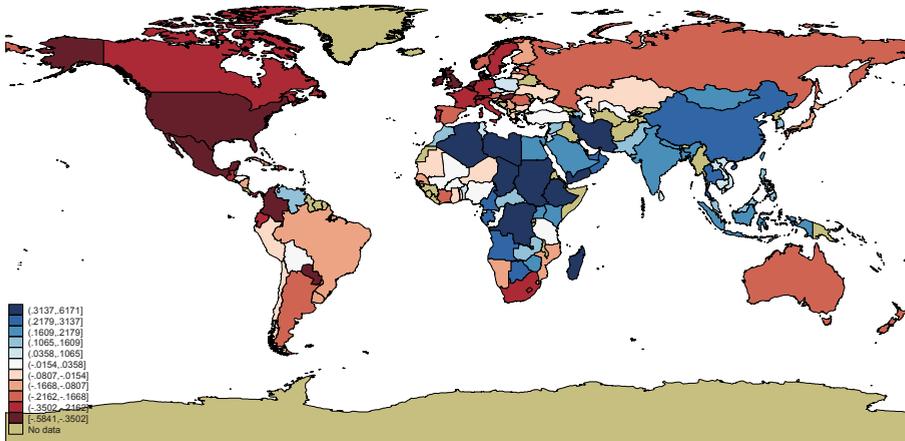


Figure 11c: FINANCE scores Factor 2 (IPD-FINANCE)

The spatial correlations for the third set of factors is relatively lower (0.44). The institutional variables that load high and positive are '*strong presence of the government*' (A3) and '*competition in the good market*' (B3). Those that load strongly negative include '*society's responsiveness to change and innovation*' (with political proximity to Japan) (A2), '*security and enforcement of property rights laws*' (B1), and '*non-meritocratic weak labour market*' (D1). From this we see that the goods and services market sector of IPD is relatively more important in this factor. We don't see the financial sector of the IPD having any impact on the factor. The finance variables that load high and positive are the ratio of deposit money bank assets to GDP, and bank deposits to GDP. The ratio of financial system deposits to GDP has a strong negative loading. Here again a competitive and open financial sector, which is relatively privatized is what seems to determine a high score for countries. Here too, the institutional clusters are stronger.

### Institutions and Finance (factor 3)

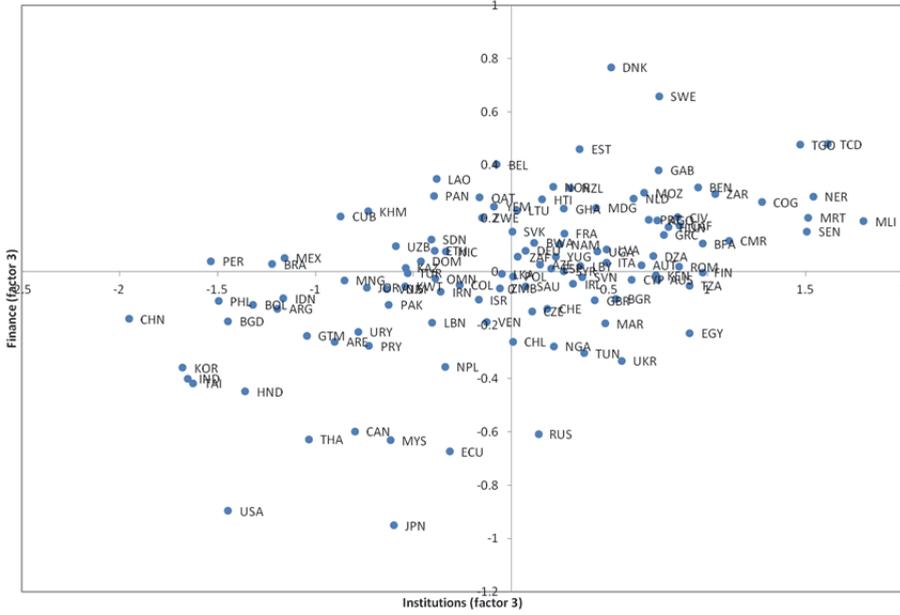


Figure 12a: Spatial correlation between institutions and financial development (factor 3)

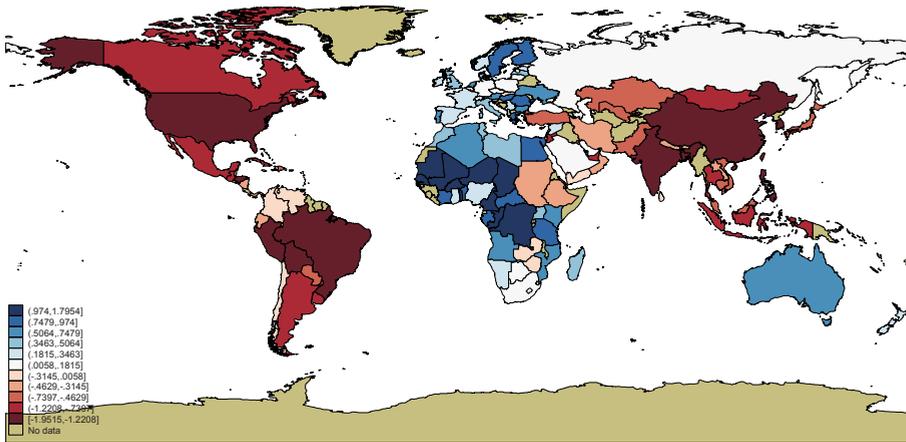


Figure 12b: IPD scores Factor 3 (IPD-FINANCE)

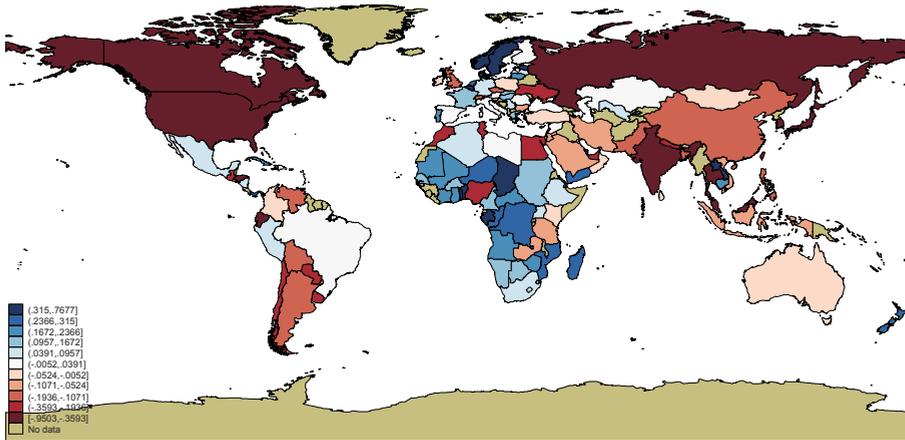


Figure 12c: FINANCE scores Factor 1 (IPD-FINANCE)

## 5. Conclusion

This paper analyzes institutional quality and economic performance through the lens of a spatial PCA and a spatial CCA. From the first canonical components of each spatial CCA we find results that conform to current literature. First, our analysis reiterates the institutional hypothesis that institutional quality of a country is strongly related to its level of GDP per capita. The more developed a country is, the better its institutional quality and vice versa, characterized by a high level of democracy, strong enforcement of property rights etc (as seen in Figure 1b and 1c). We also get results that are consistent with the geography hypothesis. We find that higher institutional quality is strongly related to ‘favorable’ geographic conditions, both in the case of physical geographic conditions and soil quality conditions in a country (Figure 4b, 4c, 7b and 7c). Further, we find that a larger financial sector is strongly correlated to a higher level of GDP in particular and a higher level of economic performance in general (Figure 10b and 10c). These results show that the influence of institutions varies across the levels of economic performance and financial performance.

Looking at the second canonical functions (Factor 2), we find other stylized facts. Firstly, we find that catching up based growth is related to market steering (Figure 2b and 2c). Secondly, we find that geography does not necessarily define the institutional structure of countries. We find that relative influences of neighbouring countries matters and need not be directly related to similar geographic conditions (Figure 5b, 5c, 8b and 8c). In other words, geography is not destiny. Thus unlike what most studies on the relationship between geography and economic performance say, countries do have a choice in developing and fostering any kind of institutional set up, though they are likely to be influenced by their neighbours to an extent. Both relative location as well as common historical backgrounds could define the institutional set up of countries. For example, Australia and New Zealand are often in the same cluster as Western Europe when look-

ing at institutional quality, but not necessarily when we consider their economic performance, and especially when we consider geographic similarities.

In this study, the causality of these relationships is not touched upon. This is one apparent disadvantage of the canonical correlation analysis. Further, the validity of these conclusions is based on the quality of the data resulting from a survey. It is important to keep in mind that biases are inherent to perceptions indices. However, one can still conclude, given the above correlations, that there is no single/global solution to address under-development and growth. The role of institutions in the economic performance of a country needs to be addressed at regional and country specific levels. Another drawback of this study is that it uses a cross-section of data. Institutions evolve over long periods of time, and are highly dependent on the historical socio-political and economic events of a country. To analyze the quality of institutions at a single point in time can be limiting. For example, the under-development of certain parts of the world in the process of the development of Europe in the colonial era cannot be discounted in understanding growth processes today. However, the lack of time-series data limits us to looking at a snap-shot picture of the current institutional set up across countries.

These results show us that we need to first understand the multi-dimensional nature of institutions before understanding their impact on or by the other factors discussed above. We find that similar geographic environment can lead to different institutional environment and also to different economic performance. Furthermore, there is a need to understand them in a spatial context as seen from the above results. While there is a lot of empirical work in the area, there is also a need for more theoretical understanding of this multi-dimensionality of institutions. Most theoretical attempts until now have been based in a very Europe-centric framework. In order to fully understand economic growth and development in the context of institutional quality, we need to look beyond theories that explain the development of the now developed countries. Our results show that spatial clusters are not always based on the level of economic development of countries. This is one major drawback which hinders current policy effectiveness.

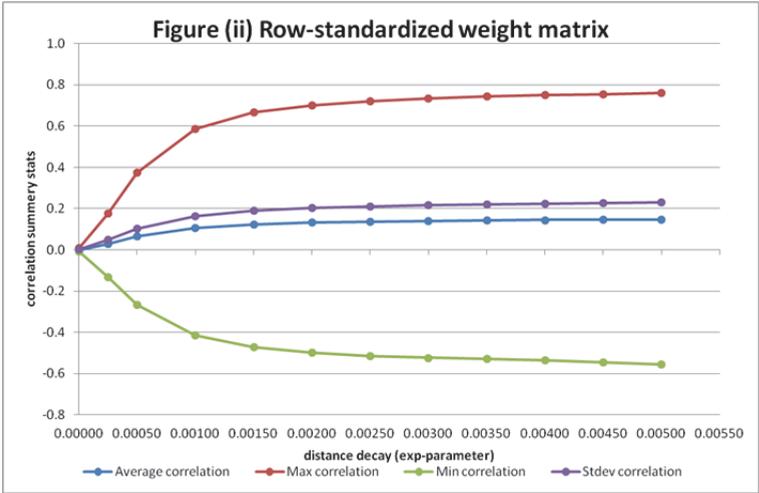
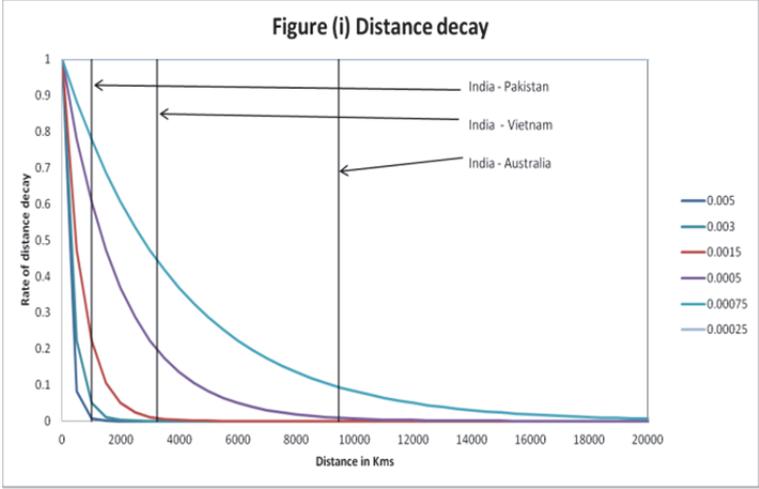
## Appendix A: Weights matrix

We choose a set of geographic distance weights for the connectivity matrix. We use an exponential distance decay function to obtain the weights matrix, given by the formula,  $w_{ij} = e^{-0.0015d_{ij}}$  where  $w_{ij}$  is the spatial weight between countries  $i$  and  $j$ , and  $d_{ij}$  is the bilateral distance between the centroids of the two countries  $i$  and  $j$ . The choice of the connectivity matrix has implications on the results obtained from the empirical analysis (Ward and Gleditsch 2007). Other types of connectivity matrices include binary contiguity matrices, bilateral trade matrices etc<sup>30</sup>.

We describe our choice using the Example of India. India's immediate neighbours (those that share a border with it) include: Pakistan, China, Nepal, Bhutan, Bangladesh, Myanmar (Burma), Sri Lanka and Afghanistan (in the disputed area). Tajikistan, Thailand, Iran, Vietnam etc are further away. The question then becomes one of how much weight should Thailand and Vietnam get, to depict the influence they exert on India. We choose that distance decay measure that gives importance to immediate neighbours. The weights given to Vietnam are significantly lower and those give to Australia get an insignificant weight. Figure (i) shows the decay in graphical terms for India with respect to Pakistan, Vietnam and Australia. The steeper the fall of the curve, the higher is the decay, i.e., relatively closer countries also exert lesser influence on the given country. In our case, we find that the nearest neighbours already have a fairly low weight. Since we consider distance between centroids of countries we note that larger countries tend to give lower weights to neighbours than smaller countries. Figure (ii) shows the effect of each of these distance decay values have on the summary statistics on the spatial lag, and in turn on the spatial correlation of a row-standardized weights matrix.

---

<sup>30</sup> For more details on connectivity matrices and their effects on weight matrices and in turn on the empirical results can be found in Ward and Gleditsch 2007.



## Appendix B- IPD categories

	A. Public Institutions & Civil Society	B. Goods and Services Markets	C. Capital Markets	D. Labour Markets & Labour Relations				
1- Political Institutions	Democracy, Legality and Freedom	a10		Freedom of association and trade union pluralism	d10			
	Political stability and Legitimacy	a11						
	Decentralization	a12						
2 - Security, Law and Order	Domestic public security and Control of violence	a20						
	Control of State violence by NGOs	a21						
	External security	a22						
3 - Functioning of Public Administrations	Governance of public administration and the justice	a30	Ease of starting a business	b30				
	Autonomy of public policies	a31	Importance of the Economic	b31				
	Donors' influence	a32	Consideration of public interest in government-	b32				
	Autonomy in operation and creation of organizations	a33	Governance of natural resources	b33				
	Government capacity to reform	a34						
	Fiscal exemptions	a35						
4 - Free Operation of Markets			Privatizations	b40	Privatizations in the financial sector	c40	Flexibility in the labour market	d40
			Nationalizations	b41	Nationalizations in the financial sector	c41	Retraining and reskilling measures	d41
			Governance of privatizations	b42	Freedom in the allocation of loans	c42		
			Performance of public organizations	b43				
			Freedom of prices	b44				
		Single exchange rate	b45					
5 - Coordination of Stakeholders and Strategic visions	Capacity of the State to coordinate stakeholders	a50	Technological environment	b50	Competence of bank executives	c50	Adaptive education system	d50
	Strategic capacities	a51	Public aid for R&D	b51	Importance of venture capital	c51		
	Government's arbitration capacity	a52	Density of sub-contracting relations	b52	Sovereign wealth fund policy	c52		
	Institutional capacity	a53						
	Government political capacity	a54						
	Change, Innovation	a55						
	Cooperative behaviour in society	a56						
	Outlook of young people	a57						
6 - Security of Transactions and Contracts	Security of transactions and	a60	Information on G&S markets	b60	Financial information	c60	Respect for workers' rights	d60
	Government respect for contracts	a61	Rural land tenure: traditional property	b61			Weak employment contract rigidity	d61
	Frequency of bankruptcy	a62	Rural land tenure: public	b62				
	Enforcement of bankruptcy law	a63	Diversity of land tenure rights systems	b63				
			Government recognition of diversity of land tenure rights	b64				
			Land tenure: security of	b65				
			Land tenure: demand for land	b66				
		Land tenure and large	b67					
7 -Market regulation, Social Dialogue			Competition on G&S markets	b70	Competition within the banking system	c70	Wage bargaining at the individual level	d70
			Shareholders: weight of the government	b71	Regulation of competition in banking	c71	Strikes	d71
			Information on shareholders	b72	Monitoring and auditing in banking	c72	Management of labour	d72
			Land tenure: development	b73	Reform of financial regulations	c73		
8 - Openess to the Outside World	Free movement of people and information	a80	Openness to business	b80	Financial openness	c80	Openness to employment of non-nationals	d80
	Political proximity with big		Joint Ventures	b81				
	Emulation with neighboring	a82	Non-national access to land	b82				
9 - Social Cohesion	National sense of identity	a90		Micro lending		c90	Quality of the supply of public goods	d90
	National cohesion	a91					Weak segmentation of the labour market	d91
	Strengthening of middle classes	a92					Low incidence of child labour	d92
	Social inclusion	a93					Social mobility	d93
	Traditional solidarity	a94					Social mobility: young higher education graduates	d94
	Subsidies on commodities	a95f					Distribution of income	d95

## Appendix C: Description of Variables

Variable	Description	Dataset
LY00	GDP per capita, 2000, in intl. dollars	PWT
GR	GDP per capita, annual average growth rate 2000-2007	PWT
I	Investment share of PPP converted GDP per capita	PWT
G	Government expenditure share of PPP converted GDP per capita	PWT
OP	Net exports share of PPP converted GDP per capita	PWT
LPOP	Population, average over the period 2000-2009	PWT
C	consumption share of PPP converted GDP per capita	PWT
areakm2	land area (km <sup>2</sup> )	GEOAREA
cen_lat	latitude of country centroid	GEOAREA
elev	mean elevation (meters above sea level)	GEOAREA
distcr	mean distance to nearest coastline or sea-navigable river (km)	GEOAREA
distc	mean distance to nearest coastline (km)	GEOAREA
distr	mean distance to nearest inland navigable river (km)	GEOAREA
cen_c	distance from centroid of country to nearest coast (km)	GEOAREA
cen_cr	distance from centroid of country to nearest coast or sea-navigable river (km)	GEOAREA
lc100km	% Land area within 100 km of ice-free coast	GEOAREA
lcr100km	% Land area within 100 km of ice-free coast (or navigable river)	GEOAREA
tropical	% Land area in geographical tropics	GEOAREA
soilsui1	mean soil suitability 1, very suitable (%)	GEOAREA
soilsui2	mean soil suitability 2, moderately suitable (%)	GEOAREA
irrsuit1	Mean irrigation suitability, very suitable (%)	GEOAREA
irrsuit2	Mean irrigation suitability, moderately suitable (%)	GEOAREA
cultmaf	% cultivated land in Köppen-Geiger Af zone	GEOSOIL
cultmam	% cultivated land in Köppen-Geiger Am zone	GEOSOIL
cultmaw	% cultivated land in Köppen-Geiger Aw zone	GEOSOIL
cultmbs	% cultivated land in Köppen-Geiger Bs zone	GEOSOIL
cultmbw	% cultivated land in Köppen-Geiger Bw zone	GEOSOIL
cultmcf	% cultivated land in Köppen-Geiger Cf zone	GEOSOIL
cultmcs	% cultivated land in Köppen-Geiger Cs zone	GEOSOIL
cultmcw	% cultivated land in Köppen-Geiger Cw zone	GEOSOIL
cultmdf	% cultivated land in Köppen-Geiger Df zone	GEOSOIL
cultmdw	% cultivated land in Köppen-Geiger Dw zone	GEOSOIL
cultme	% cultivated land in Köppen-Geiger E zone	GEOSOIL
cultmh	% cultivated land in Köppen-Geiger H zone	GEOSOIL
dbacba	Deposit Money Bank Assets / (Deposit Money + Central) Bank Assets	FINANCE
llgdp	Liquid Liabilities / GDP	FINANCE
cbagdp	Central Bank Assets / GDP	FINANCE
dbagdp	Deposit Money Bank Assets / GDP	FINANCE
pcrdbgdp	Private Credit By Deposit Money Banks / GDP	FINANCE

Variable	Description	Dataset
pcrdbofgdp	Private Credit By Deposit Money Banks And Other Financial Institutions / GDP	FINANCE
bdgdp	Bank Deposits / GDP	FINANCE
fdgdp	Financial System Deposits / GDP	FINANCE
bcdb	Bank Credit / Bank Deposits	FINANCE
overhead	Bank Overhead Costs / Total Assets	FINANCE
netintmargin	Net Interest Margin	FINANCE
concentration	Bank Concentration	FINANCE
roa	Bank ROA	FINANCE
roe	Bank ROE	FINANCE
costinc	Bank Cost-Income Ratio	FINANCE
zscore	Bank Z-Score	FINANCE
nrbloan	Loans From Non-Resident Banks (Amt Outstanding) / GDP	FINANCE
offdep	Offshore Bank Deposits / Domestic Bank Deposits	FINANCE

## Appendix D: Factor Loadings on IPD variables (First Stage)

### Sector A

	Variable Description	A_mspc1	A_mspc2	A_mspc3	A_mspc4
A10	Democracy, Legality and Freedom	0.3152	-0.044	-0.2097	0.0212
A11	Political stability and Legitimacy	-0.1679	0.1504	-0.0952	0.0314
A12	Decentralization	0.1112	-0.1891	-0.2707	0.1913
A20	Domestic public security and Control of violence	-0.028	0.1732	0.2465	0.1358
A21	Control of State violence by NGOs	0.023	0.0969	0.051	-0.2903
A22	External security	-0.0576	-0.1279	0.0236	0.0985
A30	Governance of public administration and the justice system	0.1334	-0.3579	0.2562	-0.2996
A31	Autonomy of public policies	-0.0692	0.1265	0.1068	-0.0114
A32	Donors' influence	0.0365	-0.0311	0.0169	0.0915
A33	Autonomy in operation and creation of organizations	-0.0644	-0.2088	0.1485	0.1997
A34	Government capacity to reform	0.1149	-0.2476	-0.4647	-0.1155
A35	Fiscal exemptions	0.0893	0.0013	-0.0245	0.0402
A50	Capacity of the State to coordinate stakeholders	0.0764	0.0946	0.1469	0.2406
A51	Strategic capacities	-0.1537	-0.1231	-0.0242	-0.223
A52	Government's arbitration capacity	-0.0538	0.0861	0.0775	0.0947
A53	Institutional capacity	0.037	0.1769	0.1357	0.2138
A54	Government political capacity	0.0443	-0.0729	-0.0072	-0.118
A55	Change, Innovation	-0.2244	0.4501	0.2883	0.1008
A56	Cooperative behaviour in society	-0.0932	0.0186	0.1724	-0.0716
A57	Outlook of young people	-0.0432	0.0061	0.1432	0.1307
A60	Security of transactions and contracts	-0.1139	0.2703	-0.0983	0.3528
A61	Government respect for contracts	0.0136	-0.0536	-0.0899	0.0086
A62	Frequency of bankruptcy	0.0502	-0.1086	-0.052	0.0425
A63	Enforcement of bankruptcy law	0.0277	-0.0805	0.0141	-0.0074
A80	Free movement of people and information	-0.0589	0.0358	-0.0988	0.1371
A8010	Political proximity with USA	0.0104	0.1287	0.0758	-0.0382
A8011	Political proximity with an EU country	0.3398	-0.3352	0.0075	-0.3613
A8012	Political proximity with Japan	-0.1893	0.2258	0.0432	0.0569
A8013	Political proximity with China	-0.0872	0.0055	0.2162	0.1288
A82	Emulation with neighboring countries	0.0206	0.1615	-0.0019	-0.0139
A90	National sense of identity	-0.0189	0.0857	-0.0563	0.125
A91	National cohesion	-0.0543	-0.111	-0.2794	-0.0754
A92	Strengthening of middle classes	0.0381	-0.0169	0.1111	0.003
A93	Social inclusion	0.6866	0.0348	0.1784	-0.3662
A94	Traditional solidarity	-0.1157	0.0572	0.2386	-0.0754
A95f	Subsidies on commodities	0.2017	-0.1924	-0.2161	-0.1719

**Sector B**

IPD Variable	Variable Description	B1	B2	B3
B30	Ease of starting a business	0.1493	-0.2875	0.134
B31	Importance of the Economic zones	-0.0581	0.3448	-0.291
B32	Consideration of public interest in government-business relations	-0.1657	0.2344	-0.1635
B33	Governance of natural resources	0.0596	-0.0041	-0.1982
B40	Privatizations	0.2403	-0.1203	0.2834
B41	Nationalizations	0.2046	-0.1175	0.0058
B42	Governance of privatizations	-0.0012	-0.1228	0.1439
B43	Performance of public organizations	0.1198	0.0989	0.0031
B44	Freedom of prices	0.0926	-0.1119	-0.0488
B45	Single exchange rate	-0.0281	-0.0955	0.1305
B50	Technological environment	0.249	0.3117	0.0456
B51	Public aid for R&D	-0.0651	-0.2377	0.1204
B52	Density of sub-contracting relations	0.1326	-0.2395	0.0017
B60	Information on G&S markets	0.2	0.031	0.1038
B61	Rural land tenure: traditional property	-0.3423	-0.2251	0.0548
B62	Rural land tenure: public property	-0.1466	-0.0065	-0.1727
B63	Diversity of land tenure rights systems	0.1445	0.0081	-0.061
B64	Government recognition of diversity of land tenure rights systems	-0.1606	-0.0675	-0.1109
B65	Land tenure: security of ownership	0.3153	0.2348	0.1814
B66	Land tenure: demand for land	0.1838	-0.0004	0.2013
B67	Land tenure and large investors	-0.0805	0.1109	-0.0997
B70	Competition on G&S markets	-0.132	0.0535	-0.6298
B71	Shareholders: weight of the government	0.0345	0.2188	0.1918
B72	Information on shareholders	0.0755	-0.3091	0.2116
B73	Land tenure: development policies	0.1072	-0.1641	0.1125
B80	Openness to business	-0.0083	0.1836	0.2
B81	Joint Ventures	0.306	0.1094	-0.0458
B82	Non-national access to land	-0.4915	-0.3394	-0.1356

**Sector C**

IPD Variable	Variable Description	C1	C2
C40	Privatizations in the financial sector	0.0595	-0.0467
C41	Nationalizations in the financial sector	0.1985	-0.0939
C42	Freedom in the allocation of loans	0.1531	-0.3117
C50	Competence of bank executives	0.1356	0.2854
C51	Importance of venture capital	0.368	-0.3331
C52	Sovereign wealth fund policy	-0.1862	0.5076
C60	Financial information	-0.3638	-0.0751
C70	Competition within the banking system	-0.1536	-0.0024
C71	Regulation of competition in banking	0.4207	0.2234
C72	Monitoring and auditing in banking	0.3515	0.3023
C73	Reform of financial regulations	0.0707	-0.0965
C80	Financial openness	0.3052	0.2545
C90	Micro lending	-0.4364	0.4699

**Sector D**

IPD Variable	Variable Description	D1	D2	D3
D10	Freedom of association and trade union pluralism	0.2443	-0.4614	0.1625
D40	Flexibility in the labour market	-0.0765	-0.3393	-0.0836
D41	Retraining and reskilling measures	-0.1675	-0.003	-0.0593
D50	Adaptive education system	0.0027	-0.2803	0.0394
D60	Respect for workers' rights	-0.0871	0.0819	0.0529
D61	Weak employment contract rigidity	-0.1747	0.3401	-0.0011
D70	Wage bargaining at the individual level	0.0235	0.2155	-0.1371
D71	Strikes	-0.0806	0.2255	0.3431
D72	Management of labour	-0.2551	-0.1942	0.2904
D80	Openness to employment of non-nationals	-0.1283	-0.1238	-0.1059
D90	Quality of the supply of public goods	0.0617	0.226	-0.3641
D91	Weak segmentation of the labour market	-0.1874	-0.1019	0.2309
D92	Low incidence of child labour	-0.6413	0.4341	0.4832
D93	Social mobility	-0.5463	0.0561	-0.3529
D94	Social mobility: young higher education graduates	0.0135	0.2443	-0.2891
D95	Distribution of income	0.1754	0.0892	-0.3149



## CHAPTER 4

# **Economic Growth, Institutions and Geography<sup>31</sup>**

---

<sup>31</sup> This chapter has also been published as a UNU-MERIT working paper 2013-056 (Bhupatiraju and Verspagen 2013b)

## 1. Introduction

Much of the empirical growth literature has focused on examining the causes of differences in growth across countries. Many theories have been put forth in this regard and in the recent decades interest in explaining and testing these differences has surged. One answer that has (re)gained popularity is that the “institutional setting for technological change”<sup>32</sup> determines growth in a country. In this context, the “institutional setting” (or, institutions) refers to formal rules (constitutions, laws and regulations, political systems, etc.) and informal rules (value systems, beliefs, social norms, etc.) that humans use when interacting within a wide variety of repetitive and structured situations at multiple levels of analysis (e.g., North, 2005; Ostrom, 2005). Differences in the types and quality of institutions ultimately result in different growth patterns (Acemoglu et al. 2001, 2002, 2004; Keefer and Knack, 1997; Rodrik et al, 2004; Easterly and Levine, 2003; and Bardhan, 2005).

A second influential and not so new theory, which draws on the works of Montesquieu and Weber<sup>33</sup> among others, states that these differences are due to an exogenously given set of geographic factors which are conducive for growth. Thus factors such as soil quality, weather conditions, and natural resource availability affect productivity and attitudes directly (Diamond, 1997; Gallup et al.1998; Sachs, 2001; McArthur and Sachs, 2001; Masters and McMillan, 2001). A third account sees differences in economic growth arising due to a country’s openness to trade, rather than as a result of its resource endowments or institutions. This stream of literature suggests that countries that follow open trade policies gain more in terms of growth performance (Frankel and Romer, 1999; Dollar and Kraay, 2001; Sachs and Warner, 1995).

These three approaches have been identified in the literature as the “deep determinants” of growth, as opposed to more “proximate determinants” such as capital resource accumulation and technological change (Rodrik et al, 2004). For most part, these three strands of literature have evolved independent of each other. More recently, studies have attempted to determine which of these three factors are most important (or “primary”) in accounting for economic growth and its difference among countries (Rodrik et al, 2004; Basu, 2008; Easterly and Levine, 2003).

In this paper, we attempt to re-examine the impact of institutions, trade and geography on economic growth and development using the Rodrik et al.(2004) framework. Our contribution is to use more comprehensive measures for economic growth, for institutions, and for geography, along with the more commonly used measure of trade openness. Economic growth (or development), institutions and geography are all multifaceted phenomena, and to capture them in a single indicator, as is often done in the empirical literature that we referred to above, is problematic. For example, economic growth is often captured by the level of GDP per capita, rather than the growth rate of that variable, and other aspects of growth, such as investment or structural change, remain poorly represented. Institutions are often captured by a single indicator, such as a

---

<sup>32</sup> Darwin, J., (2008) “After Tamerlane”, Bloomsbury Press. Pg 190-191

<sup>33</sup> Montesquieu C de S, (1748), “The Spirit of the Laws”, Paris

measure for “rule of law”. However, we know that the institutional setting has many more dimensions, and works on the factors that cause economic growth in many different subtle ways. The same holds for geography. This is often captured by a single variable such as distance from the equator, but there are obviously many more and different aspects to geography. Our aim is to capture these multi-faceted phenomena in a more comprehensive way, and see whether, in this way, we can confirm the causal links between them that have been found in the existing literature.

The rest of the paper is organized as follows: Section 2 briefly reviews related literature outlining the inter-relationship between institutions, trade and geography. Section 3 describes the econometric methodology and Section 4 describes the data used. Section 5 presents the results, and Section 6 concludes the paper.

## 2. Related Literature

### 2.1. “Institutions matter”

The quality of institutions as a determinant of economic growth and development is not a new theory. The revival of interest in this area of research is often attributed to Douglass North, whose definition of “institutions” is often the benchmark for what constitutes them.<sup>34</sup> Building on the theoretical literature, there have been numerous empirical studies using institutions as an explanatory variable in the past two decades. Among the many reasons that have boosted interest in this area of research is the development of various econometric methods that have enabled us to deal with problems such as endogeneity, reverse causality etc., which are often encountered in economic research<sup>35</sup>, namely the 2SLS-IV methods and those that followed it.

One such influential paper that highlights the importance of institutions in studying economic growth, while correcting for endogeneity problems, is Acemoglu et al.(2001). They make a case for differences in economic institutions as the “fundamental cause of different patterns of economic growth”. They introduce ‘settler mortality’<sup>36</sup> as an instrument variable to deal with the issue of endogeneity. Their results showed that institutions in fact, had a large impact on economic growth while geography did not matter (at least not in a direct way). An important critique of this paper is that the instrument they used is not universally applicable because not all poor countries were colonized.

Hall and Jones (1999) use the percentage of population speaking English (ENGFRAC), and percentage of population speaking a European language (EURFRAC) as instruments for institutions that are in turn measured by the ICRG survey indicators. They use distance from the equator to capture geographic characteristics. They conclude that differences in social infrastructure had a large impact on cross-country differences in income

---

<sup>34</sup> There have been many different propositions on measuring institutions. Among the more common ones include measures of corruption and democracy by various organizations.

<sup>35</sup> Vieira, F. V., and Damasceno, A. O. (2009).

<sup>36</sup> Settler mortality uses data of mortality rates of soldiers, sailors, priests/bishops (European-settlers) in their colonies per 1000 settlers.

per capita. Other studies which use some measure of ‘Western European influence’ as an instrument for institutions include La Porta et al. (1997, 1998, 1999) Mauro (1995). La Porta et al. use legal origin as a proxy for institutions and find that common law countries perform best, and civil law countries perform worst. Mauro uses ethno-linguistic divisions as a proxy for corruption and finds that corruption lowers investment and thereby growth in countries.

Easterly and Levine (2003) estimate a 2SLS-IV model using settler mortality as an instrument for institutions and distance from the equator, other geographic dummy variables such as landlockedness, crops, minerals etc to capture geography. They find that geography has an indirect effect on income differences via institutions. In their words institutions “seem to be a sufficient statistic” to account for economic growth. They do not find any direct impact of trade openness either.

Concluding, we can summarize that institutions are seen as an important determinant of economic growth, but also that they are themselves endogenous. Most of the instruments that have been proposed for institutions are somehow related to geography, which points to the fact that geography and institutions are themselves related. This means that one must indeed revert to econometric methods such as instrumental variables in order to separate the impact of these two long-run determinants of economic growth.

## **2.2. “Geography and Economic development”**

The view that the geography of a country determines its economic prosperity dates back at least to Montesquieu.<sup>37</sup> There are two key versions of this hypothesis. The first states that productivity of technologies of the temperate climate zones are higher compared to technologies of the tropics.<sup>38</sup> A second view states that the disease environment is a burden to the country, and therefore, is directly linked to the poverty that is prevailing in it. This view states that due to the high disease prone climate in a country, productivity of workers in that economy will be low. For example, the occurrence of frost kills pests and other disease causing insects or bacteria. Many tropical countries do not have many days of frost in a year. Also, the temperatures soar high enough to not allow workers to work in agricultural fields in some seasons. All this affects growth in an economy. This outlook has become popular over the past decade and has been treated in empirical studies such as Sachs (2001) and Gallup and Sachs (2001).

While this view seems more plausible than the previous one, the causality of poverty in the wake of the disease environment and economic development is not very clear. McArthur and Sachs (2001) conclude that institutions and geographically-related variables affect the GNP per capita of countries. However, if this were the main cause of growth or poverty in countries, then there would be no escape from being poor, and this contra-

---

<sup>37</sup> Montesquieu, C de S., “The Spirit of the Laws”.

<sup>38</sup> Myrdal, Gunnar (1968) *Asian Drama – An Inquiry into the Poverty of Nations*. Vol.3 (Harmondsworth, Penguin).

dicts the fact that some countries with “bad geography” have in fact developed to high income levels.

Like in the case of institutions, the empirical literature has examined the importance of geography as a determinant of growth. Some of the variables used to study this hypothesis include the latitude of the country (Sala-i-Martin, 1997a, b; Bloom and Sachs, 1998; Rodrik et al, 2004), and this usually results in a positive and significant impact of the latitudinal position of a country on its growth levels. Using disease ecology as the independent variable measured by proportion of population at risk of malaria transmission, life expectancy at birth, and infant mortality rate, a negative and significant relationship between a geographically based health measure and GNP per capita is also found (Easterly and Levine, 2003; and Sachs, 2003). The number of frost days in a year, availability of arable land, the length of the coastline, rainfall, and temperature are other variables that have been used to study the effect of geography of the country on its growth rates (Bloom and Sachs, 2003; Masters and Sachs, 2001; and Bloom et al, 2003). Here, a positive and significant effect of the first three variables on growth is found, while the last two variables are found to have a negative and significant effect. Landlocked economies have been found to have lower growth (Easterly and Levine, 2001; Sachs and Gallup, 1999).

### **2.3. “Trade promotes growth”**

International trade openness has often been cited as increasing productivity, more so in less developed countries, allowing them to catch-up or converge towards the income level of richer nations. Policy recommendations usually involve adoption of more open or liberal trade policies in order to boost growth. However, like institutions, openness may be an endogenous variable, and hence a proper test for the impact of openness on growth must account for this.

Frankel and Romer (1999) use a gravity model to propose a new instrument for overall trade share of countries as a function of geographical distance from its partners. They use this instrument in a 2SLS-IV estimation to test the relationship between trade and income level of a country. They find that overall trade of a country has a large and positive impact on income, but their result is “only moderately significant”.

Sachs and Warner (1995) find that those countries that have robust trade policies converge, but add that this is not a sufficient condition for growth, as overall macroeconomic policies should accompany the trade policies. Rodriguez and Rodrik (2001) find alternative evidence, however, claiming that there is no significant relationship between open trade policies and economic growth. They further question the methodology used in Sachs and Warner (1995) and Edwards (1998), among others claiming that the measure of openness used tends to give a biased positive and significant relationship, and furthermore is open to different interpretations.

Dollar and Kraay (2001) test and find “that trade has a strong positive effect on growth”. However, Easterly and Levine (2003) do not find any direct impact of trade on growth once they control for institutions. Rodrik et al.(2004) come to a similar conclusion regarding the impact of trade on growth.

From the empirical evidence and counter-evidence provided, it is clear that we do not have a clear winner in the “primacy” contest. Institutions, geography and trade openness all have their own supporters as the primal factor causing growth. Most of the empirical studies however, use a uni-dimensional measure of GDP per capita or growth rates to test for the above mentioned relationships. In this paper we test the hypothesis using a different measure of growth to see if the significance of institutions, trade or geography still holds.

### 3. Econometric Approach

Rodrik, et al.(2004) set out to estimate a regression model that encompasses all three “primal” factors behind economic growth, as we briefly surveyed them in the previous section. By including all three factors in one model, they aim to determine which of the three factors dominates over the other<sup>39</sup>. Following their framework, we start with the following equation:

$$\log y_i = \alpha + \beta_1 \text{INS}_i + \beta_2 \text{OPENC}_i + \beta_3 \text{GEOSOIL}_i + \varepsilon_i \quad (1)$$

where  $\log y_i$  is the log of GDP per capita (in the year 2009, in international dollars) of country  $i$ , and  $\text{INS}_i$ ,  $\text{INT}_i$  and  $\text{GEOSOIL}_i$  are the measures of institutional quality, trade share and soil quality conditions of the country  $i$ , respectively, and  $\varepsilon_i$  is the usual random error term.

Our brief survey of the literature has already emphasized that institutions and trade openness are both likely to be endogenous variables, and hence we need to account for this in the estimations. Cross-section growth regressions have often used the two stage least squares, or instrumental variables (2SLS-IV) estimation technique to study differences in income across countries. This method is one way of dealing with endogeneity. Rodrik et al.(2004) also use the 2SLS-IV approach with instrumental variables (IV) for institutions and trade.

This approach amounts to first estimate the following regression equations:

$$\text{INS}_i = \mu + \beta_5 \text{SM}_i + \beta_6 \text{FR}_i + \beta_7 \text{GEOAREA}_i + \beta_8 \text{GEOSOIL}_i + \varepsilon_{\text{INS}_i} \quad (2)$$

$$\text{OPENC}_i = \gamma + \beta_9 \text{SM}_i + \beta_{10} \text{FR}_i + \beta_{11} \text{GEOAREA}_i + \beta_{12} \text{GEOSOIL}_i + \varepsilon_{\text{OPENC}_i} \quad (3)$$

where  $\text{SM}_i$  is settler mortality and  $\text{FR}_i$  is the constructed Frankel Romer instrument for trade share below.  $\text{GEOAREA}$  is a geographical variable that is related to the physical characteristics of countries. This variable will be presented and discussed below. Equation (2) estimates the institutional quality variable as a function of settler mortality, as well as the specific geography variables in our equation (1), and the Frankel-Romer instrument for trade. Equation (3) does the same for the trade openness. The predicted values from equations (2) and (3) can then be used instead of the  $\text{INS}$  and  $\text{INT}$  variables (respectively) in the estimation of equation (1).<sup>40</sup>

<sup>39</sup> This is our preferred model, since it takes into account all three hypothesis. We retain the IV approach due to the fact that it accounts for the endogenous nature of institutions with respect to the independent variables

<sup>40</sup> We also experimented with  $\text{GEOAREA}$  as an additional exogenous variable in equation (1), i.e., not as an instrument. This yields broadly the same results as we report below, although the instruments are somewhat weaker in this case.

As already briefly mentioned above, SM is available only for a limited group of countries (those that were colonized). Therefore, instead of using SM as an instrument, we also use two language variables, ENGFRAC (fraction of population speaking English) and EURFRAC (fraction of population speaking other European languages). This yields the following two equations for the first stage estimations:

$$INS_i = \mu + \beta_4 ENGFRAC_i + \beta_5 EURFRAC_i + \beta_6 FR_i + \beta_7 GEOAREA_i + \beta_8 GEOSOIL_i + \varepsilon_{INS_i} \quad (2a)$$

$$OPENC_i = \gamma + \beta_9 ENGFRAC_i + \beta_{10} EURFRAC_i + \beta_{11} FR_i + \beta_{12} GEOAREA_i + \beta_{13} GEOSOIL_i + \varepsilon_{OPENC_i} \quad (3a)$$

In order to estimate equations (1) to (3), or (1) to (3a), we first have to construct the FR instrument of trade openness as a share of GDP for the year 2006, as the data used for Frankel & Romer (1999) use data for the year 1985. We start by estimating the Frankel and Romer (1999) gravity equation where bilateral trade between countries  $i$  and  $j$  is determined by the geographical distance between the two countries, their respective areas and populations, pairwise country landlockedness, a dummy variable for whether a common border between the countries exists, as well as some interaction terms with the common border dummy. The FR instrument thus tries to obtain a measure of the “geographic component of countries’ trade”. The equation estimated to obtain FR is as follows:

$$\text{Log}(\tau_{ij}/\text{GDP}_i) = a_0 + a_1 \ln D_{ij} + a_2 \ln N_i + a_3 \ln A_i + a_4 \ln N_j + a_5 \ln A_j + a_6 \ln (L_i + L_j) + a_7 \ln B_{ij} + a_8 B_{ij} \ln D_{ij} + a_9 B_{ij} \ln N_i + a_{10} B_{ij} \ln A_i + a_{11} B_{ij} \ln N_j + a_{12} B_{ij} \ln A_j + a_{13} B_{ij} (L_i + L_j) + e_{ij}. \quad (4)$$

Where  $\tau_{ij}$  is the bilateral trade between countries  $i$  and  $j$  (measured as exports plus imports),  $\text{GDP}_i$  is the nominal GDP of the country,  $D_{ij}$  is the bilateral distance between the country pairs  $i$  and  $j$ ,  $A$  is the area of the country,  $N$  is the population of the country,  $L$  is the dummy for landlockedness of the country,  $B_{ij}$  is the dummy for a common border between countries  $i$  and  $j$ , and  $e_{ij}$  is the random error term. Results of this estimation are discussed in Section 5A.

## 4. Data

### 4.1. Construction of the variables

The International Country Risk Guide (ICRG), Corruption Perceptions Index (CPI), and the World Bank’s World Governance Indicators (WGI) are a few among the many different indicators available that measure institutional quality. The WGI have the advantage that they cover a wide range of institutional factors. They consist of six aggregated indicators, for six different aspects of governance and institutions. In this study, we use the Institutional Profiles Database (IPD)<sup>41</sup>, which offers us 367 elementary variables which are then aggregated into three-digit (133 variables) and two-digit (93 variables)

<sup>41</sup> Freely available at <http://www.cepii.fr/anglaisgraph/bdd/institutions.htm>

levels. This database allows us to capture the multi-dimensionality of institutions, and does not require us to work with a fixed weighting scheme for the underlying indicators as in the WGI.<sup>42</sup> Using this multi-dimensional approach is our key contribution. Rather than relying on a single measure of institutions (and growth, and geography), as Rodrik et al. (2004) do, we will test the growth – institutions relationship taking a broader perspective than just one aspect of institutions.

We start from the 2-digit level IPD data<sup>43</sup>, but since we have 118 observations (countries) and 93 variables, this would leave us very few degrees of freedom. We thus further aggregate these 93 variables, first to 12 new variables, and then we aggregate these 12 variables to two variables that are, each, ultimately used in the regressions as the INS variable from equation (1). This 2-stage aggregation of the IPD variables relies on (a spatial variant of) principal components analysis. In the first stage of this procedure, we start from the four “sectors” that the IPD database covers: (i) public institutions and civil society, (ii) goods and services markets, (iii) capital markets, and (iv) labour markets and labour relations. Each of the 93 aggregated variables falls into one of these sectors. For each of the sectors, we apply our spatial principal components analysis (SPCA), and produce a number of summary measures for each sector. This yields 12 new indicators. In the second stage, these 12 indicators are again summarized by SPCA, or by (regular) PCA. In the second stage, we maintain two components each for the PCA and the SPCA version, which gives us four summary variables for the institutions variable in total.

The reason we use this 2-stage procedure is that the sectors have an unequal number of variables in them, and without the two stages, the (S)PCA results would be influenced stronger by the sectors with more variables. The 2 stage procedure has already been applied to IPD by Verspagen (2012), but only for regular PCA (not SPCA). The methodology of the spatial version of PCA is described in Bhupatiraju et al. (2013). It differs from regular PCA in terms of the objective that it sets when producing the summary variables (components). In regular PCA, the objective is to maintain a maximum fit between the summary component and the original variables. Mathematically, this is achieved by looking at the correlation matrix of the original variables. The eigenvector that is associated with the largest eigenvalue of this matrix can be used as weights to construct the “best” summary measure (component).

In the spatial version of PCA (SPCA, Bhupatiraju et al, 2013), the objective is to produce a summary measure that produces maximum spatial correlation with itself. Spatial correlation is measured by the Moran coefficient, and is essentially a measure of how spatially concentrated the underlying phenomenon is. High spatial correlation means that a country with high values of a variable tends to be surrounded by other countries with high values of that variable. SPCA also works by taking eigenvalues and eigenvectors, but in this case the matrix on which this is performed is not a simple correlation matrix,

---

<sup>42</sup> Verspagen (2012) discusses the correlations between the WGI and IPD databases.

<sup>43</sup> For a detailed explanation of the database for 2009 see: Crombrughe, D de., Farla, K., Meisel, N., Neubourg, C de., Ould Aoudia, J., Szirmai, A., (2009), “Institutional Profiles Database III: Presentation of the Institutional Profiles Database 2009 (IPD 2009)” Les Cahiers de la DGTPE – n° 2009-14.

but instead some more complicated matrix product involving, among other things, the matrix of Moran coefficients between the original variables. We will discuss the differences between regular PCA and SPCA in more detail when we look at maps displaying the institutions variables below.

Rodrik et al. (2004) use distance from the equator as their preferred measure of geography. As in the case of institutions, this is a very partial indicator. The archipelago of Indonesia embraces the equator, and in that sense it is similar to a country like Congo, but obviously the two countries are a world of difference when it comes to geography and climatic characteristics. Therefore, we use a much wider range of variables that describe the physical geographic characteristics and the soil related characteristics of the economy.

The raw data for the geography variables are taken from Gallup et al. (1998). Their dataset contains many variables, related to physical characteristics such as access to waterways, soil quality, climate zone, etc. We use SPCA to summarize these data into two summary variables: one for the physical characteristics of a country (based on 15 underlying variables), and one for the soil quality (based on 16 variables). In the regressions, both variables are included, and denoted as GEOAREA and GEOSOIL, respectively.

It is not only on the institutions or geography side that we would like to extend the measurement horizon. Rodrik et al.(2004), as many others in the literature, proxy “growth” by the level of GDP per capita. This is justified by the long-run nature of the analysis. If growth rates are different between countries for a long enough period, the countries with high growth rates will also be(come) the ones with higher levels of income per capita. In practice, however, growth rates are variable over time, and there is no strong correlation at all between growth and the level of income. Therefore, in a strict sense, the regressions of Rodrik et al. (2004) have very little to say about growth, but instead tell us something about the correlation between income levels on the one hand and institutions, trade and geography on the other hand.

Therefore, we will extend our horizon with respect to the dependent variable in the regressions (equation 1) to include also a number of other variables related to economic growth and development. We include annual growth rates of GDP per capita, average of consumption, investment and government expenditure as a percentage of GDP, an indicator of openness measured as exports plus imports as a percentage of GDP and population growth rates. All these variables are taken over the period of 2000-2009. Like before, we will use SPCA to summarize these data into two summary variables, and both of these will be used as dependent variable in the estimation of equation (1).

Since we use cross-section data for the year 2009, we have to construct the Frankel Romer instrument (FR) for 2006 using equation 4. In order to construct this, we use the “square” gravity data set from Head et al.(2008) for bilateral distances, area and population and landlockedness data.<sup>44</sup> This data set contains bilateral trade flows obtained from the International Monetary Fund’s *Direction of Trade Statistics* (DoTS), as well as distance and size data. Bilateral distance between countries is measured using the great

---

<sup>44</sup> Freely available at <http://www.cepii.fr/anglaisgraph/bdd/gravity.htm>.

circle formula, between the biggest cities of the two countries. The data on area (measured in square kilometers), population, and dummies for landlockedness, as well as common borders, are also taken from the CEPII gravity data set compiled by Head et al. (2006).

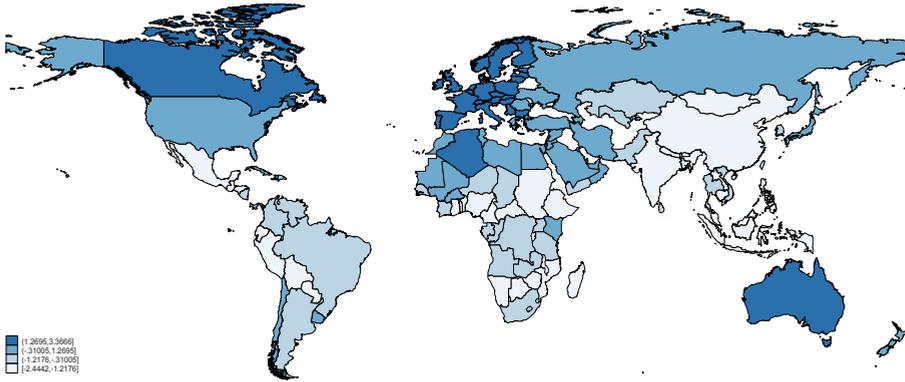
The other instrument that we use is settler mortality (SM), as first used by Acemoglu et al. (2001), from which source we also obtain these data. Using the above mentioned data we perform the 2SLS-IV regressions, the results of which are presented below. In order to include non-colonized countries in the analysis, we also use the Hall and Jones (1999) suggested ENGFRA (fraction of population speaking English) and EURFRA (fraction of population speaking other European languages) variables as a substitute instrument for MS.

## 4.2. Mapping the data

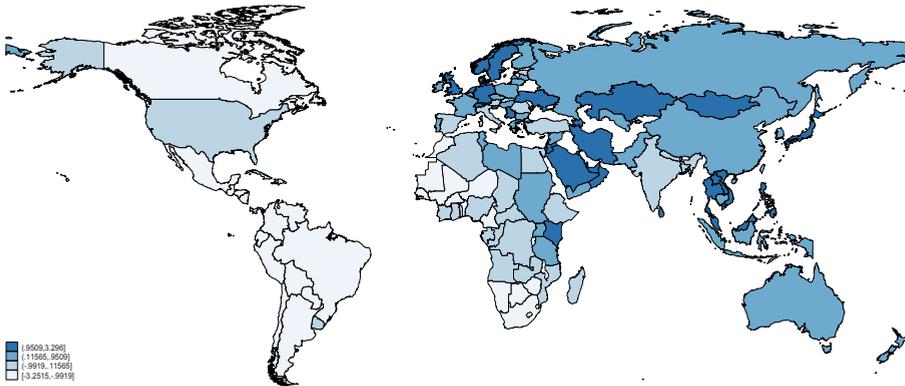
The summary variables for institutions, geography (both independent variables) and economic development (dependent variable) bring out the salient differences between countries in the underlying datasets that have many more variables. The SPCA procedure that we use to construct these summary variables is designed to interpret differences and similarities between countries in these underlying variables in a spatial (geographical) way. We opted for the SPCA (instead of the regular PCA) method because geography seems to play a large role in institutions. In the standard literature (summarized in Section 2 above), this is reflected in the fact that the instruments that are used for institutions almost always have a geographical nature. Settler mortality (one of the most used instruments) is very specific to particular environments, and depends on climate, diseases, etc. Language (another popular instrument) also has a clear geographical nature. Rather than just depending on the instruments to bring out the geographical correlation of institutions, we bring in geography in the construction of the variables, by means of the SPCA procedure.

We perform the SPCA analysis as detailed in Chapter 3, and we provide maps that illustrate the patterns for the main variables in the analysis. We start, in Map 1a and 1b, with the institutions variable.

These maps clearly bring out the spatial concentration of institutions. The first component is a clearly a measure of “Western” institutions, where we use “Western” for describing Europe and its offshoots (North America, Australia and New Zealand). All these areas are darkly colored (high values), with the exception of the USA, which scores somewhat lower. This first component broadly measures the institutional side which Verspagen (2012) has dubbed as the first stylized fact of the institutions – development relation, i.e., that developed nations tend to have the type of institutions that are highly valued in the “Western World”, such as democracy, civil rights, transparent justice systems and free markets. From our SPCA exercise, it appears that these kind(s) of institutions are strongly spatially concentrated at a global scale.



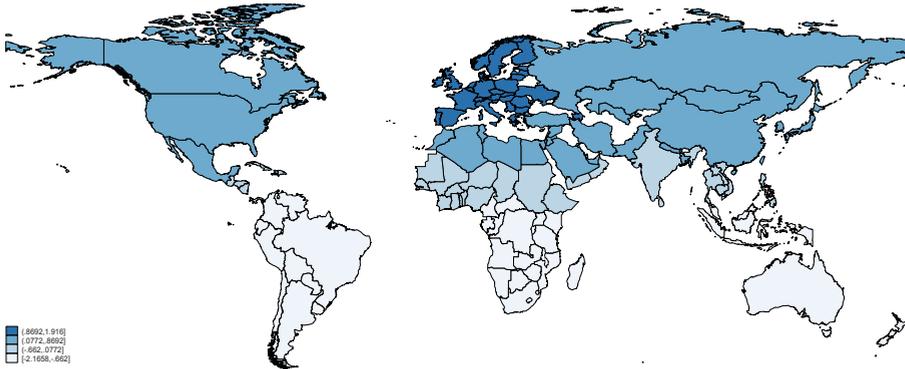
**Map 1a. Institutions from the IPD database, 2009, first component (INS1-SPCA)**



**Map 1b. Institutions from the IPD database, 2009, second component (INS2-SPCA)**

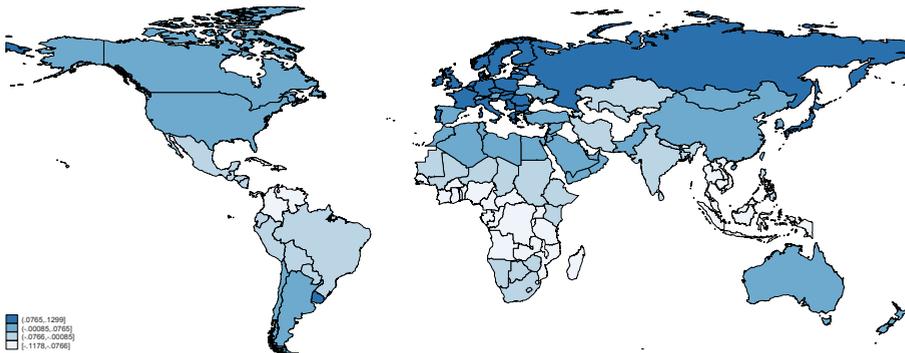
In Map 1b, we display the second component for institutions. This is again clearly spatially concentrated, but now the spatial pattern points to high values in the Eastern hemisphere, as well as in some parts of Europe. Broadly speaking, the kind of institutional features that are stressed in this second component are those that are associated to a stronger role of the government in social and economic life. Africa is an interesting area for this indicator, as it shows a broad East-West division in itself (similar to the global pattern). Summarizing, the two maps (indicators), give us two rather distinct views on institutions at a global level, and it will be interesting to see how these are correlated to growth in the estimation of equation (1).

Map 2 displays the first component for the physical characteristics dataset (GEOAREA). Here, the high values are concentrated on the northern hemisphere, which is because latitude has a high weight in the fact. On the Northern hemisphere, Europe scores particularly high and this is mainly due to the relatively high weight of access to waterways and the small size of a country.



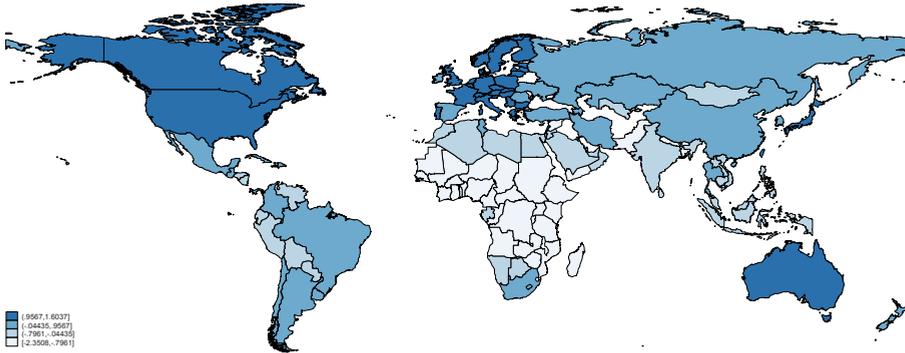
**Map 2. Physical geographical characteristics, first component**

In Map 3, the soil characteristics summary variable is displayed. This is the other geography related variable (GEOSOIL) that we include in the regressions as an independent variable. This variable is clearly also related latitude, but less so than the GEOAREA variable. However, latitude does not enter in this component, and the impression that latitude affects this component stems from the influence of climate. The climate zones that are weighted strongly negative are savannah climates, dry climates, humid subtropical climates, and Mediterranean climates.

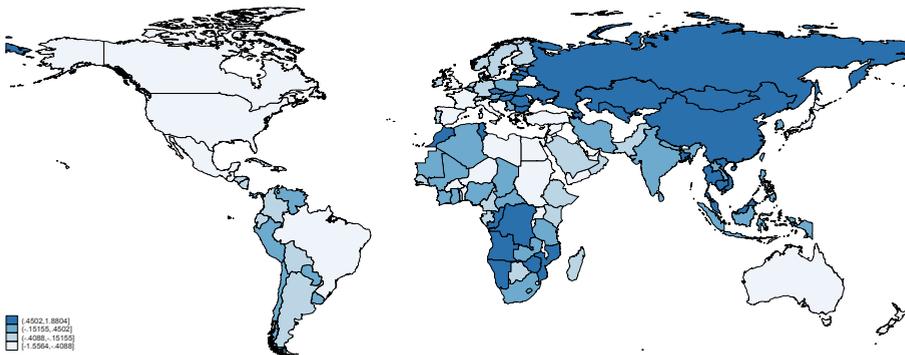


**Map 3. Soil characteristics, first component**

Finally, Maps 4a and 4b show the results of the SPCA procedure on the economic development variables. With the first factor here, in Map 4a, we capture the general development level. This variable is mostly influenced by the initial level of GDP per capita (in 2000), but also by the share of consumption in GDP and, negatively, by population growth. All of these phenomena (low population growth, high income and high consumption) are characteristics of developed countries. Therefore, this component captures the economic side of the stylized fact that was already referred to above, i.e., that developed countries tend to have developed institutions.



**Map 4a. Economic development, first component**



**Map 4b. Economic development, second component**

Map 4b shows the second factor from the economic development dataset. This captures economic growth in a much more direct way, but in combination with some of the other indicators. We see high values in Asia, South-Eastern Europe and Central/Southern Africa. What characterizes these areas, is a relatively low initial (2000) GDP per capita level, combined with a relatively high growth rate, high openness to trade, a high investment rate and low population growth. Clearly, the indicator captures an overall growth or development strategy, rather than just the growth rate of GDP per capita.

## Empirical Results

### A. Construction of the Frankel and Romer instrument

Following Helpman et al. (2008), we estimate a tobit model to take into account the missing and zero trade observations among the countries, rather than dropping those observations which have either no trade or missing trade values as in Frankel and Romer (1999). The estimated tobit equation yields broadly the same results as the OLS estimates. Table 1 shows the results of the estimated bilateral trade gravity coefficients.

**Table 1: Bilateral Trade Equation (Tobit)**

VARIABLES	(1)	(2)	(3)
	logTij_GDPi	logTij_GDPi	sigma
Dij	-1.054*** (0.0251)	-1.054*** (0.0250)	
Ni	-0.0958*** (0.0161)	-0.0959*** (0.0161)	
Ai	-0.00741 (0.0146)	-0.00736 (0.0146)	
Nj	0.803*** (0.0162)	0.803*** (0.0162)	
Aj	-0.170*** (0.0144)	-0.170*** (0.0144)	
Lij	-0.681*** (0.0318)	-0.681*** (0.0318)	
B	1.033 (1.182)	1.032 (1.181)	
BDij	0.316 (0.215)	0.316 (0.215)	
BNi	0.131 (0.108)	0.131 (0.108)	
BAi	-0.320*** (0.117)	-0.320*** (0.117)	
BNj	0.0688 (0.107)	0.0687 (0.107)	
BAj	0.0753 (0.117)	0.0754 (0.117)	
BLij	0.690*** (0.165)	0.689*** (0.165)	
Constant	6.575*** (0.286)	6.576*** (0.285)	1.922*** (0.0121)
Observations	12,644	12,644	12,644
R-squared	0.339		

Notes: Standard errors in parentheses\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The dependent variable is  $\log(T_{ij}/GDP_i)$ . Column (1) gives the results of the OLS estimation, while Columns (2) and (3) report the Tobit estimation.

The results are slightly different from those of Frankel & Romer (1999), especially for those that involve the interaction terms. We find that distance has a large, negative and significant impact on bilateral trade as expected. Bilateral trade is strongly increasing in  $j$ 's population and decreasing in  $j$ 's area and  $i$ 's population, at a lower magnitude. Country  $i$ 's area is the only non-significant variable among these. These results show that country  $i$  would trade more with countries that are closer to it, and those that have a large population but smaller areas.<sup>45</sup> The results also show that if either country is landlocked, trade between them is strongly negatively affected and this is significant. Sharing a border has a large effect on trade; however this estimate is not significant. This is in agreement with what Frankel and Romer (1999) find. According to them a possible explanation for this insignificance is because “only a small fraction of country pairs share a border”, and therefore are imprecisely estimated. Taking into account a common border between trading partners gives us different results. However, only two of the interaction terms are statistically significant. Country  $i$ 's area across a shared border is negatively significant -0.32 as opposed to -0.007 when the common border is not taken into account, while landlockedness of either country across a shared border is positively significant at 0.69 as opposed to -0.681 when a common border is not taken into account, showing that geography determines trade partially. Out of the 12644 observations only one observation is left censored at -10.67.

Using these estimated results, we obtain the fitted values of  $T_{ij}(\hat{T}_{ij})$ . We aggregate these values over  $j$  to get our instrument  $\hat{T}_i$ . This is the “geographic component” of country  $i$ 's total trade. It should be noted here that correlation between the actual and the constructed trade share is 0.58 after controlling for size, which is a rather moderate correlation. This could be due to the imputed values of the geographically constructed bilateral trade data for which values are either missing or zero.<sup>46</sup> Since the imputation is based on the available geography data, Taiwan's trade share is not estimated due to missing GDP values.

## B. Estimation of the RST model

Along with the above constructed FR instrument ( $\hat{T}_i$ ) we also use the settler mortality instrument (SM), ENGFRAC and EURFRAC for the estimation of equation (1). As mentioned above, the IPD database has 118 countries that are used as the basis for the analysis. We perform the estimations on two sub-sets of these 118 countries. The smaller set has 52 countries for which the SM data is available, and the larger set has 104 countries for which ENGFRAC and EURFRAC data are available. We estimate and present the OLS as well as the 2SLS-IV estimations.

Table 2 below shows the results of these estimations. Columns (1) and (2) give the results of the OLS estimation of Equation 1, using the two institutional components in separate equations. We find that GDP per capita is positively and significantly related to

<sup>45</sup> The results are interpreted as elasticities since they are log-transformed values.

<sup>46</sup> Noguera and Siscart (2005).

GEOSOIL, and to both institutions variables. Columns (3) to (6) show the results of the 2SLS-IV estimation, with settler mortality as an instrument in column (3) and fraction of English and European languages as instruments in columns (4) to (6). The regression with settler mortality as an instrument has fewer observations (due to missing data for countries that were not colonized), and generally show less significant results. In column (3) none of the variables is significant. The 2SLS-IV estimates in columns (4) to (6) always yield a significant impact of institutions. The first component for institutions has a positive and significant impact in GDP per capita, the second component is negative, but only significant if it is entered in an equation without the first institutions component. OPENC and GEOSOIL also enter significantly in some of the equations, but these effects are not robust across the equations. OPENC has a negative and significant sign in equation (4), which is against expectations. INS1-SPCA and INS2-SPCA are the first and second institutional components of an SPCA of the IPD 2009 data.

**Table 2. Regression results with GDP per capita (2009) as dependent variable, Institutions with SPCA**

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	logy (OLS)	logy (OLS)	logy (IV-1)	logy (IV-2)	logy (IV-2)	logy (IV-2)
GEOSOIL	6.402*** (1.976)	10.44*** (1.074)	33.14 (24.78)	-8.156 (7.576)	15.15*** (1.826)	-4.335 (5.149)
OPENC	0.00132 (0.00253)	0.00211 (0.00239)	-0.0284 (0.0411)	-0.0306** (0.0131)	0.0168 (0.0132)	-0.0159 (0.0124)
INS1-SPCA	0.323*** (0.0977)		-3.137 (2.869)	1.364*** (0.476)		1.222*** (0.309)
INS2-SPCA		0.167** (0.0779)			-0.508** (0.233)	-0.314 (0.209)
Constant	8.865*** (0.228)	8.798*** (0.225)	9.452*** (2.807)	11.51*** (1.080)	7.547*** (1.128)	10.27*** (1.056)
Observations	118	118	52	104	104	104
Underid. test (reject $H_0$ )			N	Y	Y	N
Weak instruments			Y	Y	Y	Y
Hansen J statistic (reject $H_0$ )			N	N	Y	N

Notes: Robust standard errors in parentheses \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Columns labeled OLS are OLS regressions, columns labeled IV-1 use settler mortality as instruments, and columns labeled IV-2 use ENGFRAC and EURFRAC as instruments. The row “Underid. test (reject  $H_0$ )” reports results of Kleibergen-Paap rk LM test, a N indicates that under-identification is an issue. The row “Weak instruments” reports results of Kleibergen-Paap rk Wald F test, Y indicates that the instruments are weak. The row “Hansen J-statistic (reject  $H_0$ )” reports results of the Hansen J test for over-identification. A Y indicates that over-identification is an issue.

The results for the first stage regressions for Table 2 and further tables are given in the Appendix. We did perform tests for under-identification, over-identification and weak instruments. These results are reported at the bottom of the table. We used a 10% cut-off value for deciding on these tests. In these tests, we generally (also for the estimations below) find that the instruments are weak. This is a common problem with IV estimations, and was also reported as an issue in the original Rodrik et al. (2004) paper. Our tests do indicate that institutions and trade openness are endogenous variables hence we

cannot go on the OLS results alone. However, finding strong enough instruments for these variables is a problem that the literature still needs to address.

In terms of under- or over-identification, equation (4) in Table 2 performs best. It passes both tests (but the test for weak instruments), while equations (3) and (6) do not pass the under-identification test, and equation (6) does not pass the over-identification test. Thus, equation (4) provides the best model for the general development level in Table 2. What does this table suggest in terms of the primacy of institutions in terms of explaining the level of GDP per capita? Certainly, the results in the table point to an important role of institutions in explaining the development level (GDP per capita). The two institutional variables are the only ones that are systematically significant in the 2SLS-IV regressions. The results for trade and geography are sometimes significant, which would point to an independent role of geography and/or trade, but this does not seem to be robust. We also find evidence that there are multiple sides to institutions, as both the first and the second component that we derived from the SPCA are significant. The first component seems to dominate (column 6), and the sign of the second component is negative, which leads us to conclude that what is related to the development level of a country is the degree to which institutions are of the “Western” type.

**Table 3. Regression results with GDP per capita (2009) as dependent variable, Institutions with regular PCA**

VARIABLES	(1) logy (OLS)	(2) logy (OLS)	(3) logy (IV-1)	(4) logy (IV-2)	(5) logy (IV-2)	(6) logy (IV-2)
GEO SOIL	11.54*** (1.067)	11.40*** (1.066)	8.000 (5.657)	13.79*** (2.247)	11.99*** (1.579)	14.48*** (3.347)
OPENC	0.00316 (0.00247)	0.00335 (0.00251)	-0.0186 (0.0250)	-0.0132 (0.0102)	-0.00150 (0.00838)	-0.0160 (0.0137)
INS1-PCA	0.330 (0.346)		-7.425 (9.845)	5.265* (2.824)		6.341 (4.406)
INS2-PCA		-0.124 (0.294)			-2.347 (2.295)	2.149 (5.679)
Constant	8.709*** (0.228)	8.693*** (0.229)	10.43*** (1.548)	9.936*** (0.834)	9.065*** (0.702)	10.17*** (1.140)
Observations	118	118	52	104	104	104
Underid. test (reject $H_0$ )			N	N	N	N
Weak instruments			Y	Y	Y	Y
Hansen J statistic (reject $H_0$ )			N	N	Y	N

Notes: Robust standard errors in parentheses \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Columns labeled OLS are OLS regressions, columns labeled IV-1 use settler mortality as instruments, and columns labeled IV-2 use ENGFAC and EURFAC as instruments. The row “Underid. test (reject  $H_0$ )” reports results of Kleibergen-Paap rk LM test, a N indicates that under-identification is an issue. The row “Weak instruments” reports results of Kleibergen-Paap rk Wald F test, Y indicates that the instruments are weak. The row “Hansen J-statistic (reject  $H_0$ )” reports results of the Hansen J test for over-identification. A Y indicates that over-identification is an issue.

We now proceed to see whether and how these results change if we use a regular PCA method (instead of the spatial PCA) to summarize the institutions variables (INS1-PCA and INS2)-PCA. These results are in Table 3. The main difference as compared to Table 2 is that the significance of the institutions variables is now much weaker. This holds for both the OLS and 2SLS-IV results. In fact, the only time that one of the institutions variables is significant, is in column (4), and this is only weakly significant (10%). Similarly, OPENC is never significant in Table 3. However, GEOSOIL is significant in all columns but equation (3).

All of the columns (3) to (6) in Table 3 suffer from under-identification. This can be interpreted as saying that the instruments are not good enough. With institutions that are not summarized in a geographical way, finding good instruments seems even more of a problem than in Table 2. In addition, equation (5) also seems to suffer from over-identification.

The conclusion seems to be that it makes a difference whether we weight the underlying institutional variables from IPD in a spatial way, or without spatial weights. Non-spatially weighted institutions are difficult to instrument, and (therefore) do not seem to have a robust impact on the development level, while spatially weighted institutions do have such an impact. This could be because instruments themselves have an underlying spatial narrative. For example, settler mortality is available for those countries that were colonized. These colonized countries are mostly clustered in small groups in some parts of the world. As a preliminary conclusion, it seems fair to that institutions indeed are a (the) primal factor behind economic development, but only if we take into account that institutions themselves are strongly geographically concentrated.

In order to test the effect of institutions on the broader aspects of development, including the growth rate of GDP per capita itself, we also estimate the regression models with a different dependent variable. For this purpose, we used the first and second components of the broad economic development dataset as displayed in Maps 4a and 4b above. The results for the first factor were very similar to Tables 2 and 3, and therefore, we do not document these results. This is because the first component (Map 4a) mainly captures the development level, and hence is not too different from GDP per capita itself. Table 4 documents the results for the second factor (“growth”), using institutions with SPCA as the independent variable.

In the OLS regressions (columns 1 and 2), only the first institutions component enters significantly, but with a negative sign. OPENC also enters significantly twice, and GEOSOIL significant and negative once. Now, trade openness is significant in the 2SLS-IV results with settler mortality, but none of the other variables are significant (column 3). In the 2SLS-IV results with language instruments (columns 4 to 6), the second institutions component is never significant, but the first component remains significant and negative. GEOSOIL (twice positive and once negative) and OPENC (twice positive) are also significant. Thus, these results suggest that in terms of the more dynamic aspects of economic development, institutions are not the only “primal” factor that contributes. Geography and trade also contribute. Moreover, “Western” institutions seem to contribute negatively to development in these regressions.

Like before, the instruments are weak, and all columns except (4) seem to suffer from either over- or under-identification. Thus, equation (4) is the preferred specification in Table 4, and it points to a negative impact of institutions, a positive impact of trade openness and a positive impact of GEOSOIL. Institutions no longer trump the other factors when we consider a more dynamic picture of economic development.

**Table 4. Regression results with second economic development component (growth) as dependent variable, Institutions with SPCA**

VARIABLES	(1) dev2 (OLS)	(2) dev2 (OLS)	(3) dev2 (IV-1)	(4) dev2 (IV-2)	(5) dev2 (IV-2)	(6) dev2 (IV-2)
GEOSOIL	0.125 (1.074)	-1.508** (0.684)	-1.262 (6.180)	7.446* (3.804)	-2.528*** (0.806)	7.142** (3.318)
OPENC	0.00730*** (0.00171)	0.00661*** (0.00156)	0.0239** (0.0116)	0.0193*** (0.00709)	0.00191 (0.00578)	0.0181** (0.00755)
INS1-SPCA	-0.102* (0.0588)		0.225 (0.711)	-0.618** (0.240)		-0.606*** (0.214)
INS2-SPCA		0.00820 (0.0504)			0.121 (0.0957)	0.0250 (0.127)
Constant	-0.618*** (0.135)	-0.559*** (0.131)	-1.692** (0.798)	-1.689*** (0.577)	-0.238 (0.494)	-1.590** (0.625)
Observations	118	118	52	104	104	104
Underid. test (reject $H_0$ )			N	Y	Y	N
Weak instruments			Y	Y	Y	Y
Hansen J statistic (reject $H_0$ )			N	N	Y	N

Notes: Robust standard errors in parentheses \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Columns labeled OLS are OLS regressions, columns labeled IV-1 use settler mortality as instruments, and columns labeled IV-2 use ENGFRAC and EURFRAC as instruments. The row “Underid. test (reject  $H_0$ )” reports results of Kleibergen-Paap rk LM test, a N indicates that under-identification is an issue. The row “Weak instruments” reports results of Kleibergen-Paap rk Wald F test, Y indicates that the instruments are weak. The row “Hansen J-statistic (reject  $H_0$ )” reports results of the Hansen J test for over-identification. A Y indicates that over-identification is an issue.

In Table 5, we repeat the estimations of table 4, but with using the regular PCA method to construct the institutions components. As in Table 3, the results seem to suffer from under-identification (all columns 3 – 6). We have only one significant result for institutions in the table, in column (4), i.e., the 2SLS-IV results. As in Table 4, the sign of institutions is negative. In column (4), we also have a significant and positive effect of trade openness, but GEOSOIL, which was significant and positive in Table 4, is now significant and negative.

Thus, for the broad development indicator that includes growth, we find some similar, and some different results as compared to just using GDP per capita. What is similar is that we have stronger results for institutions when we weight institutions in a spatial way. What is different is, first, the sign of the institutions variable, and, second, whether or not (some of) the other primal factors still matter. Our analysis of the broader development variable both stresses that institutions as they are usually measured have an adverse effect on development, and that, besides institutions, geography and trade also matter.

**Table 5. Regression results with second economic development component (growth) as dependent variable, Institutions with regular PCA**

VARIABLES	(1) dev2 (OLS)	(2) dev2 (OLS)	(3) dev2 (IV-1)	(4) dev2 (IV-2)	(5) dev2 (IV-2)	(6) dev2 (IV-2)
GEOSOIL	-1.536** (0.613)	-1.434** (0.620)	1.136 (2.335)	-2.657** (1.213)	-1.615* (0.853)	-3.042 (1.878)
OPENC	0.00678*** (0.00156)	0.00663*** (0.00159)	0.0250** (0.0115)	0.0128** (0.00542)	0.00607 (0.00469)	0.0144* (0.00744)
INS1-PCA	-0.220 (0.196)		-0.712 (3.236)	-3.040** (1.471)		-3.637 (2.408)
INS2-PCA		0.131 (0.172)			1.385 (1.343)	-1.194 (3.075)
Constant	-0.573*** (0.129)	-0.561*** (0.130)	-1.813** (0.718)	-1.076** (0.439)	-0.572 (0.383)	-1.206** (0.613)
Observations	118	118	52	104	104	104
Underid. test (reject $H_0$ )			N	N	N	N
Weak instruments			Y	Y	Y	Y
Hansen J statistic (reject $H_0$ )			N	N	Y	N

Notes: Robust standard errors in parentheses \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Columns labeled OLS are OLS regressions, columns labeled IV-1 use settler mortality as instruments, and columns labeled IV-2 use ENGFAC and EURFAC as instruments. The row “Underid. test (reject  $H_0$ )” reports results of Kleibergen-Paap rk LM test, a N indicates that under-identification is an issue. The row “Weak instruments” reports results of Kleibergen-Paap rk Wald F test, Y indicates that the instruments are weak. The row “Hansen J-statistic (reject  $H_0$ )” reports results of the Hansen J test for over-identification. A Y indicates that over-identification is an issue.

## 6. Conclusions

In this paper, we test the Rodrik et al.(2014) framework to explain differences in development levels across countries by using a broader set of definitions for institutions, geography and economic variables. We use constructed indices for institutions, geography (divided into physical and soil quality) and economic variables (which take into account the variance in a set of seven variables related to the level of economic development/growth of countries).

We find from our analysis that institutions do play a primary role in explaining cross-country differences in income per capita. This part of our analysis largely confirms the Rodrik et al. (2004) results. We do find, however, that it matters how a summary variable is created from our underlying multi-dimensional institutions dataset. The results are reasonably strong if we summarize institutions in a spatial way (i.e., create a summary measure that is spatially correlated). When we summarize institutions without taking into account space or geography, the results are much weaker. This suggests that institutions are indeed strongly spatially correlated, and that it matters to acknowledge this in the Rodrik et al.-type regressions. Institutions may trump the other factors (trade and geography), but they are themselves strongly geographically influenced. In other words, even

if we do not see a direct impact of geography on GDP per capita, we should keep in mind that the institutions variables are spatially weighted and so ‘space’ does matter, thus we can conclude that spatial influence of neighbours has an impact on the economic performance of the economy.

We also estimated models in which we use a broader development variable. This is a composite index, produced by the spatial principal components analysis (SPCA) that was used to construct the institutions proxies. This broader economic development level captures growth in a more direct way, as it has high values for relatively underdeveloped, but fast-growing economies that are relatively open. With this variable as the dependent variable, we find that institutions, geography and openness all matter in the final equation. Institutions are no longer the only primal factor, geography and trade also matter. In addition, we find that the institutions proxy that we use has a negative influence. In other words, institutions as we usually measure them do not seem to contribute to growth of a dynamic development pattern in general.

The role of institutions in explaining the current level of income in various countries has been tested by various studies earlier. Untangling the historical and geographic reasons for the differences in growth is not an easy task. Our analysis shows that there is a need to analyze these differences using carefully constructed indices. For this line of research to go forward, we need to look at a wider choice of instruments which may prove to be better than those currently in vogue. For example the FR instrument is widely used as a proxy for trade share (as also in this paper), but the correlation between the constructed trade share and the actual trade share is only as much as 0.6. Settler mortality on the other hand tries to bring in a historical aspect into the explanation of the quality of institutions, but is restrictive in the sense that only colonized countries can be tested and as seen by many earlier studies and by this one, is a weak instrument, much like EUR-FRAC and ENGFAC.

## Appendix

### Variable description for Tables 1-5.

Variable description	Variable code
Distance between countries $i$ and $j$	$D_{ij}$
Area of country $i$ or $j$	$A_i, A_j$
Population of country $i, j$	$N_i, N_j$
Dummy for landlockedness of country $i$ and $j$ taken together.	$L_{ij}$
Dummy for a common border between country $i$ and $j$	$B$
Interaction term of a shared border and bilateral distance ( $D_{ij}$ )	$BD_{ij}$
Interaction term of a shared border and Population in $i$	$BN_i$
Interaction term of a shared border and Area in $i$	$BA_i$
Interaction term of a shared border and Population in $j$	$BN_j$
Interaction term of a shared border and Area in $j$	$BA_j$
Interaction term of a shared border and landlockedness dummy	$BL_{ij}$
GDP per capita (in the year 2009, in international dollars)	$\log y$
Measure of institutional quality (factor 1) using SPCA	$INS1\_SPCA$
Measure of institutional quality (factor 2) using SPCA	$INS2\_SPCA$
Measure of institutional quality (factor 1) using PCA	$INS1\_PCA$
Measure of institutional quality (factor 2) using PCA	$INS2\_PCA$
Alternative measure of economic development (Factor 2)	$dev2$
Measure of agricultural productivity conducive for growth	$GEOSOIL$
Measure of physical geographic factors conducive for growth	$GEOAREA$
Trade openness	$OPENC$
Instrument for trade openness (FR instrument)	$\log T_{ij\_GDPI}$
Fraction of population speaking English	$ENGFRAC$
Fraction of population speaking an European language	$EURFRAC$

First stage results for IV regressions:

(i) First stage IV regression results for Estimation (3) of Table 2

VARIABLES	(1) inst1	(2) openc
soill	9.340*** (1.895)	-111.6 (90.11)
settlemortality	0.000163 (0.000217)	0.000287 (0.0103)
logeTijhat	0.0170 (0.276)	28.52** (13.11)
area1	0.0738 (0.149)	-5.660 (7.076)
Constant	-0.412 (0.727)	-5.888 (34.59)
Observations	52	52
R-squared	0.378	0.167

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(ii) First stage IV regression results for Estimation (4) of Table 2

VARIABLES	(1) inst1	(2) openc
soill	11.69*** (1.567)	-8.732 (64.62)
engfrac	0.701 (0.464)	5.534 (19.13)
eurfrac	0.391 (0.248)	-4.297 (10.22)
logeTijhat	0.648*** (0.219)	29.76*** (9.017)
area1	0.223* (0.131)	-1.278 (5.394)
Constant	-1.923*** (0.632)	0.635 (26.06)
Observations	104	104
R-squared	0.725	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(iii) First stage IV regression results for Estimation (5) of Table 2

VARIABLES	(1) inst2	(2) openc
soil1	5.052** (1.999)	-8.732 (64.62)
engfrac	1.641*** (0.592)	5.534 (19.13)
eurfrac	-1.458*** (0.316)	-4.297 (10.22)
logeTijhat	0.719** (0.279)	29.76*** (9.017)
area1	0.0783 (0.167)	-1.278 (5.394)
Constant	-1.809** (0.806)	0.635 (26.06)
Observations	104	104
R-squared	0.436	0.174

Standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

(iv) First stage IV regression results for Estimation (6) of Table 2

VARIABLES	(1) inst1	(2) inst2	(3) openc
soil1	11.69*** (1.567)	5.052** (1.999)	-8.732 (64.62)
engfrac	0.701 (0.464)	1.641*** (0.592)	5.534 (19.13)
eurfrac	0.391 (0.248)	-1.458*** (0.316)	-4.297 (10.22)
logeTijhat	0.648*** (0.219)	0.719** (0.279)	29.76*** (9.017)
area1	0.223* (0.131)	0.0783 (0.167)	-1.278 (5.394)
Constant	-1.923*** (0.632)	-1.809** (0.806)	0.635 (26.06)
Observations	104	104	104
R-squared	0.725	0.436	0.174

Standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

(v) First stage IV regression results for Estimation (3) of Table 3

VARIABLES	(1) ins_1	(2) Openc
soil1	0.308 (0.612)	-111.6 (90.11)
settlermortality	1.20e-05 (7.00e-05)	0.000287 (0.0103)
logTijhat	0.0365 (0.0891)	28.52** (13.11)
area1	0.0221 (0.0481)	-5.660 (7.076)
Constant	-0.0206 (0.235)	-5.888 (34.59)
Observations	52	52
R-squared	0.019	0.167

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(vi) First stage IV regression results for Estimation (4) of Table 3

VARIABLES	(1) ins_1	(2) openc
soil1	-0.495 (0.471)	-8.732 (64.62)
engfrac	0.256* (0.139)	5.534 (19.13)
eurfrac	0.0332 (0.0745)	-4.297 (10.22)
logTijhat	0.119* (0.0657)	29.76*** (9.017)
area1	-0.0125 (0.0393)	-1.278 (5.394)
Constant	-0.328* (0.190)	0.635 (26.06)
Observations	104	104
R-squared	0.062	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(vii) First stage IV regression results for Estimation (5) of Table 3

VARIABLES	(1) ins_2	(2) openc
soil1	-0.504 (0.574)	-8.732 (64.62)
engfrac	-0.0776 (0.170)	5.534 (19.13)
eurfrac	-0.0381 (0.0908)	-4.297 (10.22)
logeTijhat	-0.0693 (0.0801)	29.76*** (9.017)
area1	0.0593 (0.0479)	-1.278 (5.394)
Constant	0.197 (0.232)	0.635 (26.06)
Observations	104	104
R-squared	0.024	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(viii) First stage IV regression results for Estimation (6) of Table 3

VARIABLES	(1) ins_1	(2) ins_2	(3) openc
soil1	-0.495 (0.471)	-0.504 (0.574)	-8.732 (64.62)
engfrac	0.256* (0.139)	-0.0776 (0.170)	5.534 (19.13)
Eurfrac	0.0332 (0.0745)	-0.0381 (0.0908)	-4.297 (10.22)
logeTijhat	0.119* (0.0657)	-0.0693 (0.0801)	29.76*** (9.017)
area1	-0.0125 (0.0393)	0.0593 (0.0479)	-1.278 (5.394)
Constant	-0.328* (0.190)	0.197 (0.232)	0.635 (26.06)
Observations	104	104	104
R-squared	0.062	0.024	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(ix) First stage IV regression results for Estimation (3) of Table 4

VARIABLES	(1) inst1	(2) openc
soil1	9.340*** (1.895)	-111.6 (90.11)
settlermortality	0.000163 (0.000217)	0.000287 (0.0103)
logeTijhat	0.0170 (0.276)	28.52** (13.11)
area1	0.0738 (0.149)	-5.660 (7.076)
Constant	-0.412 (0.727)	-5.888 (34.59)
Observations	52	52
R-squared	0.378	0.167

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(x) First stage IV regression results for Estimation (4) of Table 4

VARIABLES	(1) inst1	(2) openc
soil1	11.69*** (1.567)	-8.732 (64.62)
engfrac	0.701 (0.464)	5.534 (19.13)
eurfrac	0.391 (0.248)	-4.297 (10.22)
logeTijhat	0.648*** (0.219)	29.76*** (9.017)
area1	0.223* (0.131)	-1.278 (5.394)
Constant	-1.923*** (0.632)	0.635 (26.06)
Observations	104	104
R-squared	0.725	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(xi) First stage IV regression results for Estimation (5) of Table 4

VARIABLES	(1)	(2)
	inst2	openc
soil1	5.052** (1.999)	-8.732 (64.62)
engfrac	1.641*** (0.592)	5.534 (19.13)
eurfrac	-1.458*** (0.316)	-4.297 (10.22)
logeTijhat	0.719** (0.279)	29.76*** (9.017)
area1	0.0783 (0.167)	-1.278 (5.394)
Constant	-1.809** (0.806)	0.635 (26.06)
Observations	104	104
R-squared	0.436	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(xii) First stage IV regression results for Estimation (6) of Table 4

VARIABLES	(1)	(2)	(3)
	inst1	inst2	openc
soil1	11.69*** (1.567)	5.052** (1.999)	-8.732 (64.62)
engfrac	0.701 (0.464)	1.641*** (0.592)	5.534 (19.13)
eurfrac	0.391 (0.248)	-1.458*** (0.316)	-4.297 (10.22)
logeTijhat	0.648*** (0.219)	0.719** (0.279)	29.76*** (9.017)
area1	0.223* (0.131)	0.0783 (0.167)	-1.278 (5.394)
Constant	-1.923*** (0.632)	-1.809** (0.806)	0.635 (26.06)
Observations	104	104	104
R-squared	0.725	0.436	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(xiii) First stage IV regression results for Estimation (3) of Table 5

VARIABLES	(1) ins_1	(2) openc
soil1	0.308 (0.612)	-111.6 (90.11)
settlermortality	1.20e-05 (7.00e-05)	0.000287 (0.0103)
logeTijhat	0.0365 (0.0891)	28.52** (13.11)
area1	0.0221 (0.0481)	-5.660 (7.076)
Constant	-0.0206 (0.235)	-5.888 (34.59)
Observations	52	52
R-squared	0.019	0.167

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(xiv) First stage IV regression results for Estimation (4) of Table 5

VARIABLES	(1) ins_1	(2) openc
soil1	-0.495 (0.471)	-8.732 (64.62)
engfrac	0.256* (0.139)	5.534 (19.13)
eurfrac	0.0332 (0.0745)	-4.297 (10.22)
logeTijhat	0.119* (0.0657)	29.76*** (9.017)
area1	-0.0125 (0.0393)	-1.278 (5.394)
Constant	-0.328* (0.190)	0.635 (26.06)
Observations	104	104
R-squared	0.062	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(xv) First stage IV regression results for Estimation (5) of Table 5

VARIABLES	(1) ins_2	(2) openc
soil1	-0.504 (0.574)	-8.732 (64.62)
engfrac	-0.0776 (0.170)	5.534 (19.13)
eurfrac	-0.0381 (0.0908)	-4.297 (10.22)
logeTijhat	-0.0693 (0.0801)	29.76*** (9.017)
area1	0.0593 (0.0479)	-1.278 (5.394)
Constant	0.197 (0.232)	0.635 (26.06)
Observations	104	104
R-squared	0.024	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

(xvi) First stage IV regression results for Estimation (6) of Table 5

VARIABLES	(1) ins_1	(2) ins_2	(3) openc
soil1	-0.495 (0.471)	-0.504 (0.574)	-8.732 (64.62)
engfrac	0.256* (0.139)	-0.0776 (0.170)	5.534 (19.13)
eurfrac	0.0332 (0.0745)	-0.0381 (0.0908)	-4.297 (10.22)
logeTijhat	0.119* (0.0657)	-0.0693 (0.0801)	29.76*** (9.017)
area1	-0.0125 (0.0393)	0.0593 (0.0479)	-1.278 (5.394)
Constant	-0.328* (0.190)	0.197 (0.232)	0.635 (26.06)
Observations	104	104	104
R-squared	0.062	0.024	0.174

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**CHAPTER 5**  
**Multi-level Determinants of Inward  
FDI Ownership**

## 1. Introduction

The influence of firm idiosyncratic characteristics on their decision to enter foreign markets has been dealt with in much detail in both theoretical and empirical literature. There are a host of empirical studies that analyze firm level determinants of FDI decisions as well as their modes of entry. Most often, these studies analyze firm specific advantages such as technological intensity, innovation capacity, product and process differentiation from domestic firms, capital intensity, R&D activities, advertising (marketing and distribution skills) etc to explain the entry of Multinational Enterprises (MNE's) into host countries (Blonigen 2005, Lee and Rugman 2012). The above mentioned variables describe the advantages that certain firms tend to have over others.

When analyzing why some countries are better able to attract more FDI than others, studies look at the prevailing macroeconomic situation in the country, as well as how conducive the institutional set up is, for entering and doing business in that country. Exchange rate fluctuations, inflation variation, the availability or the lack of skilled labour and unskilled labour at lower wages, tax rates on corporate profits, etc are among some of the aspects that are considered under the macroeconomic stability of host countries that influence MNE entry decisions. A second set of factors that are considered are the bureaucratic, judicial and the political situation of a country. A country with a weak legal protection of property rights is unlikely to attract MNE, considering the high risk of appropriation that would prevail. An excessive amount of bureaucratic hurdles will also discourage new entrants, since it would increase the cost of setting up a business. These costs are further hiked if the bureaucracy is coupled with higher levels of corruption. Thus the institutional quality of a country and its credibility are often taken as important considerations for an MNE to enter a new market (Naude and Krugell 2007, Sharma and Bandara 2010, Seyoum 2011). Rule of law, democracy, low levels of corruption and flexible labour institutions are all said to be important country characteristics that most attract FDI inflows (Stein and Daude 2002, Daude and Stein 2007, Delbecque et al.2008, Pajunen 2008). The historical and current institutional set up of a country also influences its view on adopting foreign investment friendly policies. For example, while post-colonial Brazil welcomed and encouraged FDI to play an important role in its industrialization and development, a newly independent India restricted entry of foreign capital and preferred to look inwards, rather than looking towards FDI related growth and development policies, due to the differences in the colonial institutions of the two countries. (Baer and Sirohi 2013).

The impact of the quality of institutions on MNE decisions seems arguably simple and quite straight forward, but there is no clear consensus on how much and how relevant institutional quality actually impedes FDI inflows. Studies such as Wei (2000), Daude and Stein (2007), Dollar et al.(2005, 2006) and Kinda (2010) conclude that a weak institutional set up discourages FDI inflows into countries. On the other hand, studies such as Wheeler and Mody (1992), Singh and Jun (1995) do not find any significant evidence to conclude the same.

Dunning (1988, 2001) along with other studies (Rugman 1980, 1981) has emphasized that MNE's have firm-specific advantages that go hand-in-hand with host country conditions in encouraging FDI. FDI decisions are neither made solely based on firm characteristics, nor are they based entirely on host country conduciveness for business. Both these levels, independently and interactively form a certain environment that firms decide to take advantage of. In order to get a better understanding on what drives FDI inflows into countries, we need to analyze these hierarchical interactions in a more thorough fashion. Econometric studies that analyze what drives FDI inflows have either studied firm-level characteristics or have studied host country characteristics. To account for country level characteristics, these studies often used country dummies. In their paper "What matters to African firms?" Gleb et al.(2007) use the Enterprise Surveys to do a firm level analysis on firm perceptions of constraints as a function of firm characteristics, using sector and country fixed effects. They find that the constraints reported by firms strongly reflect country characteristics and their levels of income. Kinda (2010) analyzes the impact of infrastructure problems using firm level data and country fixed effects to find that they strongly discourage FDI inflows. Smarzynska and Wei (2000) use firm-level data and test if corruption in host countries has an impact on inward US FDI. They find that it does indeed discourage US firms from investing in countries with a higher level of corruption. Kaufmann and Wei (1999) using data from firm-level surveys, find that firms that bribe more, face more bureaucracy rather than less.

Some country level studies that analyze the determinants of FDI include: Al-Sadig (2009) who uses country level panel data to estimate the effects of corruption on the location of FDI using fixed effects to control for country specific effects on corruption. He finds that controlling for institutional quality, the level of corruption in the country does not have a strong impact in reducing FDI inflows. Anghel (2005) analyzes the effects of institutions on FDI inflows using cross-section data. They find that poor institutional quality significantly affects FDI inflows. Wernick et al. (2009) also find that FDI inflows into emerging and developing nations are dependent on strong institutions and business friendly macroeconomic policies and in turn influence economic growth. Seyoum (2009) also analyzes the impact of the quality of formal institutions in a country on inward FDI. He finds that they strongly influence inward FDI. Seyoum (2011) finds that the existing informal institutions in countries also have an indirect effect on attracting/discouraging inward FDI. Aristotelous and Fountas (1996) use pooled country level data to analyze macroeconomic determinants of FDI inflows into EU. They find that market size, and real exchange rate and tariff have a significant effect on inward FDI. Walsh and Yu (2010) use a GMM dynamic estimation on aggregate FDI inflows into developing countries as dependent on various macroeconomic, developmental and institutional variables. They find that inward FDI into the primary sectors are not as much affected by the various determinants, as are the secondary and tertiary sectors.

Very often, there is a tendency to extrapolate results found at the micro-level to a macro-level (Meentemeyer 1989). Results of macro-level or micro-level studies that study the various determinants of FDI tend to generalize their results across the hierarchical levels. For example Busse and Hefeker (2007) perform a cross-country analysis and find that the stability of governments and their democratic accountability, the presence of reli-

gious tensions, corruption and excessive bureaucracy, among others impact (total) inward FDI volumes. They conclude that “political risk and institutional indicators matter the most when multinational corporations confront decisions about where to invest in developing countries”. While aggregated FDI flows are at a macro level, firm decisions are much more micro-level phenomena, that are influenced by other micro-level as well as macro-level variables.

In order to see if indeed such generalizations could be made, we perform a multi-level analysis that takes into account determinants of FDI at both the macro-level and the micro-level, and the inter and intra level interactions. Farla (2013) uses a multilevel model estimate the impact of foreign equity ownership and institutional quality on a firm's investment decisions. The author finds that firms with 100% foreign equity do not invest in fixed capital, and also that marcoeconomic and institutional variables do not influence firm level decisions. The aim of this paper is to empirically analyze these multilevel interactions between firm-level advantages and host country characteristics that determine the amount of inward FDI in countries using multilevel model. In order to do this, we use the firm-level data from the Enterprise Surveys, which survey numerous developing countries, along with some macro-level variables collected from various sources, explained later in the paper. We specify a multilevel logit model to identify those firm-characteristics and host-country characteristics that raise the probability of inward FDI. The main results show that while firm-level perceptions of infrastructure and financial constraints affect FDI decisions, institutional quality measured at the country level does not have any significant impact on the same.

The rest of the paper is organized as follows. Section 2 describes the data used. Section 3 outlines the empirical methodology. Section 4 presents the results and the last section concludes the paper.

## **1. Data**

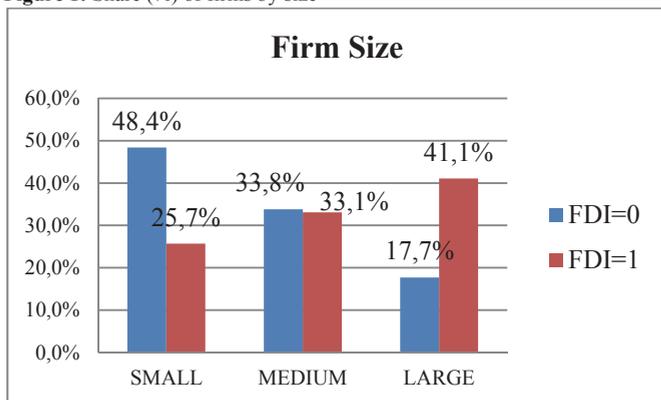
### **1.1. Micro data**

The analysis is based on the Enterprise Surveys data, collected and standardized in various developing countries over the years 2006-2011 by the World Bank. This comprehensive data set provides us a rich source of information on firm characteristics such as their size, ownership structure, firm perceptions of obstacles they face, such as corruption, telecommunications, electricity etc., as well as their sales and other production and finance variables. Compiled using country surveys during the period of 2006-2011, these data are a stratified random sample based on firm size, region and sector they belong to. The dataset analyzes 103 developing countries, with 20 countries surveyed twice during the time period 2006-2011. After omitting observations that have missing or incomplete data, we have 43,403 firms, of which about 5,010 firms have some level of FDI. Our

dependent variable for this analysis is foreign ownership (FDI)<sup>47</sup>. We use a dummy variable with a value of 1 if a firm has FDI and 0 otherwise. Approximately 12-13 percent of the firms surveyed had 10-100% of foreign ownership.

The independent variables of interest include firm structural characteristics and other characteristics. SIZE refers to the size of a firm measured by the number of employees a firm hires. It is a categorical variable ranging from small firms with 0-20 employees, medium firms with 20-99 employees to large firms with 100 or more employees. We take size dummies MEDIUM for (20-99 employees) and LARGE for 100 or more employees, with SMALL being the base. LNAME accounts for the age of the firm in logs of the number of years a firm has been established at the time of the survey.

**Figure 1:** Share (%) of firms by size



We construct index variables by performing Principle Component Analysis (PCA). The aim of doing a PCA is to reduce a number of correlated variables to a smaller number of latent variables that explain most of the variance in the data.

The survey captures the degrees of obstacles faced by firms, based on their perceptions. PHYSICAL 1 and PHYSICAL 2 are index variables capturing physical infrastructure obstacles faced by firms. These variables include constraints faced by firms with respect to (i) transportation of goods, supplies and inputs, (ii) electricity to current operations, and if they are accessible to clients and suppliers, using (iii) email and (iv) website. Transportation and electricity obstacles are categorical variables that take five values, namely, no major obstacle, minor obstacle, moderate obstacle, major obstacle or very severe obstacle. Email and web are dummy variables that account for firms either using or not using these facilities to communicate with clients and suppliers. They account for the lack of telecommunication infrastructure within a country<sup>48</sup>. In order to keep in line with the way the other obstacles have been defined, we define the email and web dum-

<sup>47</sup> To qualify as FDI, a firm should have at least 10% of their capital owned by foreign persons or organizations, as per the standard IMF definition.

<sup>48</sup> Of those firms which use emails to communicate with their clients and suppliers only 60 % of them also use a website as well, and of those firms which do not use emails about 3.3 % use a website.

mies such that if firms use email and web they are 0. PHYSICAL 1 captures the lack of internet facilities as a channel of communication between firms and their clients and suppliers, while PHYSICAL 2 captures transportation and electricity obstacles. Together these two components explain 72% of the variance in the data. We expect PHYSICAL 1 and PHYSICAL 2 to have a negative sign.

INFRA 1 and INFRA2 are index variables that capture the firm's perception on obstacles to both physical and financial obstacles all taken together. In other words, we perform a PCA using email, web, transportation, electricity and financial access (FINANCIAL). FINANCIAL is a categorical variable that measures how much of an obstacle access to finance is taking five values, namely, no major obstacle, minor obstacle, moderate obstacle, major obstacle or very severe obstacle. We expect obstacles to financial access to have a negative impact on FDI. We do this to see how all obstacles taken together affect the results. We begin the analysis using INFRA 1 and INFRA 2. We then proceed to using PHYSICAL 1, PHYSICAL 2 and FINANCIAL separately to assess their impact on FDI. Table 1a shows the results of the PCA's all index variables. INFRA 1 has a first component which explains 32% of the variance. All the variables in INFRA 1 have high and similar weights, thus representing an overall lack of infrastructural facilities that pose a hindrance to businesses. INFRA 2 explains 29.59% of the remaining variance. We expect INFRA 1 and INFRA 2 to have a negative sign as well.

INS is another index variable that captures the institutional obstacles firms face. This variable includes (i) corruption as an obstacle to current operations, (ii) crime, theft and disorder as an obstacle and (iii) property rights i.e., if the firm considers the court system as fair, impartial and uncorrupted. Here again, obstacles to corruption and crime are categorical variables taking five values ranging from no obstacle to severe obstacle. Property rights on the other hand is a categorical variable taking four values, namely, strongly disagree, tend to disagree, tend to agree and strongly agree. For INS we find that the first component explains about 51% of the variance. We use only the first component and this captures the quality of the institutional set up in which firms operate. We expect a negative coefficient on this variable.

**Table 1a: Principal Components of all index variables**

VARIABLES	INFRA 1	INFRA 2	PHYSICAL 1	PHYSICAL 2	INS
Email	0.5064	-0.4756	0.7024	-0.0002	
Web	0.5162	-0.4634	0.7008	0.0109	
Electricity	0.433	0.4212	0.0825	0.7078	
Transportation	0.3239	0.5527	-0.0932	0.7063	
Financial	0.4299	0.276			
Corruption					0.6496
Crime					0.5698
Proprietary					-0.5034
Eigen Value	1.60265	1.47968	1.52976	1.35277	1.54971
Explained %	0.3205	0.2959	0.3824	0.3382	0.5166

SKILL refers to the degree of obstacle firms face due to lack of skilled labour i.e., inadequately educated workforce<sup>49</sup>. This is a categorical variable with the same scale of five values ranging from no obstacle to very severe obstacle. This seems particularly important when we study FDI into developing countries. Although it is likely that foreign firms enter less developed countries for their cheap labour, inadequate education would pose a problem for firms. We therefore expect a negative relationship between SKILL and FDI. AGLOM measures the attractiveness of a sector-region for foreign firms to invest. This is constructed by taking the number of foreign firms present in a given sector in a given region. The Enterprise Surveys classify the region from within a country that the firm belongs to. For example, in Chile firms from the regions of Antofagasta, Bío Bío, Los Lagos, Santiago, Valparaíso are included in the survey. Similarly in Afghanistan firms from six regions, namely Hirat, Jalalabad, Kabul, Kandahar, Mazar and 5 Smaller Cities (Bamyan, Ghazni, Khost, Kunduz, Pul-E-Khumri) taken together are surveyed. AGLOM thus measures the attractiveness of these sector-regions to foreign firms. Table (A) in the Appendix gives the variable descriptions.

There are 15 broad sectors under which each firm has been classified in the data. In this chapter we re-classify these into 8 sectors as follows: 1) Textiles, Leather and Garments; 2) Food; 3) Metals and machinery, Auto and Auto parts; 4) Chemicals and Pharmaceuticals; 5) Electronics, Wood and furniture, Non-metallic and plastic materials and Other manufacturing; 6) retail and wholesale trade; 7) Hotels and restaurants and Other services and 8) Construction, transportation etc. The reclassification is done to avoid having too few observations in region-sector averages which will be discussed below.

## 1.2. Macro data

Some countries are able to attract more FDI than others. These cross country differences in FDI inflows come from differences in their macroeconomic and institutional condi-

<sup>49</sup> For some firms, the questions on the degree of obstacles do not apply. For these firms, it has been assumed that that particular obstacle is not an obstacle.

tions. When considering macroeconomic differences between countries, an obvious first choice is GDP per capita (we use the log of current USD prices taken from the World Bank database (2012)) which we refer to as LGDPPC<sup>50</sup>. As we study FDI inflows into developing countries, we are interested in the impact the general level of development has on the dependent variable. We thus take GDP per capita as an indicator of the level of development of a country. The level of GDP per capita reflects investment opportunities. Asiedu (2002) explains that GDP per capita and the risk of investment are inversely related, implying that the returns have to be increased in order to compensate this higher rate of risk. Since we use mainly developing countries we expect the sign to be negative. TAXRATE refers to the total tax rate as a percentage of commercial profits. This variable is taken from the World Bank database (2012). This tax rate measures the "amount of taxes and mandatory contributions payable by businesses after accounting for allowable deductions and exemptions as a share of commercial profits". Personal income tax, sales taxes, value added taxes and goods and services taxes are excluded. Host countries try to attract FDI by providing various incentives for firms to encourage investments in their economies. One such way of promoting FDI has been to reduce corporate tax rates. Studies such as Hartman 1984, 1985 find that transfer FDI<sup>51</sup> is not sensitive to tax rates, while Sato (2012), Bénassy-Quéré et al. (2005) etc, find a significant negative impact of the same on FDI. Devereux and Freeman (1995) on the other hand find that the country context tax influences FDI. Studies like those mentioned above, tend to take country or industry level tax rates to understand firm level decisions (Blonigen 2005).

Host country attractiveness is further enhanced by macroeconomic stability in markets. High inflation tends to increase uncertainty and is assumed to discourage FDI in general. Furthermore this uncertainty limits a firm's price setting plans in the long run (Sato 2012, Buckley et al.2007). INFLATION is measured by the growth rate of the GDP deflator, showing the rate of price changes in a year, taken from the World Bank database (2012). We expect a negative relationship between the two.

OPENK accounts for openness to trade measured by net exports as a percentage of GDP, taken from the Penn World Table 7.1 (Heston et al.2012). Trade openness measures the level of trade protection of a country. One expects openness to be positively correlated to FDI (Orts & Alguacil 2004, and Alguacil et al.2010). We use the financial openness indicator (KAOPEN) developed by Chinn and Ito (2008). KAOPEN measures the "extent and intensity" of capital controls of countries. Adams (2009) and Alfaro et al.(2010) among others argue that when host country financial markets are open, they allow for more efficient movement of foreign capital. Both these measures of openness are expected to have a positive effect on FDI. Additionally we control for LABOURPARTICIPATION, which is the labour participation measured by the proportion of working age population which is economically active (aged 15 years or older). Higher labour participation rates imply more willing workers, and thus a higher labour supply. A higher la-

---

<sup>50</sup> GDP per capita are taken according to the country-year of the Enterprise Surveys.

<sup>51</sup> Blonigen (2005) refers to FDI through new capital transferred to host countries as transfer FDI as opposed to FDI financed by retained earnings.

bour supply leads to wage competition effects that foreign firms then would want to take advantage of. We thus expect a positive relationship between labour participation and FDI inflows.

Institutional quality, especially in developing countries is likely to be an important determinant of FDI inflows. Good institutions tend to reduce transaction costs, by eliminating corruption, providing guaranteed property rights, ensuring justice etc. To capture the institutional conditions in a country, we use legal rights and business days. LEGAL-RIGHTS measures the degree of protection given to lenders and borrowers, given the country’s bankruptcy laws. BIZDAYS measures the number of calendar days required to start a business (even if it requires additional costs, the fastest number of days are taken). We expect that a better institutional environment in a country increases its attractiveness as a destination for FDI.

We also use the World Governance Indicators (WGI) to account for institutional quality at the country level as an indicator of the country institutional quality. The WGI include Voice and Accountability, Political Stability and Absence of Violence, Government, Regulatory Quality, Rule of Law, Control of Corruption<sup>52</sup>. Since these variables are highly correlated to each other, we do a PCA on the 6 world governance indicators. The first latent component of this PCA explains 79% of the variance. Table 1b shows the weights of each of the six variables in the first Eigen vector. This variable we call WGI. We can see that all the variables have similar weights, contributing almost equally to the scores.

**Table 1b: Variables of the WGI indicator**

Variable	WGI
Voice	0.409
Stability	0.341
effectiveness	0.423
Regquality	0.403
Ruleoflaw	0.437
corruption	0.43
Eigen Value	4.733
Expl. Proportion	0.79

<sup>52</sup> The definitions for each of the variables are in Table (B) the appendix.

## 2. Empirical Methodology

Our aim in this paper is to determine which characteristics influence if a firm invests in foreign markets or not. We also want to find out if these characteristics are conditional on the host country characteristics, i.e., why is one country able to attract more FDI than another country. A multi-level analysis is required when we know that a hierarchical structure exists in the data. As is apparent, firms are nested within countries and therefore firm-level decisions are usually not independent of country characteristics, be it policy/law related or just dependent on the current macro-economic situation in the country, which is the assumption of single level models (Goldstein 2003, Goedhuys and Srholec 2010, Srholec 2010). We therefore use a multi-level model<sup>53</sup> with firms at level-1 and country characteristics at level-2. This allows us to explain a part of the total variance by country context, with more accurate standard errors.

Since our dependent variable is a dummy variable, we have to account for this by using a multi-level logit model:

$$E(FDI_{ij} = 1 | \beta_j) = \varphi_{ij} \quad (1)$$

$$\text{Log} \left[ \frac{\varphi_{ij}}{1 - \varphi_{ij}} \right] = \beta_{0j} + \beta_1 X_{ij} + \beta_2 Z_j + u_j + e_{ij} \quad (2)$$

for firm  $i$  and country  $j$ , where  $\beta_0 + \beta_1 X_{ij} + \beta_2 Z_j$  is the deterministic part of the model and  $u_j + e_{ij}$  is the random part. Further,  $u_j \sim N(0, \sigma_u^2)$ ; and  $e_{ij} \sim N(0, \sigma_e^2)$ . From equation (2) we obtain different average outcomes for each group (country).

FDI is the binary dependent variable (that is always at the lowest level, i.e., firms). The level-1 predictors which are firm level variables given by  $X_{ij}$  include LNAGE, SIZE (MEDIUM, LARGE), AGLOM, INFRA, PHYSICAL, FINANCIAL, INS and SKILL. Level-2 predictors which are country level variables given by  $Z_j$ , include LGDPPC, TAXRATE, LABOURPARTICIPATION, INFLATION, OPENK, KAOPEN, LEGAL-RIGHTS, BIZDAYS and WGI.

The reasons for using a multi-level model as opposed to a regular fixed-effects model are also evident. If we use a non-hierarchical model, the statistical significance of level-1 variables is overestimated. Further, the standard errors of level-2 or group level variables will also be affected (Goldstein 2003). Unlike in the fixed-effects model where the groups are taken into account by taking dummies for each group, a multi-level model allows us to distinguish between observed and unobserved group characteristics (Rabe-Hesketh and Skrondal 2008).

Perceptions of firms tend to be endogenous, i.e. identical conditions can be perceived differently by different firms based on their size, ownership etc., (Kinda 2010). Gleb et al.2007 find that larger firms tended to complain more about physical infrastructure constraints, whereas smaller firms cited financial obstacles more frequently. In order to overcome this endogeneity (since the data is collected based on the perceptions of firms),

---

<sup>53</sup> These models are also known as random-coefficient or mixed effects models.

infrastructural obstacles are thus are instrumented by sector-region averages of these variables. Based on Dollar et al.2005, Aterido et al.2007 and Kinda 2010, these sector-region averages tend to be exogenous by default, since only 12% of our firms are FDI firms.

### 3. Results

We begin by analyzing the firm level determinants of FDI flows. Estimation (1) of Table 2 shows the 'null' or the 'empty' model. This model estimates:  $FDI_{ij} = \beta_{0j} + u_j + e_{ij}$  where  $\beta_{0j}$  represents the average over country fixed effects, and  $u_j$  and  $e_{ij}$  are the random effects at the country and firm level respectively. The purpose of this model is to investigate, without specifying any particular variables yet, how much of the FDI inflows can be attributed to the idiosyncratic characteristics of the firms themselves, and how much can be attributed to the country into which these investments flow. The 'random' part of the model, gives us these details. We can calculate the share of each level as a percentage of total variance. The variance partition coefficients (VPC's) measure the proportion of variation that lies at each level. When there are only two levels, the VPC's are the same as the inter class correlation coefficients (ICC's) which explain the correlation of observations within groups (countries), given by  $\rho$ . For a two level model the VPC and ICC at the firm level are given by  $ICC_e = VPC_e = \frac{\sigma_e^2}{(\sigma_e^2 + \sigma_u^2)}$  and the country level VPC and ICC are given by  $ICC_u = VPC_u = \frac{\sigma_u^2}{(\sigma_e^2 + \sigma_u^2)}$  where  $\sigma_e^2$  is the firm level variance of  $e_{ij}$  and  $\sigma_u^2$  is the variance of  $u_j$ . These ICC's allow us to see how much of the variance of FDI decisions can be attributed to country level variables as well as to the sectors.

Estimation (1) gives us an intra-class correlation of ( $\rho = 0.12$ ) to represent the variance attributed to differences between countries in receiving FDI inflows, i.e., the country context explains about 12% of FDI inflows. Although this is a relatively small part of the variance, it is not negligible and is significant at the 1% level. This implies that there are significant contextual differences between countries. We proceed further by introducing firm level explanatory variables in the remaining estimations of Table 2.

Following Kinda (2010) we first estimate the impact of INFRA on FDI prospects (estimation (2) in Table 2, and then estimate the impact of PHYSICAL and FINANCIAL infrastructure constraints separately in estimation (3). In Table 2 we do not instrument any of the variables. Estimation (2) shows the results of a fixed effects logit regression. As we can see, all the covariates are highly significant. We find that the age of the firm (LNAGE) has a negative effect, indicating that older firms are less likely to be foreign owned (partly or wholly). SIZE of the firm has a positive impact, meaning that larger firms are more likely to be foreign owned. AGLOM is strongly significant, implying that presence of other firms in the same sector in the region has a positive effect on attracting further FDI. This result is different from what Kinda (2010) finds in all regressions. This could be due to the fact that we use different sector categorizations. Further, INFRA 1 (infrastructure constraints) has a highly significant negative impact, implying that the

lack of email and web (internet) infrastructural constraints are not likely to discourage FDI inflows. This could be because firms which do not have email and web do not need it for their operations. While 86% of firms with FDI use email and 60% use websites. However, obstacles of electricity and transportation seem to have an unexpected positive sign. The coefficients on this variable are small (0.06). Among the firms that have FDI, about 24% of them think transportation is a major to severe obstacle and 37% think electricity is a severe obstacle. The positive coefficient on this variable may be interpreted as follows: the lack of infrastructure is dependent upon the level of overall development in a country. The rate of return on investment in countries with a lower level of development could be higher than in those countries. Estimation (3) shows the estimations of a fixed effects logit regression when the obstacles are divided into PHYSICAL (PHYSICAL 1 and PHYSICAL 2) and FINANCIAL infrastructure constraints also show similar effects. Obstacles to both PHYSICAL 1 and FINANCIAL infrastructure have a significantly negative effect on FDI ownership. The coefficient on PHYSICAL 1 shows that the availability of internet resources increases FDI, while the lack of financial access decreases FDI. PHYSICAL 2 again has a small but positive coefficient and is highly significant.

To check the robustness of the fixed effects results, we have random effects estimations in (4) and (5), which have results similar to (2) and (3) respectively. The interclass correlations for both the estimations is approximately 20% ( $\rho=0.20$ ). This is higher than before (in estimation 1) because now a part of the firm-level variance in equation (1) is captured by firm-level variables, and hence the firm-level variance due to unobserved firm-level variables is reduced. However, still the firm-level variance is much higher than country level variance, in turn showing that firm characteristics are more relevant than country characteristics in explaining inward FDI decisions of firms. However, we cannot disregard country effects in explaining FDI, showing that multi-level interactions should indeed be taken into account, while analyzing the determinants of FDI.

**Table 2: Estimations of firm level variables with out INS and SKILL**

	(1)	(2)	(3)	(4)	(5)
	FDI	FDI	FDI	FDI	FDI
VARIABLES		fixed effects	fixed effects	random effects	random effects
LNAGE		-0.28*** (0.02)	-0.28*** (0.02)	-0.28*** (0.02)	-0.29*** (0.02)
MEDIUM		0.7*** (0.04)	0.7*** (0.04)	0.7*** (0.04)	0.7*** (0.04)
LARGE		1.71*** (0.048)	1.69*** (0.048)	1.7*** (0.048)	1.69*** (0.048)
AGLOM		0.03*** (0.002)	0.03*** (0.002)	0.03*** (0.001)	0.03*** (0.001)
INFRA 1		-0.36*** (0.017)		-0.35*** (0.017)	
INFRA 2		0.254*** (0.016)		0.251*** (0.016)	
PHYSICAL 1			-0.40*** (0.018)		-0.39*** (0.018)
PHYSICAL 2			0.06*** (0.015)		0.063*** (0.015)
FINANCIAL			-0.18*** (0.013)		-0.18*** (0.013)
Constant	-1.91*** (0.07)			-2.34*** (0.12)	-2.03*** (0.122)
$\rho$	0.12			0.20	0.19
LR Chi2	1644.21			1969.04	1956.48
LR df(1)	***			***	***
# of firms	43,403	43,403	43,403	43,403	43,403
# of countries	103	103	103	103	103

Standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

All estimations include Sector dummies.

The Hausman test results in our case show that fixed effects models are preferred to random effects models (as is typically the case in cross-country yearly data). Since we want to fit a multi-level model with both firm and country level variables fitting a fixed effects model is not possible. The Enterprise Surveys measure obstacles faced by firms based on their perceptions. When using these subjective measures to account for the investment climate in a country, we are bound to encounter endogeneity issues. Endogeneity arises due to the reason that firms may assess the same investment climate differently, based on their own performance or efficiency. For example, more efficient firms are likely to find a way around obstacles of financial access while those that are not as

efficient may face more problems in obtaining finance. Dollar et al. (2005) use sector-region averages as instrument. They reason that these sector-region averages would allow us to capture the investment climate that is relevant to all firms in a particular sector, located in that particular region. Lall and Mengistae (2005) argue that location decisions of small and medium firms are more or less exogenous and thus this would be a valid instrument. Since our dataset consists mainly of small and medium size firms, we also use the same method of instrumenting. We ensure that there are at least four firms in each sector-region. This way of instrumenting has also been used by Kinda (2010) and Aterido et al. (2007) among others.

Using this method of instrumenting for INFRA 1, INFRA 2, PHYSICAL 1, PHYSICAL 2, and FINANCIAL, we estimate the random effects model. Estimations (6) onwards, in Table 3 give the results of the iv-logit estimations. While the results remain similar in terms of direction and significance the coefficients on INFRA1 and INFRA 2 in estimation (6) are larger. In the case of INFRA 2 it makes a particularly large difference. Similarly, in estimation (7), the coefficients on PHYSICAL 2 when instrumented for, are larger, than in the case of PHYSICAL 1 and FINANCIAL.

We add other firm level variables, namely, INS and SKILL to the firm level explanatory variables (instrumented by their sector-region averages). Estimation (8) and (9) shows these results. We again find that the coefficients on INFRA 2 and PHYSICAL 2 are higher when we take into account the two new variables. INS has an unexpected positive sign but is not significant in the estimations (8) and (9). On the other hand SKILL is highly significant at a 1% level, but has a negative coefficient showing that FDI inflows find lack of education an obstacle. The unexplained variance given by  $\rho$  is between 0.2 and 0.19 as in earlier estimations. The LR tests are significant and thus show that a multi-level understanding of the dependent variable is necessary.

**Table 3: Estimations of firm level variables without INS and SKILL**

	(6)	(7)	(8)	(9)	(10)	(11)
	FDI	FDI	FDI	FDI	FDI	FDI
VARIABLES	iv-logit	iv-logit	iv-logit	iv-logit	iv-logit	iv-logit
LNAGE	-0.25*** (0.02)	-0.25*** (0.02)	-0.25*** (0.02)	-0.25*** (0.02)	-0.25*** (0.02)	-0.25*** (0.02)
MEDIUM	0.9*** (0.04)	0.9*** (0.04)	0.9*** (0.04)	0.9*** (0.04)	0.9*** (0.04)	0.9*** (0.04)
LARGE	2.03*** (0.05)	2.03*** (0.05)	2.04*** (0.05)	2.04*** (0.05)	2.03*** (0.05)	2.04*** (0.05)
AGLOM	0.025*** (0.001)	0.025*** (0.001)	0.025*** (0.001)	0.025*** (0.001)	0.025*** (0.001)	0.025*** (0.001)
INFRA 1	-0.38*** (0.05)		-0.39*** (0.05)		-0.38*** (0.05)	-0.39*** (0.05)
INFRA 2	0.37*** (0.05)		0.45*** (0.05)		0.38*** (0.048)	0.45*** (0.05)
PHYSICAL 1		-0.48*** (0.05)		-0.54*** (0.05)		
PHYSICAL 2		0.18*** (0.06)		0.22*** (0.06)		
FINANCIAL		-0.26*** (0.05)		-0.2*** (0.05)		
INS			0.02 (0.06)	0.027 (0.06)	-0.04 (0.05)	
SKILL			-0.21*** (0.05)	-0.2*** (0.05)		-0.21*** (0.05)
Constant	-2.56*** (0.13)	-2.13*** (0.16)	-2.19*** (0.16)	-1.835*** (0.188)	-2.47*** (0.12)	-2.136*** (0.146)
$\rho$	0.21	0.21	0.21	0.20	0.21	0.20
LR Chi2	1656.07	1669.54	1552.97	1565.04	1650.35	1555.42
LR df(1)	***	***	***	***	***	***
# of firms	43,403	43,403	43,403	43,403	43,403	43,403
# of countries	103	103	103	103	103	103

Standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

All estimations include Sector dummies.

Table 4 shows the correlations between the explanatory variables. To see if these correlations affect the results we drop INS (in estimation 10) and SKILL (in estimation 11)<sup>54</sup>. We find that the results are robust to these changes. We further check the variance inflation factor and find that it is at an acceptable level<sup>55</sup>.

**Table 4: Correlations between explanatory variables**  
**a) for estimations (6) and (8)**

	INFRA 1	INFRA 2	INS	SKILL
INFRA 1	1.00			
INFRA 2	-0.23	1.00		
INS	0.03	0.49	1.00	
SKILL	-0.26	0.64	0.54	1.00

**b) for estimations (7) and (9)**

	PHYSICAL 1	PHYSICAL 2	INS	SKILL	FINANCIAL
PHYSICAL 1	1.00				
PHYSICAL 2	0.15	1.00			
INS	-0.25	0.38	1.00		
SKILL	-0.55	0.23	0.54	1.00	
FINANCIAL	0.25	0.58	0.27	0.20	1.00

So far we find that all the firm level variables that we include have a significant impact on the dependent variable. Further, we find that the (remaining) variance at the firm level is much larger than the variance at the country-level in explaining FDI inflows. However, we find that the estimated values of  $\rho$  indicates a country level dependence of about 0.2 which is consistently significant (as given by the LR tests) and thus indicate the presence of country level effects that cannot be ignored when analyzing FDI inflows.

In order to take into account these country level effects in a more explicit way, and further analyze which specific aspects of the country context are important in understanding these FDI inflows in Table 5, we introduce country level explanatory variables. As explained in the previous section, these variables include LGDPPC, WGI, BIZDAYS, INFLATION, LABOURPARTICIPATION, LEGALRIGHTS, TAXRATE, KAOPEN, and OPENK. We first introduce the macroeconomic variables in estimations (12) and (15).

In the estimations that include country variables, we first note that the coefficients on the firm level variables remain more or less the same as in previous estimations. Among the macro level variables, the GDP per capita of the host country (LGDPPC) has a negative and strongly significant impact. GDP per capita reflects the overall development of a

<sup>54</sup> We show the results for the estimations with INFRA 1 and INFRA 2; the results are similar for estimations with PHYSICAL 1 and PHYSICAL 2 and FINANCIAL as well and hence have not been included here.

<sup>55</sup> The general rule of thumb says that a VIF>10 is harmful.

country; we find that FDI is attracted to countries with a lower level of development as it might imply higher economic profitability. This is in line with our finding that lack of proper infrastructural facilities of transportation and electricity lead to higher FDI ownership. This is perhaps substituting for trade. LABOURPARTICIPATION is significant at a 5% level and has the expected sign. TAXRATE has a positive but a very small coefficient. It does not have the expected sign across the estimations and is not stable across all specifications. INFLATION has a negative impact, suggesting that macroeconomic stability discourages FDI. This variable is significant across all the specifications as well. Both these variables are significant at the 5% level. In (11) and (14) we exclude LGDPPC and include the institutional variables. WGI is negative and significant at the 5% level, implying that low institutional quality at the country level is detrimental to FDI flows as we had expected. This is interesting since the institutional variables at the firm level were never significant. LEGAL RIGHTS and BIZDAYS are also negative but insignificant. When we include both LGDPPC and the institutional variables, (as in estimations 14 and 17) we find that WGI loses significant. This is expected since LGDPPC and WGI are correlated<sup>56</sup>. Further the interaction terms of the size of the firm and WGI shows that while large firms are not affected by the country's institutional set up, it is the medium sized firms that are impacted negatively as compared to small firms. This could be because larger firms can "grease" their way through and small firms aim to establish themselves in foreign countries for much smaller profit margins. It is the medium sized firms that do not profit from paying their way through, and yet cannot discount possible institutional hurdles that might arise after entry. OPENK is positive and significant at a 5% level implying that the more trade openness a country has, the greater the FDI, implying vertical FDI. However, the coefficients are rather small in this case.

Looking at the inter class correlations, we find that the values of  $\rho$  in estimations (10) and (15) drop down to 0.22 to around 0.15 - 0.16 when we include country level determinants. This means that although the country characteristics that were introduced through the new variables are significant, there is still a sizable country variation that is unobserved. Comparing to  $\rho = 0.124$  of estimation (1), we also note that the firm-level variables that introduced captured a larger part of firm-variation than the country-level variables did for country variation. While the country level variables don't exhaustively explain the country variance, LGDPPC and LABOURPARTICIPATION are the strongest results. We find that the LR tests are significant.

---

<sup>56</sup> The correlation between WGI and LGDPPC is nearly 0.68.

**Table 5: Estimations of firm and country level variables without INS and SKILL**

VARIABLES	(12) FDI	(13) FDI	(14) FDI	(15) FDI	(16) FDI	(17) FDI
LNAGE	-0.25*** (0.022)	-0.25*** (0.022)	-0.25*** (0.022)	-0.25*** (0.022)	-0.25*** (0.022)	-0.25*** (0.022)
MEDIUM	0.89*** (0.042)	0.88*** (0.042)	0.89*** (0.042)	0.89*** (0.042)	0.89*** (0.042)	0.89*** (0.042)
LARGE	2.026*** (0.046)	2.018*** (0.046)	2.024*** (0.046)	2.025*** (0.046)	2.017*** (0.046)	2.023*** (0.046)
AGLOM	0.024*** (0.001)	0.025*** (0.001)	0.025*** (0.001)	0.024*** (0.001)	0.025*** (0.001)	0.025*** (0.001)
INFRA 1	-0.43*** (0.049)	-0.41*** (0.05)	-0.43*** (0.05)			
INFRA 2	0.37*** (0.045)	0.36*** (0.045)	0.37*** (0.045)			
PHYSICAL 1				-0.52*** (0.048)	-0.5*** (0.048)	-0.52*** (0.049)
PHYSICAL 2				0.16*** (0.055)	0.16*** (0.055)	0.16*** (0.055)
FINANCIAL				-0.24*** (0.05)	-0.24*** (0.05)	-0.24*** (0.05)
LGDPPC	-0.24*** (0.065)		-0.25*** (0.078)	-0.22*** (0.065)		-0.24*** (0.077)
WGI		-0.07* (0.041)	0.008 (0.047)		-0.062 (0.04)	0.013 (0.046)
WGI*MEDIUM	-0.06*** (0.018)	-0.06*** (0.018)	-0.06*** (0.018)	-0.06*** (0.018)	-0.06*** (0.018)	-0.06*** (0.018)
WGI*LARGE	-0.025 (0.019)	-0.025 (0.019)	-0.025 (0.019)	-0.026 (0.019)	-0.026 (0.019)	-0.026 (0.019)
BIZDAYS		-0.001 (0.001)	-0.001 (0.001)		-0.001 (0.001)	-0.001 (0.001)
INFLATION	-0.017** (0.007)	-0.017** (0.007)	-0.017** (0.007)	-0.016** (0.007)	-0.016** (0.007)	-0.017** (0.007)
LABOUR PARTICIPATION	0.022*** (0.008)	0.028*** (0.008)	0.02** (0.008)	0.023*** (0.008)	0.028*** (0.008)	0.021** (0.008)
LEGALRIGHTS		-0.029 (0.029)	-0.03 (0.028)		-0.03 (0.029)	-0.032 (0.028)
TAX RATE	0.003* (0.001)	0.003* (0.001)	0.002 (0.001)	0.003* (0.001)	0.003* (0.001)	0.0024 (0.001)
KAOPEN	-0.054 (0.048)	-0.045 (0.05)	-0.058 (0.049)	-0.058 (0.048)	-0.051 (0.05)	-0.064 (0.049)
OPENK	0.005**	0.004**	0.005**	0.005**	0.004**	0.005**

VARIABLES	(12)	(13)	(14)	(15)	(16)	(17)
	FDI	FDI	FDI	FDI	FDI	FDI
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	-2.49***	-4.51***	-2.08**	-2.25**	-4.11***	-1.79*
	(0.883)	(0.660)	(1.002)	(0.884)	(0.662)	(1.004)
$\rho$	0.15	0.16	0.15	0.15	0.15	0.15
LR Chi2	1318.53	1318.00	1306.16	1332.83	1331.82	1320.51
LR df(1)	***	***	***	***	***	***
# of firms	43,403	43,403	43,403	43,403	43,403	43,403
# of countries	103	103	103	103	103	103

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

All estimations include Sector dummies.

In Table 6 we include INS and SKILL at the firm level along with the country level variables. Like in the case of estimations (6) to (9) we find that institutional obstacles at the firm level have no significant impact on FDI decisions. On the other hand, SKILL is highly significant and negative. The remaining firm level variables do not change much from the previous estimations. We reiterate that firms entering developing country markets are attracted by profits but are dissuaded if they do not have the required skill in production of goods. Firms may enter developing country markets which have the required skill and also have relatively cheap labour as compared to their country of origin, which can be seen from the significant LABOUR PARTICIPATION variable at the country level.

For the remaining country level variables, the results are very similar to those discussed in Table 4 above. For these estimations, the inter class correlations are 0.14 - 0.15. Hence we again find that the overall development of a country and its macroeconomic situation are more important than the institutional quality at the country level. Recognizing that WGI and LGDPPC are correlated, we still find that the LGDPPC has a greater significance than WGI and thus we conclude that LGDPPC has a bigger influence on FDI than WGI. Here again we find that medium sized firms are strongly affected by the country institutional set up as compared to small firms. Large firms on the other hand are not significantly affected<sup>57</sup>.

<sup>57</sup> Including WGI as a geographically weighted measure (taking into account spatial influence of neighbouring countries in a country's institutions -as discussed in Chapters 3 & 4) we find that the results are not influenced in anyway . The reason for this could be that the scaling effect at the country level is lost on firm-level decisions. That is, the influence of the institutional quality of neighbouring countries on a country's own institutions has no bearing a firm's decision to enter a country. Further, the model's parsimony might be lost in including such scaling effects when we attempt to understand firm-level decisions given a country's macroeconomic and institutional structure,

**Table 6: Estimations of firm and country level variables with INS and SKILL**

	(18)	(19)	(20)	(21)	(22)	(23)
	FDI	FDI	FDI	FDI	FDI	FDI
VARIABLES	iv-logit	iv-logit	iv-logit	iv-logit	iv-logit	iv-logit
LNAGE	-0.25*** (0.022)	-0.25*** (0.022)	-0.25*** (0.022)	-0.25*** (0.022)	-0.25*** (0.022)	-0.25*** (0.022)
MEDIUM	0.89*** (0.042)	0.89*** (0.042)	0.89*** (0.042)	0.89*** (0.042)	0.89*** (0.042)	0.89*** (0.042)
LARGE	2.03*** (0.046)	2.02*** (0.046)	2.03*** (0.046)	2.03*** (0.046)	2.02*** (0.046)	2.03*** (0.046)
AGLOM	0.03*** (0.001)	0.025*** (0.001)	0.025*** (0.001)	0.025*** (0.001)	0.025*** (0.001)	0.025*** (0.001)
INFRA 1	-0.44*** (0.05)	-0.42*** (0.05)	-0.44*** (0.05)			
INFRA 2	0.44*** (0.05)	0.444*** (0.05)	0.443*** (0.05)			
PHYSICAL 1				-0.57*** (0.05)	-0.56*** (0.05)	-0.57*** (0.05)
PHYSICAL 2				0.19*** (0.057)	0.2*** (0.057)	0.186*** (0.06)
FINANCIAL				-0.21*** (0.05)	-0.21*** (0.05)	-0.21*** (0.05)
INS	0.027 (0.055)	0.018 (0.055)	0.027 (0.055)	0.031 (0.055)	0.024 (0.055)	0.033 (0.055)
SKILL	-0.2*** (0.053)	-0.22*** (0.053)	-0.2*** (0.053)	-0.19*** (0.053)	-0.2*** (0.053)	-0.19*** (0.053)
LGDPPC	-0.21*** (0.065)		-0.21*** (0.078)	-0.12*** (0.065)		-0.21*** (0.078)
WGI		-0.066* (0.04)	0.009 (0.047)		-0.058 (0.04)	0.007 (0.047)
WGI*MEDIUM	-0.06*** (0.018)	-0.06*** (0.02)	-0.06*** (0.018)	-0.06*** (0.018)	-0.06*** (0.018)	-0.06*** (0.018)
WGI*LARGE	-0.025 (0.019)	-0.024 (0.019)	-0.024 (0.019)	-0.025 (0.019)	-0.025 (0.019)	-0.025 (0.019)
BIZDAYS		-0.001 (0.001)	-0.001 (0.001)		-0.001 (0.001)	-0.001 (0.001)
INFLATION	-0.014** -0.007	-0.015** -0.007	-0.016** -0.007	-0.014** -0.007	-0.015** -0.007	-0.015** -0.007
LABOUR PARTICIPATION	0.023*** (0.008)	0.028*** (0.008)	0.021** (0.008)	0.023*** (0.008)	0.028*** (0.008)	0.021** (0.008)
LEGAL RIGHTS		-0.028	-0.03		-0.03	-0.031

	(18)	(19)	(20)	(21)	(22)	(23)
	FDI	FDI	FDI	FDI	FDI	FDI
VARIABLES	iv-logit	iv-logit	iv-logit	iv-logit	iv-logit	iv-logit
		(0.028)	(0.028)		(0.028)	(0.028)
TAX RATE	0.003*	0.003**	0.003	0.003*	0.003**	0.003
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
KAOPEN	-0.063	-0.055	-0.066	-0.066	-0.06	-0.07
	(0.048)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
OPENK	0.005**	0.005**	0.005**	0.005**	0.005**	0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	-2.48***	-4.19***	-2.134**	-2.265**	-3.86***	-1.86*
	(0.88)	(0.65)	(0.99)	(0.88)	(0.66)	(1.00)
$\rho$	0.15	0.15	0.14	0.14	0.15	0.14
LR Chi2	1237.5	1222.48	1226.21	1248.48	1234.81	1237.56
LR df(1)	***	***	***	***	***	***
# of firms	43,403	43,403	43,403	43,403	43,403	43,403
# of countries	103	103	103	103	103	103

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

All estimations include Sector dummies.

## 4. Conclusions

Attracting FDI inflows has been argued to help developing countries induce economic growth in them. While many policies are made by these countries to attract MNE's to invest in their markets, firm decisions on entering new markets are based both on their own idiosyncratic characteristics as well as host country characteristics. Using firm and country data in a multilevel framework, we analyze how obstacles faced by firms both at the operational level and the macroeconomic and institutional situation in a country affect FDI inflows into developing countries. We find from our analysis that a firm's age and size are strongly statistically significant in all the regressions. We consider constraints that are divided into physical infrastructure and financial access constraints, skill and institutional constraints. When considering a micro level analysis we find that internet availability problems are detrimental to FDI decisions. The lack of physical infrastructure facilities such as electricity and transportation on the other hand are found to be encouraging inward FDI. Lack of skilled labour is found to be detrimental to FDI while institutional obstacles such as corruption and crime are not seen as an important determinant.

As emphasized by Dunning (2001) and Rugman (1980, 1981) we can say that FDI inflows, while largely explained by firm specific idiosyncrasies, do go hand in hand with host country conditions. Host country conduciveness does play a small but significant role in explaining the attractiveness of FDI inflows. We find that the macroeconomic variables such as LGDPPC, LABOUR PARTICIPATION, and to a smaller extent IN-

FLATION, OPENK and TAXRATES affect FDI decisions on FDI inflows. Institutions measured as WGI at the country level are never significant except when considered without accounting for overall development. Since the overall development of a country has a bearing on the institutional quality of the country this is expected. This is consistent across all specifications.

While many macro studies such as Al-Sadig (2009), Seyoum (2009) and Anghel (2005) among others have found institutions to be an important element in attracting FDI, we find that this is not the case here, when we take both firm level and country level variables into account in a multilevel regression. One reason being that overall development of a country already captures the country institutional set up. Further, country level perceptions measures of institutions are often biased by the macroeconomic situation of the country. At the firm level however, firms might not be as biased towards their perceptions of crime. Their perception on crime and corruption are based on day to day activities of businesses. This creates a disparity between the levels of importance given to institutions at both levels. Secondly, profitability of MNE's is dependent on how they can overcome these problems at the least possible cost in the host country. It is likely that certain business costs, even if large, are worth incurring, if long term profits are anticipated. This is further confirmed by the fact that while foreign firms are more likely to be larger firms, these costs do not deter them from entering a country with a bad institutional set up, if the profitability is high.

We could conclude from this that FDI inflows into a country are largely influenced by firm idiosyncratic factors and partially on the general macroeconomic conditions and hardly so on the institutional environment of host countries. As economic fundamentals tend to increase returns to investment, a certain level of uncertainty of crime, corruption and expropriation of wealth seem to be unimportant. FDI decisions are thus neither made solely based on firm characteristics, nor are they based entirely on host country conduciveness for business. We find that the constraints reported by firms do reflect country characteristics and their levels of income and overall development, which are small yet significant. Hence we conclude that it is necessary to look at firm level activities such as FDI in a multi-level manner rather than at single levels.

## Appendix

**Table (A): Variable descriptions**

Spatial Level/ Scale	Variable	Variable description
Firm-level dependent variable	FDI	Dummy variable: 0= domestic firms 1= at least 10% of their capital owned by foreign persons or organizations
Firm-level independent variables	Lnage	Age of the firm
	Size	1 = < 20 employees (base) 2= medium(20-99) 3= large(>100)
	Infra	Overall infrastructural obstacles firms face Obtained from the scores of Component 1 of PCA (electricity, transportation, email, web, financial)
	Physical	Physical infrastructure obstacles firms face Obtained from the scores of Component 1 of PCA (electricity, transportation, email, web)
	Financial	Obstacles to financial access
	Ins	Institutional quality. Obtained from the scores of Component 1 of PCA (corruption, crime and proprights)
	Skill	Obstacle due to lack of skilled labour.
	Aglom	number of foreign firms present in a sector-region
Country-level independent variables	Lgdppc	GDP per capita (log of current USD prices)
	WGI	Measure of Governance Obtained from the scores of Component 1 of PCA of the 6 World Governance Indicators
	Bizdays	number of calendar days required to start a business
	Legal rights	degree of protection given to lenders and borrowers, given the country's bankruptcy laws
	Tax rates	tax rate (% commercial profits)
	Kaopen	Financial openness measured by "extent and intensity" of capital controls of countries
	Openk	openness to trade measured by net exports as a percentage of GDP
	labourparticipation	proportion of working age population which is economically active
	inflation	growth rate of the GDP deflator

**Table (B): World Governance Indicators**

Voice and accountability	<i>Perceptions of freedom of citizen's in choosing their government as well as freedom of press</i>
Political Stability and Absence of Violence	<i>Perceptions on the stability of government and on the presence of political violence and terrorism</i>
Government Effectiveness	<i>Perceptions on the quality of public and civil services as well we effective government policies and their credibility</i>
Regulatory Quality	<i>Perceptions on government regulation, and its ability to foster private sector development</i>
Rule of Law	<i>Perceptions on the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence</i>
Control of Corruption	<i>Perceptions on corruption and the power elites and private interests.</i>

Source: [www.govindicators.org](http://www.govindicators.org)

**Table (C): List of countries in the larger sample:**

Afghanistan	Costa Rica	Latvia	Rwanda
Albania	Cote d'Ivoire	Lesotho	Samoa
Angola	Croatia	Liberia	Senegal
Argentina	Czech Republic	Lithuania	Sierra Leone
Armenia	Dominican Republic	Macedonia, FYR	Slovak Republic
Azerbaijan	Ecuador	Madagascar	Slovenia
Bahamas, The	El Salvador	Malawi	South Africa
Bangladesh	Eritrea	Mali	St. Lucia
Belarus	Estonia	Mauritania	St. Vincent & the Grenadines
Belize	Fiji	Mauritius	Suriname
Benin	Gabon	Mexico	Swaziland
Bhutan	Gambia, The	Moldova	Tajikistan
Bolivia	Georgia	Mongolia	Tanzania
Bosnia and Herzegovina	Ghana	Mozambique	Togo
Botswana	Guatemala	Namibia	Tonga
Brazil	Guinea	Nepal	Trinidad and Tobago
Bulgaria	Guinea-Bissau	Nicaragua	Turkey
Burkina Faso	Guyana,Co-op Republic of	Niger	Uganda
Burundi	Honduras	Pakistan	Ukraine
Cameroon	Hungary	Panama	Uruguay
Cape Verde	Indonesia	Paraguay	Uzbekistan
Chad	Jamaica	Peru	Venezuela, RB
Chile	Kazakhstan	Philippines	Vietnam
Colombia	Kenya	Poland	Yemen, Rep.
Congo, Dem. Rep.	Kyrgyz Republic	Romania	Zambia
Congo, Rep.	Lao PDR	Russian Federation	

**CHAPTER 6**  
**Conclusion**

This thesis sets out to empirically analyze the spatial interactions of institutions across countries. To do so, we focus on two fundamental causes of economic growth and development, namely institutions and geography. The findings contribute to our understanding of the interrelationship between institutions across countries and economic growth and development. Our interest is to link institutions and geography to economic growth and development using a spatial narrative. It further contributes to our understanding of institutional quality at the firm and country level and their influence on FDI inflows. We mainly use cross-section analysis. Furthermore, this thesis also contributes to the methodology in analyzing spatial correlation using principal components and canonical correlation analysis. Below we summarize each chapter and highlight our findings.

## **1. Summary and main findings**

In the introductory chapter, we briefly survey the literature on institutions and geography as fundamental causes of economic growth and development in an economy. We highlight the need to study not only institutions and 'absolute' geography (the physical endowments of a country, which are independent of its neighbours) but also as the 'relative' geography (or spatial interaction) of countries in order to better understand economic growth and development of countries. We also note that institutions within nation-state boundaries need to be analyzed at both country level as well as at the micro/ firm-level.

To understand the relationship between economic growth and development and institutional quality in countries by taking into account the underlying spatial relations (relative geography), we begin our analysis by proposing a method for spatial principal component analysis. This method is explained in Chapter 2 and produces summary measures (or components) that maximize spatial correlation underlying in a dataset. We propose an alternative spectral decomposition method to Wartenberg (1985). Our method spectrally decomposes a distance weighted Moran matrix to obtain the Eigen vectors and Eigen values, thus maximizing the underlying spatial correlation within the data. The other feature of our method is that we find an easy and intuitive link to spatial canonical correlation analysis. While the spatial principal component analysis is performed on one data set, spatial canonical correlation analysis finds the relationship between two sets of multiple variables  $X$  and  $Y$ , for which we obtain two sets of latent vectors, each measuring different phenomena. We provide example applications of the methods and show that our variant of spatial canonical correlation analysis may produce rather different results than spatial principal components analysis using Wartenberg's method. We also illustrate how spatial canonical correlation analysis may produce different results than spatial principal components analysis.

In chapter 3 we examine the role of institutions relative to economic performance, absolute geography and financial performance of a country. In order to do this, we use the spatial principal component analysis and a spatial canonical correlation analysis (dis-

cussed in chapter 2) to obtain multi-dimensional measure of institutions, economic performance, absolute geography and financial performance of countries. Our analysis shows that the first canonical functions in all the cases give us results that conform to current literature. That is, we find that a higher level of development is correlated to a higher level of institutional quality, deeper financial structure as well as "good" geography of the Jeffery Sachs variety. From the second canonical functions we find that economic growth is correlated to market steering. We further find that geographic conditions need not define the institutional set up of countries. A similar institutional set up need not result in a similar financial structure in countries. We show that there is a necessity to take spatial interactions with neighbouring countries into account while analyzing the relationships between institutions, geography, economic and financial performance of a country. We find that space indeed has a strong influence on the prevailing institutional and economic conditions of countries. While the impact of space on geography is very obvious, we find that it has no bearing on the financial performance of countries.

In chapter 4, we test the Rodrik et al. (2004) framework to explain differences in development levels across countries by using a broader set of definitions for institutions, geography and economic variables. We use a multi-faceted database to measure institutions in an attempt to go beyond the single-dimension measures that are often employed. Using measures that we obtain in Chapter 3, we test whether it is institutions, absolute geography or trade that "trumps" other factors in explaining development differences. The regression analysis we use is a two stage least squares – instrumental variable approach. We find that institutions trump other factors (geography and trade) when we use GDP per capita as an independent variable. When we expand the dependent variable to include other aspects of development, such as growth and investment, we find that institutions, growth and geography are all important variables. In this case, institutions no longer trump the other factors. In this case, we also find that the same institutions variable that was positively associated to GDP per capita is now negatively correlated with the more dynamic development variable.

In the last chapter (Chapter 5) we shift from a macro level analysis to a multi-level analysis. In this study, we empirically analyze the determinants of FDI inflows into developing countries. We do this by using firm-level data obtained from the Enterprise Surveys data of the World Bank and country level data from various sources. Using a multi-level logit model, we analyze how institutional and structural variables at both firm and country levels impact inward FDI. In our view, there is a gap between analysis at the country level studies and firm level studies on inward FDI. In this paper, we fill the gap by doing a multi-level regression analysis, taking into account both firm variables and country characteristics to explain inward FDI. We find that firm structural characteristics and obstacles they face most affect inward FDI. While some macroeconomic variables such as GDP per capita, inflation and openness have a significant influence, other variables

that measure institutional quality of a country do not have any statistically significant influence on FDI inflow.

From the above analysis we find that neither institutions, nor geography or trade theories can single handedly explain the economic performance of a country. While all the three hypotheses tend to use 'absolute' geography as the backcloth in their explanations of what is the most fundamental cause of economic growth. In this study we have shown that the spatial influence of neighbours has an impact on the economic performance of an economy. We find that similar geography does not necessarily result in the same level of economic development. We do find clusters of institutional quality but these do not always reflect the level of development in the countries. We find that taking conventional measures such as GDP per capita give us results that are in line with conventional theories of growth, but these do not always hold water when we take a more multi-dimensional measure of economic performance. Similarly, when we take single variables such as democracy to measure institutional quality we find our results in line with mainstream theories, which is not so when we take multi-dimensional measures. Studying institutional quality within a country, we find that firms are mainly attracted by economic profitability in a foreign country. Their perceived sense of institutional quality does not seem to be as big a hindrance at the micro-level. Firms seem to be influenced by the level of GDP per capita which reflects institutional quality measured as an aggregate at the country level influences the level of FDI in a country. This means that more importance needs to be given to firm level workings of institutions and their interactions with country level perceptions in order for countries to attract more FDI which would in turn help the country's economic performance.

## **2. Further Research and limitations of the study**

In Chapter 3, we construct latent variables that take into account the multi-dimensional nature of institutions, absolute geography, economic development and financial development using spatial PCA and spatial CCA. These methods prove to be very useful when tackling with a large number of variables that measure similar phenomena, and are possibly highly correlated. These latent variables however are very sensitive to the data used. This is an often mentioned criticism for both principal component analysis as well as canonical correlation analysis. Thus, the validity of these conclusions are often based entirely on the choice of the data.

A related limitation is that of causality. Does a higher level of development lead to better institutions or do better institutions lead to a higher level of development. Current research on institutions and economic development assumes that causality runs from the former to the latter, and reverse causality is often neglected. The methodology used in Chapter 2, Canonical correlation analysis does not allow for such directionality, either from institutions to development or the other way around. While we do not set out to address this question, this study neglects causality in Chapter 2. One way of incorporat-

ing reverse causality is to use an instrumental variable approach, which we do in Chapter 3.

The data on institutional quality used in this thesis are perceptions indices. Institutions are by definition harder to measure than other macroeconomic policy variables. Macro level measures of institutions are often indices constructed based on perceptions. These measures are oftentimes biased towards European/American institutions. As Chang (2011) puts it, these measures try and fit a certain description (such as the liberalization narrative) thus showing a rather biased picture on their influence on economic growth and development. We thus expect that institutional quality of countries might be overestimated for those countries that have high growth rates, if only because they are expected to have better institutions. The inherent biases in perceptions indices might undermine the conclusions of the study, thus making it difficult to make concrete policy recommendations.

In Chapter 3, we use an instrumental variable approach. We instrument for institutions using settler mortality and percentage of English and European language speaking population, which are popular in this literature. These are rather weak instruments as has been pointed out by Albouy (2012) the settler mortality data suffers from various problems. Similarly, the Hall and Jones (1999) instruments also suffer from weak correlation to institutional quality. For this line of research to go forward, we need to look for a wider choice of instruments, which prove to be better than those that are currently in vogue.

The findings in this thesis have a number of implications for future research. Firstly, we confirm that relative geography needs to be given more attention in order for us to get a better understanding on the relationship between various “fundamental sources of growth” and growth and development itself. By including relative geography in the picture, this study moves away from fatalistic explanations of absolute geography, as well as bring into picture the need for specificity of circumstances, rather than sweeping generalizations that may or may not hold across time and space, in understand the links between institutional quality and economic performance.

From our analysis in Chapters 3 and 4, we know that institutions that are important for growth and development when bundled together are different from those that seem to be important for growth at a given time. Further research in this field will need to explain which specific institutional characteristics within classes of institutions matter for growth and which matter for development.

Finally, it should be noted that multi-level modelling has not been used much in economics, even when analyzing those economic phenomena that we know have explicit micro-meso-macro dimensions. Further research on institutional quality and its influence on economic growth and development would benefit from this type of analysis.

The national innovation systems literature emphasizes on how institutions are important for encouraging and fostering innovation in a country. This thesis has only touched upon institutions in relation to economic and financial performance and geography. To take this research forward in this direction, it would be interesting to see how institutions and innovation systems are correlated spatially amongst countries, and what institutional aspects matter most in forstering growth..

# Bibliography

- Abreu, M., De Groot, H. L. F., and Florax, R. J.G.M., (2004). "Space and Growth: A Survey of Empirical Evidence and Methods". Tinbergen Institute Working Paper No. TI 04-129/3.
- Acemoglu, D., S. Johnson, and J. Robinson, (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 5, 1369-1401.
- Acemoglu, D., S. Johnson, and J. 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, and J. Robinson, (2004), "Institutions as the Fundamental Cause of Long-run Growth," National Bureau of Economic Research Working Paper no.10481.
- Acemoglu, D., S. Johnson, and J. Robinson, (2005), "The Rise of Europe: Atlantic Trade, Institutional Change and Growth", *American Economic Review*, June 2005, volume 95, pp. 546- 579.
- Acemoglu, D., (2008), "Interactions between Governance and Growth" in North, Douglass., Acemoglu, Daron., Fukuyama, Francis., Rodrik, Dani., (April 2008), "Governance, Growth, and Development Decision-making". The World Bank.
- Acemoglu, D., (2009), *Introduction to Modern Economic Growth*. Princeton University Press.
- Adams, S. (2009). "Foreign Direct investment, domestic investment, and economic growth in Sub-Saharan Africa". *Journal of Policy Modeling*, 31(6), 939-949.
- Aghion, P., & Howitt, P. (1992). "A Model of Growth through Creative Destruction". *Econometrica*, 60(2), 323-51.
- Alesina, A., S. Ozler, N. Roubini, and P. Swagel, (1996), "Political Instability and Economic Growth," *Journal of Economic Growth*, 1, 2, 189-211.
- Alfaro, L., Chanda, A., Kalemli-Ozcan, S., & Sayek, S. (2010). "Does foreign direct investment promote growth? Exploring the role of financial markets on linkages". *Journal of Development Economics*, 91(2), 242-256.
- Alguacil, M., Cuadros, A., & Orts, V. (2011). "Inward FDI and growth: The role of macroeconomic and institutional environment". *Journal of Policy Modeling*, 33(3), 481-496.
- Al-Sadig, A. (2009). "Effects of Corruption on FDI Inflows", *The Cato J.*, 29, 267.
- Anghel, B. (2005). "Do institutions affect foreign direct investment?" *International Doctorate in Economic Analysis*, 2-40.
- Aristotelous, K., & Fountas, S. (1996). "An Empirical Analysis of Inward Foreign Direct Investment Flows in the EU with Emphasis on the Market Enlargement Hypothesis". *JCMS: Journal of Common Market Studies*, 34(4), 571-583.
- Aterido, R., Hallward-Driemeier, M., & Pages, C. (2007). "Investment climate and employment growth: The impact of access to finance, corruption and regulations across firms". *Inter-American Development Bank Working Paper No. 626*.
- Baer, W., & Sirohi, R. A. (2013). "The Role of Foreign Direct Investments in the Development of Brazil and India: A Comparative Analysis". *Kyklos*, 66(1), 46-62.
- Bardhan, P., (2005), "Institutions matter, but which ones?" *Economics of Transition* 13.
- Barro, R. (1996). "Democracy and Growth," *Journal of Economic Growth* 1, 1-27.
- Barro, R.J. 1999. "Determinants of democracy." *Journal of Political Economy*, 107 (6): 158- 183.
- Basu, S. R., (2008), "A New Way to Link Development to Institutions, Policies and Geography," *Policy Issues in International Trade and Commodities Study Series No 37*.
- Beck, T., Demirgüç-Kunt, A., and Levine, R., (2000), "A New Database on Financial Development and Structure," *World Bank Economic Review* 14, 597-605.
- Beck, N., Gleditsch, K. S. And Beardsley, K. (2006), "Space Is More than Geography: Using Spatial Econometrics in the Study of Political Economy". *International Studies Quarterly*, 50: 27-44.
- Bénassy-Quéré, A., Coupet, M., & Mayer, T. (2007). "Institutional determinants of foreign direct investment". *The World Economy*, 30(5), 764-782.
- Bennedsen, M., Malchow-Moller, N., Vinten, F., (2005-01). " Institutions and Growth- a Literature Survey". *CEBR Report*.

- Bhupatiraju, S., & Verspagen, B., (2013a). "Economic development, growth, institutions and geography", UNU-MERIT Working Paper 2013-056
- Bhupatiraju, S., Verspagen, B., & Ziesemer, T., (2013b) "Summarizing large spatial datasets: Spatial principal components and spatial canonical correlation", UNU-MERIT Working Paper 2013-011
- Blaut, J. M. (1999), "Environmentalism and Eurocentrism". *Geographical Review*, 89: 391–408.
- Blonigen, B. (2005). "A Review of the Empirical Literature on FDI Determinants," *Atlantic Economic Journal*, International Atlantic Economic Society, vol. 33(4), pages 383-403, December.
- Bloom, D. and Sachs, J., (1998), "Geography, Demography, and Economic Growth in Africa," mimeo, Harvard Institute for International Development.
- Bloom, D., Canning, D., and Sevilla, J., (2003), "Geography and Poverty Traps," *Journal of Economic Growth*, 8, 355-378.
- Brohman, J (1995) "Universalism, Eurocentrism, and Ideological bias in development studies: From modernization to neo-liberalism", *Third World Quarterly*, 16 (1), pp. 121- 140.
- Buckley, P. J., Devinney, T. M., & Louviere, J. J. (2007d). "Do managers behave the way theory suggests? A choice theoretic examination of foreign direct investment location decision making". *Journal of International Business Studies*, 38 (7), 1069–1094.
- Butkiewicz, J. L. & Yanikkaya, H., (2006). "Institutional quality and economic growth: Maintenance of the rule of law or democratic institutions, or both?", *Economic Modelling*, Elsevier, vol. 23(4), pages 648-661, July.
- Busse, M., & Hefeker, C. (2007). "Political risk, institutions and foreign direct investment". *European Journal of Political Economy*, 23(2), 397-415.
- Cameron, C. A., and Trivedi, P.K., (2009), *Microeconometrics: Methods and Applications*. Cambridge University Press.
- Chang, H., & Evans, P., (2000), "The Role of Institutions in Economic Change".
- Chang, H., (2002). "Kicking Away the Ladder: Development Strategy in Historical Perspective", Anthem Press.
- Chang, H., (2007). "Chapter 2: Understanding the Relationship between Institutions and Economic Development- Some Key Theoretical Issues in Institutional Change and Economic Development", ed. Chang, Ha-Joon.
- Chang, H., (2011). *Institutions and economic development: theory, policy and history*. *Journal of Institutional Economics*, 7, pp 473-498
- Cheshire, P., and Magrini, S., (2008) "Urban Growth Drivers in a Europe of Sticky People and Implicit Boundaries", *Journal of Economic Geography*.
- Cheshire, P, and Magrini, S., (2008) "Urban Growth Drivers in a Europe of Sticky People and Implicit Boundaries", *Journal of Economic Geography*.
- Chinn, M. D., & Ito, H. (2008). "A new measure of financial openness. *Journal of Comparative Policy Analysis*", 10(3), 309-322.
- Cho, W. T . (2003) "Contagion Effects and Ethnic Contribution Networks". *American Journal of Political Science* 47:368–387.
- Clark , D. (1975). "Understanding Canonical Correlation Analysis. Concepts and techniques in modern geography", No.3. Norwich: Geo Abstracts Ltd, 1975.
- Cliff, A.D. and J.K. Ord, (1981), *Spatial Processes: Models & Applications*, Taylor & Francis
- Darlauf, Steven N., Johnson, Paul A., Temple, Jonathan R. W., 2004; "Growth Econometrics".
- Daude, C., & Stein, E. (2007). "The quality of institutions and foreign direct investment". *Economics & Politics*, 19(3), 317-344.
- de Crombrughe, D., Farla, K., Meisel, N., de Neubourg, C., Ould Aoudia, J., & Szirmai, A., (2010). " Institutional Profiles Database III". Technical report, Agence Française de Développement (AFD), Maastricht University (UM), and CEPII, 2010. Presentation of the Institutional Profiles Database 2009 (IPD 2009).
- Delbecque, V., Méjean, I., & Patureau, L. (2008). "Social competition and firms' location choices". Centre D'études Prospectives Et d'Informations Internationales (CEPII) Working Paper, 12.
- Diamond, J., (1997), *Guns Germs and Steel*, New York: W.W. Norton & co.
- Dollar, D. and Kraay, A., (2002), "Growth is Good for the Poor," *Journal of Economic Growth*, 7, 3, 195-225.
- Dollar, D. and Kraay, A., (2003), "Institutions, Trade and Growth: Revisiting the Evidence," *Journal of Monetary Economics*, 50, 1, 133-162.

- Dollar, D. and Kraay, A., (2004), "Trade, Growth and Poverty," *Economic Journal*, Royal Economic Society, vol. 114(493), pages F22-F49, 02.
- Dollar, D., Hallward-Driemeier, M., & Mengistae, T. (2005). "Investment Climate and Firm Performance in Developing Economies". *Economic Development and Cultural Change*, 54(1), 1–31
- Dollar, D., Hallward-Driemeier, M., & Mengistae, T. (2006). "Investment climate and international integration". *World Development*, 34 (9), 1498–1516.
- Dunning, J. H., (1988), "Trade, Location and Economic Activity and the Multinational enterprise: A Search for an Eclectic Approach", Ch. 1, Pg 13-40, in Dunning, John H., "*Explaining International Production*", Unwin Hyman, London, ISBN- 0-04-44517-9.
- Dunning, J. H., (2001). "The eclectic (OLI) paradigm of international production: past, present and future". *International journal of the economics of business*, 8(2), 173-190.
- Easterly, W. & Levine, R., (2001), "It's Not Factor Accumulation: Stylized Facts and Growth Models," *World Bank Economic Review*, 15, 177-219.
- Easterly, W., & Levine, R., (2003): 'Tropics, Germs, and Crops: How Endowments Influence Economic Development', *Journal of Monetary Economics*, 50, 1, 3-39.
- Engerman, S. L. and Sokoloff, K. L., (1997). "Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States." In Stephen Haber, ed., *How Latin America Fell Behind*. Stanford.
- Engerman, S. L. & Sokoloff, K. L., (2002). "Factor Endowments, Inequality and Paths of Development Among New World Economies." *Economia* 3 (Fall): 41-109.
- Engerman, S. L. & Sokoloff, K. L., (2005). "Colonialism, Inequality, and Long-Run Paths of Development," NBER Working Paper 11057.
- Fingleton, B. & López-Bazo, E., (2006), Empirical growth models with spatial effects. *Papers in Regional Science*, 85: 177–198.
- Frankel, J. & Romer, D., (1999), "Does Trade Cause Growth?," *American Economic Review*, 89, 3, 379-399.
- Gallup, J.L., Sachs, J.D. and Mellinger, A. D., (1999), Geography and Economic Development, *International Regional Science Review*, vol. 22, pp. 179-232.
- Gallup, J. L. and Sachs, J. D., (2001), "The Economic Burden of Malaria," *The Supplement to The American Journal of Tropical Medicine & Hygiene*. Vol. 64, no. 1, 2, January/February.
- Gelb, A., Ramachandran, V., Shah, M. K., & Turner, G. (2007). What matters to African firms. The relevance of perceptions data, *World Bank Policy Research Working Paper*, (4446).
- Gleditsch, K. S. and Ward, M. D. (2000), War and Peace in Space and Time: The Role of Democratization. *International Studies Quarterly*, 44: 1–29
- Goedhuys, M. & Srholec, M., 2010. "Understanding multilevel interactions in economic development," MERIT Working Papers 003, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).
- Goldstein, H. (2003). *Multilevel statistical models*. London: Edward Arnold.
- Grossman, G. M., & Helpman, E. (1991). Quality ladders in the theory of growth. *The Review of Economic Studies*, 58(1), 43-61
- Hall, R., and Jones, C., (1999), "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 114, 1, 83-116
- Harvey, D (2011). "Roepke lecture in economic geography—Crises, geographic disruptions and the uneven development of political responses." *Economic Geography* 87(1) pp. 1-22.
- Hartman, D. G. (1984). "Tax policy and foreign direct investment in the United States". *National Tax Journal*, 475-487.
- Hartman, D. G. (1985). "Tax policy and foreign direct investment". *Journal of Public economics*, 26(1), 107-121.
- Helpman, E., Melitz, M., and Rubinstein, Y., (2008), "Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics*, MIT Press, vol. 123(2), pages 441-487, 05.
- Heston, A., & Summers, R., and Aten, B. (2012): Penn World Table Version 7.1. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- Hossein M. H. & Kaneko, S., (2012). "Causality between pillars of sustainable development: Global stylized facts or regional phenomena?," *Ecological Indicators*, Volume 14, Issue 1
- Johnson, R.A. and Wichern, D.W., (2007), *Applied Multivariate Statistical Analysis*, 6<sup>th</sup> edition, Pearson Prentice-Hall, Upper Saddle River, New Jersey.

- Jombart, T. (undated), "A tutorial for the spatial Principal Component Analysis using the R package adegenet" 1.2-2, mimeo.
- Kaufmann, D., & Wei, S. J. (1999). "Does 'grease money' speed up the wheels of commerce?" (No. w7093). National bureau of economic research.
- Kaufmann, D., Kraay, A., & Zoido-Lobaton, P., (1999). "Aggregating Governance Indicators". World Bank Policy Research Working Paper No. 2195.
- Kaufmann, D., Kraay, A., & Zoido-Lobaton, P., (2000). "Governance matters," *Finance & Development*, 37(2), 10.
- Kaufmann, D., Kraay, A., & Zoido-Lobaton, P., (2002). "Governance Matters II – Updated Indicators for 2000/01." World Bank Policy Research Working Paper No. 2772, Washington, D.C.
- Kaufmann, D., Kraay, A., & Mastruzzi, M., (2009). "Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996-2008". World Bank Policy Research Working Paper No. 4978.
- Kaufmann, D., Kraay, A., & Mastruzzi, M., (2010). "The Worldwide Governance Indicators: Methodology and Analytical Issues". World Bank Policy Research Working Paper No. 5430.
- Keefer, P. and Knack, S., (1997), "Why Don't Poor Countries Catch Up? A Cross-National Test of an Institutional Explanation," *Economic Inquiry*, 35, 3, 590-602.
- Khan, M.H. (2002), "State Failure in Developing Countries and Strategies of Institutional Reform." In *Toward Pro-Poor Policies: Aid Institutions and Globalization. Proceedings of the Annual Bank Conference on Development Economics*, Bertil Tungodden, Nicholas Stern, and Ivar Kolstad, eds. pp. 155–186. Oxford: Oxford University Press and World Bank 2004.
- Khan, M. H. (2004) 'What is a "Good Investment Climate"?' In: Kochendörfer-Lucius, G and Pleskovic, B, (eds.), *Investment Climate, Growth, and Poverty*. World Bank (USA), pp. 77-83.
- Khan, M. H., (2005), What is a "Good Investment Climate"?, in Kochendörfer-Lucius, G. and Pleskovic, B. (eds) *Investment Climate, Growth, and Poverty*, Washington D.C.: World Bank.
- Khan, M. H., (2006), "Governance and Anti-Corruption Reforms in Developing Countries: Policies, Evidence and Ways Forward", <http://www.g24.org/khan0906.pdf>
- Kinda, T. (2010). "Investment climate and FDI in developing countries: firm-level evidence". *World Development*, 38(4), 498-513.
- Knack, S. and Keefer, P., (1995), "Institutions and Economic Performance: Cross Country Tests Using Alternative Institutional Measures," *Economics and Politics*, 7, 3, 207-27.
- Landes, D. S. (1998). *The wealth and poverty of nations: why some countries are so rich and some so poor*. WW Norton.
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. W., (1997), "Legal Determinants of External Finance," *Journal of Finance*, American Finance Association, vol. 52(3), pages 1131-50, July
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. W., (1998), "Law and Finance," *Journal of Political Economy*, University of Chicago Press, vol. 106(6), pages 1113-1155, December.
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. W., (1999), "The Quality of Government," *Journal of Law, Economics and Organization*, Oxford University Press, vol. 15(1), pages 222-79, April.
- Lee, I. H. and Rugman, A. M. (2012). "Firm-specific advantages, inward FDI origins, and performance of multinational enterprises", *Journal of International Management*, 18, issue 2, p. 132-146
- Leeson, P., & Dean, A. (2009). The democratic domino theory: An empirical investigation. *American Journal of Political Science*, 53(3), 533–551.
- Lin J, Pana, D., Davisb, S. J., Zhang, Q., Hed, K., Wang, C., Streetsf, D. G., Wuebblesg, D.J. & Guanc, D., (2014). "China's international trade and air pollution in the United States" *Proceedings of the National Academy of Sciences* (2014): 201312860
- López-Bazo, E., et al. (1998), "Regional Economic Dynamics and Convergence in the European Union." *Annals of Regional Science* 36, 1-28.
- López-Bazo E, Vayá, E and Artís M (2004). "Regional externalities and growth: evidence from European regions", *Journal of Regional Science* 44: 43-73.
- Lucas Jr, R. E. (1988). "On the mechanics of economic development". *Journal of monetary economics*, 22(1), 3-42.
- Masters, W. and McMillan, M., (2001), "Climate and Scale in Economic Growth," *Journal of Economic Growth*, 6, 3, 167-186.
- Masters, W. and Sachs, J. D., (2001), "Climate and Development," mimeo, Purdue University.
- Mauro, P., (1995), "Corruption and Growth," *Quarterly Journal of Economics*, 110, 3, 681-713.

- Mayer, T. & Zignano, S., (2011), "Notes on CEPII's distances measures: The GeoDist database", CEPII Working Paper 2011-25, Paris: CEPII
- McArthur, J. and Sachs, J., (2001), "Institutions and Geography: Comment on Acemoglu, Johnson and Robinson," National Bureau of Economic Research Working Paper no. 8114.
- Meentemeyer, V. (1989). "Geographical perspectives of space, time, and scale". *Landscape ecology*, 3(3-4), 163-173.
- Melling, A.D., Sachs, J.D. and Gallup, J.L., (1999). "Climate, water navigability and economic development". Center for International Development Working Paper , 24, Harvard University, Cambridge, MA.
- Montesquieu C de S, (1748), *The Spirit of the Laws*. Paris
- Myrdal, G. (1968). *Asian Drama – An Inquiry into the Poverty of Nations* Vol.3 (Harmondsworth, Penguin).
- Naude, W. A. & Krugell, W. F. 2007. "Investigating geography and institutions as determinants of foreign direct investment in Africa using panel data," *Applied Economics*, Taylor & Francis Journals, vol. 39(10), pages 1223-1233.
- North, D. (1984). 'Transaction Costs, Institutions, and Economic History', *Journal of Institutional and Theoretical Economics*, 140
- North, D.C. (1990). *Institutions, institutional change, and economic performance*, Cambridge: Cambridge University Press.
- North, D. C. (2003). "The Role of Institutions in Economic Development: Gunnar Myrdal Lecture". United Nations, Geneva.
- North, D. C., (2005), "Understanding the Process of Institutional Change", Princeton, NJ: Princeton University Press.
- North, D.C., Acemoglu, D., Fukuyama, F., & Rodrik, D. (2008). "Governance, growth, and development decision-making". Washington, DC: World Bank.
- Nougar, M., and Siscart, M., (2005), "Trade Raises Income: A Precise and Robust Result," *Journal of International Economics*, 65, 447-460.
- Ostrom, E., (2005) "Understanding Institutional Diversity", Princeton, NJ: Princeton University Press
- Pajunen, K. (2008). "Institutions and inflows of foreign direct investment: A fuzzy-set analysis". *Journal of International Business Studies*, 39(4), 652-669.
- Papaioannou, E. & Gregorios, S., (2008). "Democratisation and Growth", *Economic Journal*, Royal Economic Society, vol. 118(532), pages 1520-1551, October.
- Peet, R. (2011). "Sachs' Fifth Avenue", *Dialogues in Human Geography*, 1 (1) pp. 80-83.
- Persson, T., & Tabellini, G., (2009). Democratic Capital: The Nexus of Political and Economic Change, *American Economic Journal: Macroeconomics*, American Economic Association, vol. 1(2), pages 88-126, July.
- Pomeranz, K., (2000), "The Great Divergence: China, Europe and the Making of the Modern World Economy", Princeton, NJ: Princeton University Press.
- Quah, D., (1997b), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", *Journal of Economic Growth*, 2: 27–59
- Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modelling using Stata*. STATA press.
- Rey, Sergio J. and Montouri, B. D., (1998), "US Regional Income Convergence: A Spatial Econometric Perspective", *Regional Studies*, Vol. 33.2, pp. 143- 156.
- Rey, S., (2001), "Spatial Empirics for Economic Growth and Convergence", *Geographical Analysis*.
- Rodriguez, F., and Rodrik, D., (2001), "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," NBER Chapters, in: NBER Macroeconomics Annual 2000, Volume 15, pages 261-338 National Bureau of Economic Research, Inc.
- Rodrik, D. (2003). "Introduction." in D. Rodrik, ed., *In Search of Prosperity: Analytical Narratives on Economic Growth*, Princeton University Press, pp. 1–19.
- Rodrik, D., Subramanian, A., & Trebbi, F., (2004), "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development," *Journal of Economic Growth*, 9, 2, 131-165.
- Romer, P. M. (1990). "Endogenous Technological Change". *Journal of Political Economy*, 98(5 pt 2).
- Rugman, A. M. (1980). "Internalization as a general theory of foreign direct investment: A re-appraisal of the literature". *Review of World Economics*, 116(2), 365-379.
- Rugman, A.M. 1981. *Inside the multinationals*. London: Croom Helm.
- Sachs, J. and Warner, A., (1995), "Economic Reform and the Process of Global Integration (with discussion)," *Brookings Papers on Economic Activity*, 1, 1-118.
- Sachs, J. D., (2001) "Tropical Underdevelopment." NBER Working Paper No. W8119.

- Sachs, J., (2003), "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," National Bureau of Economic Research Working Paper no. 9490.
- Sala-i-Martin, X., (1997a), "I Just Ran 4 Million Regressions," National Bureau of Economic Research Working Paper no. 6252.
- Sala-i-Martin, X., (1997b), "I Just Ran 2 Million Regressions." *American Economic Review*, 87, 2, 178-83.
- Sato, T. (2012). "Empirical Analysis of Corporate Tax and Foreign Direct Investment". *Public Policy Review*, 8(1), 1-20.
- Seyoum, B. (2009). "Formal institutions and foreign direct investment". *Thunderbird International Business Review*, 51(2), 165-181.
- Seyoum, B. (2011). Informal institutions and foreign direct investment. *Journal of Economic Issues*, 45(4), 917-940.
- Sharma, K. & Bandara, Y. (2010). Trends, patterns and determinants of Australian foreign direct investment. *Journal of Economic Issues*, 64 (3), 661-676.
- Sheppard, E. (2011). "Geography, nature, and the question of development". *Dialogues in human geography*, 1(1), 46-75.
- Singh, H., & Jun, K. (1995). "Some new evidence on determinants of foreign direct investment in developing countries". World Bank Policy Research Working Paper, (1531).
- Sobel, R.S., and Leeson, P.T., (2007), "The Spread of Global Economic Freedom" In: *Economic Freedom of the World (2007)*, Gwartney, J., Hall and Lawson, R., (Eds), Fraser Institute.
- Smarzynska, B. K., & Wei, S. J. (2000). "Corruption and composition of foreign direct investment: Firm-level evidence". (No. w7969). National Bureau of Economic Research.
- Srholec, M. (2010). "A multilevel approach to geography of innovation". *Regional Studies*, 44(9), 1207-1220.
- Srivastava, M. (2004). "Moving beyond 'institutions matter': some reflections on how the 'rules of the game' evolve and change". Crisis States Discussion Paper Series No. 4. London School of Economics.
- Stein, E., & Daude, C. (2002). "Institutions, Integration and the Location of Foreign Direct Investment. *New Horizons for Foreign Direct Investment*", (548), 101.
- Sylos Labini, P., (2001); *Underdevelopment: A Strategy for Reform*. Cambridge University Press.
- Szirmai, A. (2012). "Proximate, Intermediate and Ultimate Causality: Theories and Experiences of Growth and Development", UNU-MERIT Working Paper Series, 2012-32.
- Verspagen, B. (2010). "The spatial hierarchy of technological change and economic development in Europe," *The Annals of Regional Science*, Springer, vol. 45(1), pages 109-132, August.
- Verspagen, B., (2012), "Stylized facts of governance, institutions and economic development. Exploring the institutional profiles database", UNU-MERIT Working Paper 2012-036
- Vieira, F. V., and Damasceno, A. O. (2009), "Is There a Primary Role of Institutions on Explaining Cross Country Differences in Income Levels and Long-Run Economic Growth?", XXXVII Encontro Nacional de Economia, 2009, Foz do Iguaçu.
- Walsh, J. P., & Yu, J. (2010). "Determinants of foreign direct investment: A sectoral and institutional approach". (Vol. 10). International Monetary Fund.
- Ward, M. D. and K. S. Gleditsch, (2008), "Spatial regression models". CA: Sage Publications.
- Wartenberg, D. (1985), "Multivariate Spatial Correlation: A Method for Exploratory Geographical Analysis". *Geographical Analysis*, 17: 263–283.
- Wei, S. J. (2000). "How taxing is corruption on international investors?". *Review of economics and statistics*, 82(1), 1-11.
- Welsch, H., (2003), "Corruption, Growth and the Environment: A Cross Country Analysis," mimeo, German Institute for Economic Research.
- Wernick, D. A., Haar, J., & Singh, S. (2009). "Do governing institutions affect foreign direct investment inflows? New evidence from emerging economies". *International Journal of Economics and Business Research*, 1(3), 317-332.
- Wheeler, D., & Mody, A. (1992). "International investment location decisions: The case of US firms". *Journal of International Economics*, 33(1–2), 57–76.
- World Bank. (1992), *Governance and Development* (Washington, DC: The World Bank).
- World Bank (2012). *World Development Indicators*.

# Valorization

Sources of growth and development have preoccupied economists since the inception of the discipline itself. There has been an explosion of empirical work building on earlier theoretical models of growth and development. This has been made possible due to the improved methods of quantification and collection of data on such measures of growth. Among these discussions of sources of growth, what are called the "ultimate" sources have gained much attention in the recent past. As mentioned in the introductory chapter, institutions and geography are among those "ultimate" sources.

Studies that have talked about institutions and geography mainly fall into the debate of which of these two sources of growth are more important, i.e., is it the geographic nature of the country or is it the institutional structure and quality that prevails in the country that ultimately changes the course of the country's growth path and hence its development pattern as well. In this academic debate, economists like Jeffery Sachs, John Gallup, Andrew Mellinger among many others argue that it is the geography of a country that decides the fate of its economic growth and development. On the other side of the debate fall economists such as Daron Acemoglu, Simon Johnson, James Robinson, Kenneth Sokoloff, Dani Rodrik etc, who argue that it is institutions that are the most important factors that lead a country into positive or negative paths. When a case for geography is made in latter's argument, it is that geography influences the type of institutions that develop in a country. One explanation given by (Acemoglu et al 2005), is that geography may have played a role in determining where colonizers decided to settle (settler colonies), and where they wanted to just extract resources (extractive colonies), since they could not adapt to the environment. Thus according to them, geography played a role in the set up of historical institutions, but not necessarily current ones. The institutional hypothesis thus says that geography has an indirect role influencing economic growth and development.

When considering which countries fall into developing and which into developed groups, we find a lot of clusters across the globe. These clusters might occur because of various spatially correlated attributes. It is not unrealistic to assume that countries that are neighbours share some common climatic conditions, physical geographic endowments, soil fertility conditions etc. When considering these 'absolute' aspects of geography, we tend to take into account only those time-invariant factors that cannot be influenced directly and in a short period of time. The reason for this is because 'absolute' geography allows us to circumvent the problem of endogeneity. Yet, if we consider how institutions and geography reinforce each other and in turn institutions and growth, we find that 'absolute' geography falls short as an 'ultimate' explanatory variable.

One channel through which institutions and geography are related to each other is through space. What much of the literature on the geography and institutional hypothesis misses is the socio-spatial interaction between countries. The political and economic conditions of neighbours influence a country's own political and economic structure and vice versa. For example, if a neighbour has political unrest, it is likely that a country will

have to adjust its migration policy to adjust and accommodate incoming refugees. This not only impacts the political but also the economic fabric of the country. Neighbouring country spillover effects lead to spatial correlations that cannot be ignored in explaining the economic performance of a certain country. These spatial correlations could also lead to more direct effects on the institutional structure of a country. For example, what started off as protests against the Tunisian government in 2010 resulted not only in changes in the political system within the country, but also to major protests, government overthrows and governmental changes in many neighbouring countries such as Algeria, Libya, Yemen, etc. Thus, we take into account the 'relative' geography of countries. The main contribution of this thesis is the inclusion of this 'relative' geography in explaining the roles of geography and institutions in economic performance. We contend that it is not only the common historical experiences but also common or similar cultural practices that neighbouring countries may share, that lead to similar institutional structures.

Studying 'relative' geography to explain economic phenomena is not new. Spatial influence in regional income convergence, technology diffusion, migration, housing prices etc have been studied using explicit geographic components such as bilateral distance between countries, dyadic trade distances, distance from school districts and various forms of social distance as well. Related to institutions, Gleditsch and Ward (2000), and Cho (2003) among many others have studied the results of spatial distance on the levels of democracy in countries. Taking cue from these studies, in this thesis, we attempt to understand the underlying spatial nature of the institutional qualities of countries.

In order to understand this spatial structure, we study the institutional structure with respect to a country's economic performance, financial performance as well as its 'absolute' geography. All these measures are not easily defined by single variables. For example, GDP per capita is often used as the measure of economic development of a country. However, it does not reflect the various aspects of development that need to be accounted for when talking about development, such as growth rates of GDP, levels of consumption and savings in the economy, to determine the health of the economy overall. Similarly, we construct measures for institutional quality, financial performance and geography. We construct these measures using Principal Component Analysis (PCA) and Canonical Correlation Analysis (CCA) with spatial weighting. This is another contribution this thesis makes. The methodologies are explained in Chapter 2 of the thesis. Using our Spatial PCA and Spatial CCA, we construct and interpret the results. This method proves to be useful in the study other spatial correlations at cross country and regional levels. It is possible to use weighting schemes based on the study at hand, and this can easily be routinized on Matlab, Stata and other software.

The debate of primacy mentioned above has one more player. Open economies that trade more with other countries are said to grow faster (Frankel and Romer 1999, Dollar and Kraay 2001). Rodrik, Subramaniam and Trebbi 2004, study these three contesting theories together to see which one has the most impact on economic performance. They find that institutions are more important than both geography and trade. These studies too ignore the influence of neighbouring countries on both the external trade of a country and its institutional quality. In Chapter 4 we analyze the Rodrik, Subramaniam and Trebbi (2004) using our measures (using SPCA and SCCA). Our analysis confirms that

spatially weighting institutions do indeed strengthen the results of the Rodrik et al type regressions. This makes a case for more closely looking at the spatial influence that institutions are subject to, and thus, it becomes more policy relevant to look at these kind of influences.

Institutions, their structure, quality and governance are certainly important in determining the economic performance of a country. But how important is the perception of quality of governance important to economic entities within a country? In order to understand this, the last chapter in this thesis analyzes a multi-level framework. It studies the impact of firm and country variables on a firm's foreign investment level. Multi-level analysis is rather new to the field of economics. The novelty in this particular study is that it is applied to perceptions of institutional quality in relation to FDI. While a firm's decision to enter a country to do business is a micro level decision, institutional quality is often a country level aggregate measure. In this analysis, we attempt to look at these two different levels and their interactions with respect to inward FDI. What we find is that a firm's idiosyncratic factors and the host country's macroeconomic factors are most important in attracting inward FDI. We find, interestingly, that institutional quality does not play a significant role in FDI decisions of firms.

Much has been said about developing countries and their institutions and governance. It has often been said that improving institutional quality would lead to higher FDI inflows and hence would lead to more growth in developing economies. This has been propagated by international financial organizations such as the World Bank and the IMF. Most studies which either agree or disagree with this view are exclusively country level or exclusively firm level analyses. Our results are pertinent in that, when both levels are accounted for in a multi-level framework, we find that it is the economic fundamentals that drive firms to invest in foreign countries and not their perceptions of governance of these host countries.

This thesis thus stresses on the importance of scale in studying institutional quality and its relationship to economic performance. In a highly globalized world such as ours, it becomes even more important to study these interactions carefully. The results of this study are useful for both poor and emerging countries. These countries run the risk of falling into short run growth paths rather than the more stable long run growth paths if they overlook important aspects of growth and development. Even rich countries that are currently at the top may lose their positions to other countries if the global centre of economic activity shifts to other parts of the world.



# Samenvatting

Een populaire bevinding uit de empirische literatuur over economische groei en ontwikkeling is dat de economische en sociale instituties van een land economische activiteit stimuleren. Het beleid van deze instituties kan de groei en ontwikkeling van het land bevorderen of belemmeren. De economische ontwikkeling van landen wordt ook vaak toegeschreven aan geografische aspecten. Voorstanders van deze literatuur benadrukken dat de geografische aspecten van een land, zoals tropische ziekten of natuurlijke hulpbronnen, de groeivoorzichten van een land beïnvloeden. Voorstanders van de literatuur over instituties vinden geografische aspecten eveneens van belang maar redeneren dat geografische aspecten een effect hebben op institutionele aspecten en enkel een indirect effect hebben op economische groei en ontwikkeling. Deze literatuur houdt geen rekening met de ruimtelijke interacties tussen de landen. Door middel van ruimtelijke interactie hebben buurlanden soortgelijke instituties. Buurlanden hebben deels een gemeenschappelijke geschiedenis en beïnvloeden elkaar meer dan landen die zich in een ander werelddeel bevinden. In dit proefschrift wordt onderzocht of de resultaten van de literatuur over instituties kan worden beïnvloed door rekening te houden met deze ruimtelijke interacties.

In het eerste hoofdstuk wordt de literatuur en de motivatie voor deze scriptie samengevat. In het tweede hoofdstuk introduceren we de methodologie van ruimtelijke principale componentenanalyse (SPCA) en ruimtelijke canonische correlatie analyse (SCCA). Met behulp van SPCA en SCCA methodologie construeren we samenvattende indicatoren. De doelstelling van deze methode is om de ruimtelijke correlatie van de data te maximaliseren. Terwijl SPCA wordt gebruikt om een database met verschillende variabelen samen te vatten, wordt SCCA gebruikt om de ruimtelijke correlatie tussen twee databases samen te vatten. In dit hoofdstuk vergelijken we de twee methoden met de hulp van voorbeelden. We vergelijken onze resultaten met de Wartenberg methode van ruimtelijke principale componentenanalyse. Om de voordelen van deze specifieke werkwijzen te tonen introduceren we een voorbeeld met geografische data. Wij vinden dat onze methode resultaten geeft die makkelijker te interpreteren zijn dan de eerder beschikbare methode.

De studie in hoofdstuk 3 maakt gebruik van de methode beschreven in hoofdstuk 2. De doelstelling van hoofdstuk 3 is om indicatoren te construeren die verschillende multidimensionale verschijnselen samenvatten. We construeren nieuwe indicatoren voor de instituties, geografische dimensies en de economische en financiële ontwikkeling van landen. We evalueren hoe onze indicator van de kwaliteit van instituties gerelateerd is aan de andere bovengenoemde indicatoren. Deze studie draagt bij aan het onderzoek naar interacties tussen de bovengenoemde fenomenen door middel van een ruimtelijk verhaal. Wij vinden dat de eerste principale componenten van de SPCA en SCCA overeenkomend zijn met de resultaten beschreven in de literatuur wanneer deze ruimtelijke gewichten hebben. Maar we vinden ook dat de tweede principale componenten met ruimtelijke gewichten nieuwe resultaten tonen. Terwijl economische ontwikkeling sterk

gecorrleerd is met institutionele kwaliteit is economische groei sterk gecorrleerd met meer overheidsingrijpen. De resultaten tonen ook aan dat geografische factoren niet noodzakelijkerwijs bepalend zijn voor het economisch lot van een land. Landen met vergelijkbare geografische factoren hebben niet altijd dezelfde of soortgelijke instituties. We concluderen dat het mogelijk is om institutionele veranderingen in een land beter te begrijpen wanneer er rekening wordt gehouden met de ruimtelijke interacties.

In het volgende hoofdstuk maken we gebruik van deze ruimtelijk geconstrueerde indicatoren van instituties en geografie, en ook een proxy voor economische handel, om te analyseren welke indicatoren economische ontwikkeling het beste verklaren. Met behulp van het kader beschreven in Rodrik et al. (2004) vinden we dat instituties de belangrijkste verklarende indicator is maar enkel wanneer economische ontwikkeling wordt gemeten als BBP per capita. Wanneer we gebruik maken van een bredere definitie van economische ontwikkeling (hetgeen we construeren met behulp van SPCA) zijn de resultaten niet langer toonaangevend van het belang van instituties. De resultaten suggereren ook dat instituties een negatief effect kunnen hebben op economische ontwikkeling, hetgeen betekent dat de kwaliteit van instituties gecorrleerd is met economische ontwikkeling maar dat er geen causaal verband is.

Schaal en ruimtelijke interactie hoeft niet te worden beperkt tot het cross-country niveau. In hoofdstuk 5, verleggen we de focus naar een multi-level analyse. We analyseren de determinanten van buitenlands bezit in binnenlandse bedrijven met behulp van een multi-level kader. Met behulp van structurele en institutionele data van bedrijven en macro-economische en macro-institutionele variabelen schatten we een multi-level logit model. Dit hoofdstuk draagt bij aan de bestaande literatuur door middel van een multi-level model dat controleert voor het effect van micro en macro niveau factoren op buitenlandse investeringen. We vinden dat structurele kenmerken en (institutionele) belemmeringen op het bedrijfsniveau de belangrijkste determinanten zijn van buitenlandse investeringen. Op het macro niveau tonen de resultaten aan dat macro-economische variabelen die de winst van het bedrijf kunnen beïnvloeden belangrijk zijn maar daartegenin zijn instituties niet invloedrijk in het aantrekken van buitenlandse investeringen.

In hoofdstuk 6 van het proefschrift worden de resultaten samengevat. In conclusie, dit proefschrift probeert de ruimtelijke interacties te analyseren die invloed hebben op de kwaliteit van instituties. Een bijdrage van dit proefschrift was het ontwikkelen van een methodologie om de ruimtelijke correlatie tussen variabelen te maximaliseren. Met behulp van deze methodologie, hebben we geprobeerd om de relatie tussen instituties, geografie en economische ontwikkeling beter te begrijpen. De multi-level analyse is gericht op het ontrafelen van percepties van instituties op macro-niveau en de percepties van instituties op bedrijfs-niveau om aan te tonen of micro of macro factoren meer invloed hebben op het bedrijfsleven.

# Biography

Samyukta Bhupatiraju studied her Bachelor's degree in Chennai and completed her Masters degree in Economics at University of Hyderabad. She studied International Economics at the Kiel Institute for the World Economy in 2008-09 before joining UNU-MERIT. While at UNU-MERIT she worked on the EXPLORE project launched under the DIME (Dynamics of Institutions and Markets in Europe) network in 2009. She was also a part of the "Institutions, Governance and Long-term Growth" project by Maastricht Graduate School of Governance, Maastricht University and the Agence Française de Développement (AFD).





UNITED NATIONS  
UNIVERSITY

UNU-MERIT

## 2014

### 86. Samyukta Bhupatiraju

The Geographic Dimensions of Growth and Development

### 85. François Lafond

The evolution of knowledge systems

### 84. Annalisa Primi

Promoting Innovation in Latin America: What Countries Have Learned (and What They Have Not) in Designing and Implementing Innovation and Intellectual Property Policies

### 83. Fatoumata Lamarana Diallo

Evaluation of Meal and Deworming Programs for Primary Schools in Rural Senegal

## 2013

### 82. Anant Kamath

Information Sharing through Informal Interaction in Low-Tech Clusters

### 81. Flavia Pereira de Carvalho

What we talk about when we talk about Brazilian Multinationals: an investigation on Brazilian FDI, economic structure, innovation and the relationship between them

### 80. Jun Hou

Complementarity in Innovation and Development: A Cross-country Comparison

### 79. Rufin Baghana

Impacts of Government Incentives to R&D, Innovation and Productivity: A Microeconomic Analysis of the Québec Case

### 78. Lilia I. Stubrin

High-Tech Activities in Emerging Countries: A Network perspective on the Argentinean biotech activity

## 2012

### 77. Abdul Waheed

Innovation Determinants and Innovation as a Determinant: Evidence from Developing Countries

### 76. Bilal Mirza

Energy Poverty and Rural Energy Markets in Pakistan

### 75. Benjamin Engelstätter

Enterprise Software and Video Games: An Empirical Analysis

### Fulvia Farinelli

Natural Resources, Innovation and Export Growth: The Wine Industry in Chili and Argentina

### Rodolfo Lauterbach

Innovation in Manufacturing: From Product Variety and Labor Productivity Growth to Economic Development in Chile

### 74. Kirsten Wiebe

Quantitative Assessment of Sustainable Development and Growth in Sub-Saharan Africa.

### 73. Julio Miguel Rosa

Organizational Strategies, Firms' Performance and Spatial Spillovers. The Canadian Case in Research and Development.

### Johannes Wilhelmus Marie Boels

Joseph Schumpeter, honderd jaar economische ontwikkeling. Een historisch-theoretische beschouwing.

## 2011

### 72. Daniel Vertesy

Interrupted Innovation: Emerging economies in the structure of the global aerospace industry.

### 71. Tina Saebi

Successfully managing alliance portfolios: an alliance capability view.

### 70. Nora Engel

Tuberculosis in India - A case of innovation and control.

### 69. Evans Mupela

Connectivity and growth in Sub-Saharan Africa: The role of communication satellites

**68. Nantawan Kwanjai**

Cross cultural intelligence amid intricate cultural webs – A tale of the UnDutchables in the land of 1002 smiles

**67. Lina Sonne**

Innovation in Finance to Finance Innovation: Supporting pro-poor entrepreneur-based innovation

**2010**

**66. Fernando Santiago**

Human Resources Management Practices and Learning for Innovation in Developing Countries: Pharmaceutical Firms in Mexico

**65. Zakaria Babutsidze**

Essays on Economies with Heterogenous Interacting Consumers

**64. Bertha Vallejo**

Learning and Innovation Under Changing Market Conditions: The Auto Parts Industry in Mexico

**63. Donatus Ayitey**

Technical Change, Competitiveness and Poverty Reduction: A Study of the Ghanaian Apparel Industry

**62. Sergey Fillipov**

Multinational Subsidiary Evolution: Corporate Change in New EU Member States

**61. Asel Doranova**

Technology Transfer and Learning under the Kyoto regime; Exploring the Technological Impact of CDM projects in developing countries

**2009**

**60. Alexis Habiyaremye**

From Primary Commodity Dependence to Diversification and Growth". "Absorptive Capacity and Technological Catch Up in Botswana and Mauritius".

**59. Yoseph Getachew**

The Role of Public Capital in Economic Development

**58. Sandra Leitner**

Embodied Technological Change and Patterns of Investment in Austrian Manufacturing

**57. Semih Akçomak**

The Impact of Social Capital on Economic and Social Outcomes

**56. Abraham Garcia**

The Role of Demand in Technical Change

**55. Saurabh Arora**

Coherence in socio-technical systems: a network perspective on the innovation process

**2008**

**54. Rutger Daems**

Medicines for the developing world

**53. Johannes Hanel**

Assessing Induced Technology - Sombart's Understanding of Technical Change in the History of Economics

**52. Rifka Weehuizen**

Mental Capital: the economic significance of mental health

**51. Danielle Cloodt**

The relationship between R&D partnership formation, social embeddedness and innovative performance

**50. Sabine Fuss**

Sustainable Energy Development under Uncertainty

**2007**

**49. Tobias Kronenberg**

Reconciling Environmental Conservation with Economic Prosperity: The Feasibility of Double Dividends in the Short and Long Run

**48. Viktoria Kravtsova**

Assessing the Impact of Foreign Direct Investment in Transition Economies

**47. Suhail Sultan**

The Competitive Advantage of Small and Medium Sized Enterprises: The Case of Jordan's Natural Stone Industry

**2006**

**46. Bulat Sanditov**

Essays on Social Learning and Imitation

**45. Mamata Parhi**

Dynamics of New Technology Diffusion: A Study of the Indian Automotive Industry

**44. Andreas Reinstaller**

Social structures and the innovation process: Their role in the demand of firms and consumers

**43. Rose Kiggundu**

Innovation systems and development : the journey of a Beleaguered Nile Perch Fishery in Uganda

**42. Thomas Pogue**

The Evolution of Research Collaboration in South African Gold Mining: 1886-1933

**41. Geoffrey Gachino**

Foreign Direct Investment, Spillovers and Innovation: The Case of Kenyan Manufacturing Industry

**40. Önder Nomaler**

Technological Change, International Trade and Growth - An Evolutionary, Multi-Agents-Based Modeling Approach

**2005**

**39. Samia Satti Osman Mohamed-Nour**

Change and Skill Development in the Arab Gulf Countries

**38. Elad Harison**

Intellectual Property Rights: Economics and Policy Analysis

**37. Daniel Dalohoun**

The relationship between R&D partnership formation, social embeddedness and innovative performance: a multi-level approach of social embeddedness

**36. Müge Ozman**

Networks, Organizations and Knowledge

**35. Bas Straathof**

Product variety and economic growth - The counteracting effects of scale and idiosyncrasy

**34. Wilfred Schoenmakers**

Knowledge Flows between Multinational Companies: A Patent Data Analysis

**33. Myriam Cloodt**

Mergers and Acquisitions (M&As) in High-Tech Industries: Measuring the Post-M&A Innovative Performance of Companies

**2004**

**32. Paola Criscuolo**

R&D Internationalisation and Knowledge Transfer. Impact on MNEs and their Home Countries

**31. Maarten Verkerk**

Trust and Power on the Shop Floor

**30. Gottfried Leibbrandt**

Adoption, harmonization and succession of network technologies across countries

**29. Mark Sanders**

Skill Biased Technical change - Its Origins, the Interaction with the Labour Market and Policy Implications

**2003**

**28. Nadine Roijackers**

Inter-firm cooperation in high-tech industries: a study of R&D partnerships in pharmaceutical biotechnology

**27. Viki Sonntag**

Speed, Scale and Sustainability

**26. Masaru Yarime**

From End-of-Pipe Technology to Clean Technology

**25. Stéphane Malo**

The combinatorial Chemistry Revolution - Sustaining a Superior Performance Position through Technological Learning

**2002**

**24. Annelies Hogenbirk**

Determinants of Inward Foreign Direct Investment: the Case of the Netherlands

**2001**

**23. John Adeoti**

Technology Investment in Pollution Control in Sub-Saharan Africa: The Case of the Nigerian Manufacturing Industry

**22. Edward Huizenga**

Innovation Management: How Frontrunners Stay Ahead. An Empirical Study on Key Success Factors in the ICT sector

**2000**

**21. Machiel van Dijk**

Technological Change and the Dynamics of Industries. Theoretical Issues and Empirical evidence from Dutch Manufacturing

**1999**

**20. Jan Cobbenhagen**

Managing Innovation at the Company Level: A Study on Non-Sector-Specific Success Factors

**19. Marjolein Caniëls**

Regional Growth Differentials: The Impact of Locally Bounded Knowledge Spillovers

**1998**

**18. Aldo Geuna**

Resource allocation and knowledge production: Studies in the economics of university research

**1996**

**17. Reinoud Joosten**

Dynamics, Equilibria, and Values

**16. Hugo Kruiniger**

Investment, R&D, and the Financing Decisions of the Firm

**1995**

**15. Hans van Meijl**

Endogenous Technological Change: The Case of Information Technology. Theoretical Considerations and Empirical Results

**14. René Kemp**

Environmental Policy and Technical Change. A Comparison of the Technological Impact of Policy Instruments

**13. Rohini Acharya**

The Impact of New Technologies on Economic Growth and Trade. A Case Study of Biotechnology

**12. Geert Duysters**

The Evolution of Complex Industrial Systems. The Dynamics of Major IT Sectors

**11. Marjan Groen**

Technology, Work and Organisation, A Study of the Nursing Process in Intensive Care Units

**1994**

**10. Huub Meijers**

On the Diffusion of Technologies in a Vintage Framework; Theoretical Considerations and Empirical Results

**9. Theon van Dijk**

The Limits of Patent Protection. Essays on the Economics of Intellectual Property Rights

**8. Hans Voordijk**

Naar Integrale Logistiek in Bedrijfsketens, Ontwikkelingen in de Bouw

**1993**

**7. Paul Diederer**

Technological Progress in Enterprises and Diffusion of Innovations. Theoretical Reflections and Empirical Evidence.

**6. Ben Dankbaar**

Economic Crisis and Institutional Change. The crisis of Fordism from the perspective of the automobile industry

**5. Hanno Roberts**

Accountability and Responsibility: The Influence of Organisation Design on Management Accounting

**1992**

**4. Bart Verspagen**

Uneven Growth Between Interdependent Economies. An Evolutionary View on Technology Gaps, Trade and Growth

**3. Sjoerd Romme**

A Self-organization Perspective on Strategy Formation

**1989**

**2. John Spangenberg**

Economies of Scale, and Atmosphere in Research Organisations

**1988**

**1. John Hagedoorn**

Evolutionary and heterodox innovation analysis : a study of industrial and technological development in process control and information technology

