Maria Ziegler

Institutions, Inequality and Development

The book focuses on the linkages between institutions, inequality and development. It analyzes formal political institutions, in particular the relationship between democracy and human development. It also centers on informal social institutions leading to the exclusion of population groups such as women and indigenous people. To measure these institutions in the case of gender inequality the Social Institutions and Gender Index (SIGI) and its five subindices are proposed and for ethnic inequality dummy variables indicating ethnic origin are used. The dissertation shows that formal and informal institutions affect human development, the governance of a society and inequality.

Maria Ziegler, born in Stollberg (Germany) in 1980, studied Public Policy and Management at the University of Konstanz and the University Paris 1, Panthéon-Sorbonne. Between 2006 and 2010 she was a Ph.D. candidate at the Department of Economics at the University of Göttingen. During her studies she also worked as a consultant for various international agencies in Latin America and Africa.

www.peterlang.de

Institutions, Inequality and Development

Göttinger Studien zur Entwicklungsökonomik
Göttingen Studies in Development Economics

Bd./Vol. 31

PETER LANG

Maria Ziegler

Institutions, Inequality
and Development

PETER LANG

Internationaler Verlag der Wissenschaften

Cover design:
Olaf Glöckler, Atelier Platen, Friedberg

Cover illustration by Rolf Schinke

Meinen großartigen Großeltern

Editors Preface

Progress in achieving the Millennium Development Goals has been uneven and unsatisfactory in many dimensions. In this volume, Maria Ziegler examines the unsatisfactory progress by highlighting the role of informal and formal institutions structuring social interactions, distributing power in a society and therefore affecting the freedoms to choose a life according to one's needs and preferences.

The first essay investigates the influence of democracy on progress in education and health in developing countries. The theoretical part tries to answer the question why democracy influences health and education focusing on redistribution as well as accountability and responsiveness in political systems. Secondly, it addresses the question of whether this effect depends upon other factors such as inequality, the level of development, education of the population and ethnic diversity. Using international panel data a robust positive and significant effect of democracy promoting health and education is found. However, the interaction effects of democracy with GDP per capita, inequality, ethnic fractionalization and education turn out to be insignificant or not robust. Carefully interpreted, democratic institutions are themselves important for human development and less the circumstances under which they occur.

There is another type of institutions, namely informal social institutions that should not be neglected in the study of development outcomes. These informal institutions are often taken-for-granted, and provide role models and social exclusion mechanisms. Those social institutions that are related to gender inequality and distribute power between men and women in daily life build the focus of the next three essays.

Essay 2 centers on the measurement of social institutions related to gender inequality proposing the Social Institutions and Gender Index (SIGI) and its five subindices Family code, Civil liberties, Physical integrity, Son preference and Ownership rights, which are now officially used by the OECD Development Centre. In the first step, the five one-dimensional subindices are constructed by aggregating variables of the OECD Gender, Institutions and Development Database with polychoric principal component analysis. In a second step, the subindices are combined using the Foster-Greer-Thorbecke poverty measurement approach

to calculate the SIGI. Preliminary analyses show that the SIGI is empirically non-redundant to other gender-related indices and can be used to compare the societal situation of women in over 100 developing countries.

Essay 3 investigates whether the newly proposed indices can explain development outcomes such as female education, child mortality, fertility and governance (rule of law and civil liberties). In particular, the study aims at separating the exploratory value of the SIGI from the one of religion, region, the political system and income. The theoretical motivation is based on household bargaining and investment models. The empirical results show a robust significant effect of at least one of the social institutions indices on the development outcomes. Controlling for religion, political system, geography and the level of economic development, higher inequality in social institutions related to gender is associated with worse development outcomes.

Essay 4 concentrates on the relationship between social institutions and gender inequality and governance focusing on corruption. Embedded into the literature on gender inequality and corruption, the study highlights that a worse social status of women in a society measured by a higher inequality in social institutions related to gender is associated with a higher perceived level of corruption in a society even if one controls for representation of women in the society and democracy as well as other factors proposed by the literature.

The last essay focuses on another marginalized group, the indigenous population, whose situation is not clearly covered by the Millennium Development Goals but deserves attention as they are overrepresented among the world's poor. Essay 5 analyzes the relationship between ethnic origin and health inequality in Bolivia and shows that social exclusion and institutional mechanisms – measured with significant dummies for ethnic origin – are relevant factors for racial differences in health. However, this perspective might lead to unsuccessful policy interventions as it does not consider other factors that are associated with both ethnic origin and health, such as material wealth, urban-rural differences, geographical location and other household and maternal characteristics. The two major results are that first ethnic origin matters but that there is heterogeneity in health outcomes within the indigenous population. Secondly, health knowledge and mother's education could be responsible for health outcomes differences between ethnic groups, and the role of both variables as a pathway between ethnic origin and health outcomes should be investigated further.

Overall, the volume of Maria Ziegler makes an important contribution to the empirical literature on the linkage between institutions, inequality and economic and human development.

Acknowledgements

The way was long, the way was hard, but it was worth the effort.

First of all, I would like to thank my supervisor Professor Stephan Klasen for giving me the opportunity to write this dissertation and to work in a dynamic, international environment. He supported me during the whole time span of the dissertation, he was demanding but also understanding. His scientific input was always of high value for my work and I have to admit that I am always impressed by the diversity of knowledge he has and the range of the projects he manages. I would also like to thank my second supervisor Junior Professor Carola Grün for her kind and constructive comments, which improved my work considerably. Professor Matin Qaim agreed to be the third examiner of my doctoral dissertation and I would like to thank him for this.

My colleagues deserve mentioning as well, as they filled my days with humor, motivation, understanding and inspiration. I thank my co-authors Sebastian Vollmer, Boris Branisa and Elena Gross. In particular, I thank Michaela Beckmann for administrative help, Felicitas Nowak-Lehmann Danziger for her warm and scientific support at the end of the thesis and I thank Kenneth Harttgen, Inmaculada Martínez-Zarzoso, Adriana Cardozo, Jan Priebe, Johannes Gräb and Tobias Lechtenfeld for scientific advice. I also want to thank all my other colleagues - there are too many to name them all.

On my way many more scientists gave me helpful comments. I thank Christian Bjørnskov, Mark Dincecco, Isabel Günther, Dierk Herzer, Johannes Jütting, Denis Drechsler, Tatyana Krivobokova, Juan R. de Laiglesia, Oleg Nenadić, Jean-Marc Siroën, Stefan Sperlich, Walter Zucchini and other anonymous referees. In addition, I thank the participants of the following conferences: The Institutional and Social Dynamics of Growth and Distribution (Pisa, 2007), Conference of the International Society for Comparative Economic Studies (Sao Paulo, 2008), International Economic Association's conference (Istanbul, 2008) American Economic Association's annual conference (San Francisco, 2009), International Conference on Gender and the Global Economic Crisis (New York, 2009), North American

Summer Meeting of the Econometric Society (Boston, 2009), Far East and South Asia Meeting of the Econometric Society (Tokyo, 2009) and Singapore Economic Review Conference (Singapore, 2009).

I also want to thank the German National Academic Foundation, which provided me with a scholarship without which it would not have been possible to write this dissertation. I also acknowledge travel funds from the University of Göttingen, the Universitätsbund Göttingen and the Verein für Socialpolitik.

Finally, my deepest thanks go to my grandparents and André as well as all friends, in particular Sophia, Karin, Susanna, Laura, Franzi, Theresa, Gesine, Anne, Doreen, Knut, Julian, Felix, Marco and Tobi, who encouraged me to go this way. They gave me stability during times of crisis and support whenever I needed it. Last not least I thank my parents that believed in me and supported me as long as they could.

Contents

List of Figures

List of Tables

List of Abbreviations

BCG	Bacille Calmette-Guérin vaccine against tuberculosis
ECA	Europe and Central Asia
FPC	First Principal Component
CPI	Corruption Perception Index
Dem.	Democratic
DHS	Demographic and Health Surveys
DPT	Diphtheria, pertussis and tetanus
EAP	East Asia and Pacific
EPI	Expanded Program on Immunization
FGT	Foster-Greer-Thorbecke
GDI	Gender-Related Development Index
GDP	Gross Domestic Product
GEM	Gender Empowerment Measure
GEM revised	revised Gender Empowerment Measure
GGG	Global Gender Gap Index
GGI (capped)	Gender Gap Index
GID	Gender, Institutions and Development
HC	Heteroscedasticity-consistent
HDI	Human Development Index
ICRG	Corruption in Government Index from the International Country Risk Guide
IMF	International Monetary Fund
LAC	Latin America and Caribbean
MDGs	Millennium Development Goals
MJCA	Multiple Joint Correspondence Analysis
OECD	Organisation for Economic Co-operation and Development
PCA	Principal Component Analysis
PPP	Purchasing Power Parity
Rep.	Republic
SA	South Asia
SIGI	Social Institutions and Gender Index
UCDP/PRIO	Uppsala Conflict Data Program/ International Peace Research Institute, Oslo
UNAIDS	The United Nations Joint Programme on HIV/AIDS
UNDP	United Nations Development Programme

WECON	Women's Economic Rights index
WHO	World Health Organization
WIDER	World Institute for Development Economics Research
WOPOL	Women's Political Rights index
WOSOC	Women's Social Rights index
WOSOC	Women's Social Rights index

Introduction

The State of Development

In September 2010, the Summit on the Millennium Development Goals (MDGs) - the high-level plenary meeting of the General Assembly - will take place to review the implementation of the MDGs and to identify areas of action to achieve them by 2015. Although some progress in terms of fighting poverty and hunger, improving health and education and other aspects of the MDGs has been achieved, progress has been uneven. In 2005, there were still 1.4 billion people living on less than $1.25 a day and further progress has been reversed or delayed due to the world economic crisis (United Nations, 2009). The number of people suffering from hunger rose to 1.02 billion in 2009, 129 million children were underweight and 195 million under the age of 5 were stunted (United Nations, 2010). Progress towards achieving universal primary education in developing countries has been noticeable, though there are still more than 10% of children of primary school age that are not enrolled (United Nations, 2010). Under-five mortality declined from 93 per 1,000 live births in 1990 to 67 per 1,000 live births in 2007, which is still short of the goal of reducing child mortality by two thirds between 1990 and 2015 (United Nations, 2009).

A look at the overall performance on these or other indicators neglects large disparities due to income itself, urban-rural differences and other inequalities related to gender, language, ethnicity or disability. Social exclusion and a lack of participation have been diagnosed as the main drivers of group-based disparities and represent a further dimension of poverty (United Nations, 2009, 2010). This is mainly reflected in MDG 3, "Promote gender equality and empower women". However, progress regarding gender equality remains low. In 2007, out of 171 countries only 53 had achieved gender parity in primary and secondary education. The gender gap in secondary schooling is even more severe and particularly evident in Sub-Saharan Africa, where the girls' to boys' enrolment ratio is only 79% in 2007. Moreover, gender gaps persist on the labor market and in the political arena. For example, only 18% of the parliamentary seats were held by women in January 2009 (United Nations, 2009, 2010).

The picture given here shows that there still remains much to be done and makes clear that actions have to be identified in the multiple dimensions of development to accelerate progress towards achieving the MDGs. Such a multidimensional view towards development constitutes the basis of this work. It is inspired by Sen's notion of development as freedom or expansion of capabilities (e.g. Sen, 1999b, 2003). Sen's capability approach is based on the two concepts of functionings and capabilities. Functionings are the "doings and beings" of a person, her actions and the status that she values and enjoys, like being healthy, being educated, achieving self-respect or participating in social life. Capabilities refer to different combinations of functionings a person is able to achieve, covering the notion of freedom to choose the kind of life one would like. With his approach, Sen inspired the emergence of the pluralist and integrative conception of "human development" and its operationalization in the Human Development Index of the United Nations Development Program (UNDP). It is not only income but also health and education among other factors that enable people to live the life they value.

Sen leaves the question of what are valuable achievements and freedoms wide open and does not make explicit a list of fundamental universal capabilities (Nussbaum, 2003; Gaspers and van Staveren, 2003). However, he highlights the importance of public deliberation, participatory processes and political freedoms for social choice and the constitution of values and development goals, as in such a context people are able to advance their own case and act as agents of the development process.

Political Institutions and Human Development

Sen's discussion about valuable capabilities and the formation of these values centers on social interactions and draws attention to institutions in general. North (2001, p. 97) defines institutions as "the humanly devised constraints that structure political, economic and social interaction. They consist of both informal constraints [...] and formal rules [...]." Institutions are the rules of the game. They create order, reduce uncertainty and affect the prosperity of a nation by reducing transaction costs and regulating contract enforcement and property rights protection. A very important feature is the distributional effect of institutions (World Bank, 2005). In particular, institutions distribute power in a society and therefore they affect the capabilities of people to choose between different ways of living. Sen (1999b) emphasizes democratic political institutions that create the environment for social choice and value formation where all people can actively and equally participate in an open deliberation process. Therefore, besides the intrinsic value of democracy, democratic institutions help to produce responsive policies and to hold politicians and bureaucrats accountable. It is the purpose of the **Essay 1** to investigate whether Sen's argument withstands empirical evidence and to an-

swer the question of which political system is the best for obtaining a high level of human development measured with the non-income dimensions education and health.

There are in fact examples that challenge Sen's claim. Present-day Singapore, an autocracy, is a high income country with a high life expectancy at birth of 80 years and a high literacy rate of 94%.[1] The development path of his own country inspired Singapore's former President, Lee Kuan Yew, to put forward the famous Lee hypothesis according to which authoritarian rule is more efficient than democratic government and therefore beneficial to economic development (and thus to human development as well) (Sen, 1999a). Also, relatively poor Cuba has managed to achieve a very high life expectancy rate at birth of 79 years and an adult literacy rate of almost 100%.[2] On the other hand, the democratic country of India, for example, has a life expectancy at birth of only 63 years and a literacy rate of 66%.[3]

Sen (1999a, p. 6) calls this "sporadic empiricism" and this is certainly true. Nevertheless, controversies put forward in the theoretical literature do uphold the question about the power of democracy. First, there is a controversy concerning the contradictory effects of property rights protection and redistribution in democratic societies on growth and well-being (e.g. Mohtadi and Roe, 2003; Alesina and Rodrik, 1994; Baum and Lake, 2003). Secondly, the causal direction is not clear: is democracy a cause or a consequence of the development process (e.g. Lipset, 1959; Persson and Tabellini, 2007a; Glaeser et al., 2007)? Thirdly, there is a debate as to which conditions are necessary for democracy to have a positive effect on human development (e.g. Keefer and Khemani, 2005). Empirical studies do not provide a coherent answer to these questions and they have their limitations (e.g. Lake and Baum, 2001; Baum and Lake, 2003; Navia and Zweifel, 2003; Franco et al., 2004; Besley and Kudamatsu, 2006; Tsai, 2006; Safaei, 2006; Ross, 2006). They are either restricted to only one non-income dimension of human development or to a cross-sectional analysis leaving out developments over time. Furthermore, they do not sufficiently account for possible conditions influencing democracy's performance.

Acknowledging these shortcomings, **Essay 1**, which is based on joint work with Sebastian Vollmer, extends the existing literature in several ways. First, the essay emphasizes the redistributive effects of democracy and complements Sen's theoretical argument using the well-known median voter theory to illustrate why democracy should outperform autocracy with respect to health and education (Meltzer and Richard, 1981). A second contribution

---

[1]http://hdrstats.undp.org/en/countries/country_fact_sheets/cty_fs_SGP.html (date of access, May 2010). Reference year is 2007.

[2]See http://hdrstats.undp.org/en/countries/country_fact_sheets/cty_fs_CUB.html (date of access, May 2010). Reference year is 2007.

[3]See http://hdrstats.undp.org/en/countries/country_fact_sheets/cty_fs_IND.html (date of access, May 2010). Reference year is 2007.

consists in identifying conditions (income inequality, the level of economic development, education and ethnic fractionalization) that are assumed to affect democracy's performance. Using a panel analysis over a time span of 30 years, the relationship between political institutions, life expectancy at birth and the literacy rate is tested and interaction effects are included to account for factors that affect the functioning of democracy.

The main finding is a robust positive and significant association between democracy and the indicators of human development, even if one controls for factors like the level of economic development. Although causality is difficult to establish, besides its intrinsic value, democracy seems to be instrumental to achieving better health and education. However, the interaction effects between democracy and the presumed conditions of functioning turn out to be insignificant or not robust.

The Role of Social Institutions related to Gender Inequality

As pointed out in North's definition there are different types of institutions that together determine the extent of capability expansion or deprivation. At the level below those institutions that are mainly concerned with property rights protection, redistribution and contract enforcement in a political system, there are informal social institutions (Williamson, 2000). These are often taken for granted, shape people's identity and provide role models that help people to behave appropriately in daily life without putting efficiency at the forefront (Hall and Taylor, 1996; Peters, 2005). Some of these institutions lead to capability deprivation in the form of social exclusion. Sen's capability approach has been criticized for his view of independence, autonomy and individualism, which fails to highlight social relations (e.g. Nussbaum, 2003; Gaspers and van Staveren, 2003). However, he identifies social exclusion as a constitutive part of capability deprivation as well as a cause of capability deprivation in other dimensions (Sen, 2000b).

The implantation of the "right" formal institutions, e.g. democratic ones, to a country does not guarantee the "right" track towards development, as formal institutions interact with informal ones. Development outcomes then depend on the strength of both formal and informal institutions (Williamson, 2009). Relationships are either complementary or substitutive. Although a formal democratic system may open the space for public discussion, deliberation might be at risk because a deeply rooted power structure and elite domination hinder the participation of all citizens (Gaspers and van Staveren, 2003). The relevance of this issue becomes obvious if social exclusion mechanisms in formally democratic countries are considered.[4] For example, informal institutions that back up social exclusion mechanisms might

---

[4]Gaspers and van Staveren (2003) and Nussbaum (2003) criticize Sen's account because his notion of social justice is underelaborated as it is left to social choice. Moreover, he does not explicitly deal with the problem

hinder the extension of the franchise. Racial discrimination against African Americans in the form of 'informal' violence and intimidation or disenfranchising laws restricted the use of the formal right to vote for black people for a long time until finally in 1965 the Voting Rights Act was passed to counteract at least to some extent those discriminatory practices.[5] Another example is Switzerland, where women gained the right to vote in 1971. Including social institutions in the study of development could therefore be a valuable effort.

This is particularly evident if one considers that despite considerable progress in recent decades, gender inequality in the manifold dimensions of well-being remains pervasive in many countries of the world. **Essays 2, 3 and 4** are dedicated to the roots of these inequalities and their heterogeneity across space and time. They center on social institutions related to gender inequality that frame gender-relevant meanings, shape gender roles and become guiding principles in everyday life. Influencing the distribution of power between men and women in the private sphere of the family, in the economic sphere, and in public life, they constrain the opportunities of women and their ability to become agents of development (Sen, 1999b).

Essay 2, which is the result of joint work with Boris Branisa and Stephan Klasen, focuses on the measurement of social institutions related to gender inequality. Existing measures are outcome-focused, measuring gender inequality in well-being and agency (Klasen, 2006, 2007), e.g. the Gender-Related Development Index (GDI) and the Gender Empowerment Measure (GEM) (United Nations Development Programme, 1995) or the Global Gender Gap Index from the World Economic Forum (Lopez-Claros and Zahidi, 2005). Other measures like the Women's Social Rights index (WOSOC) of the CIRI Human Rights Data Project[6] could be partially used as a proxy for the institutional basis of gender inequality. However, it also covers outcomes of institutions and, coming from a human rights perspective, it neglects informal institutions and does not differentiate between what happens within the family and what happens in public and social life.

Given this lack of measures, **Essay 2** proposes several composite indices measuring social institutions related to gender inequality that can be used to compare the societal situation of women in over 100 non-OECD countries and allow the identification of problematic countries and dimensions of social institutions that deserve attention by policy makers and need to be scrutinized in detail. These are the Social Institutions and Gender Index (SIGI) as a multidimensional measure of deprivation of women, and its five subindices each measuring

---

that freedoms have to be curtailed if social justice and, implicitly, equality are pursued. Nussbaum (2003) therefore claims fundamental entitlements that are independent of people's preferences.

[5]See http://www.justice.gov/crt/voting/intro/intro_b.php, date of access May 2010.

[6]Information is available on the webpage of the project http://ciri.binghamton.edu/ (date of access: April 16, 2010).

one dimension of social institutions related to gender inequality (Family code, Civil liberties, Physical integrity, Son preference and Ownership rights). The one-dimensional subindices are built out of variables of the OECD Gender, Institutions and Development database[7] using the method of polychoric PCA to extract the common information of the variables corresponding to a subindex (Kolenikov and Angeles, 2009). The formula of the SIGI is inspired by the Foster-Greer-Thorbecke poverty measures (Foster et al., 1984), which offers a reasonable way to capture the multidimensional deprivation of women caused by social institutions. It has the advantage of penalizing high inequality in each dimension and of allowing for only partial compensation between dimensions.

It is widely accepted that gender inequalities not only harm the affected women but come at a cost for the whole society, leading to ill-health, low human capital, bad governance and lower economic growth (e.g. World Bank, 2001; Klasen, 2002). Due to the scarcity of cross-country level data only a few studies investigate the development impact of gender-relevant social institutions (e.g. Morrisson and Jütting, 2005; Jütting et al., 2008). Applying the newly proposed social institutions indicators, **Essay 3**, which is based on joint work with Boris Branisa and Stephan Klasen, investigates at the cross-country level their explanatory value for development outcomes (female secondary schooling, fertility rates, child mortality and governance in the form of rule of law and voice and accountability). Based on bargaining household models (e.g. Manser and Brown, 1980; McElroy and Horney, 1981; Lundberg and Pollak, 1993), models considering the costs and returns of children (e.g. Becker, 1981; King and Hill, 1993; Hill and King, 1995) as well as contributions from several disciplines on governance and democracy, we derive hypotheses on the impact of social institutions related to gender inequality. The findings from the regression analysis show that social institutions matter even if one controls for religion, political system, geography and the level of economic development; higher inequality in social institutions is associated with worse development outcomes not only for the affected women but also the whole society.

Essay 4, which was produced in collaboration with Boris Branisa, elaborates more on the linkage between social institutions related to gender inequality and governance, contributing to a separate branch of research on gender and corruption. Former research efforts showed that there is a negative statistical association between representation of women in political and economic life and corruption in a society (Swamy et al., 2001; Dollar et al., 2001). Some explanations trace this back to differences in behavior between men and women, some take a historical perspective stating that women are newcomers to the system and therefore behave less corruptly than men (Goetz, 2007) and others mention the possible omitted variable

---

[7]See Morrisson and Jütting (2005); Jütting et al. (2008)

of liberal democracy which might affect both the level of representation and corruption in a society (Sung, 2003). Swamy et al. (2001) proposed another omitted variable, "the level of discrimination against women", which we try to capture using the subindex Civil liberties. The findings of a cross-sectional regression analysis controlling for democracy and representation of women in politics and economic life suggest that corruption is higher in countries where social institutions deprive women of their freedom to participate in social and economic life. In such contexts it might therefore not be sufficient to push democratic reforms and to increase the participation of women in order to reduce corruption.

Indigenous Origin and Health Inequality in Bolivia

Recognizing the pervasiveness of gender inequality in the world, MDG 3 is dedicated exclusively to the situation of women. With respect to other groups, the MDGs are less clear. However, background documents and global initiatives draw attention to indigenous people as they are overrepresented among the world's poor at about 15% and suffer more from marginalization, poverty and problems in health and education than the non-indigenous population (Hall et al., 2006; Stephens et al., 2006; United Nations, 2010). As a response to these problems the General Assembly of the United Nations proclaimed the Second International Decade of the World's Indigenous People, which started in 2005.

Essay 5, a result of a joint project with Elena Gross, focuses on the situation of indigenous people in Bolivia and demonstrates that ethnic origin is a decisive factor for child health and reaching MDG 4, "Reduce child mortality". From a first point of view, it seems that this is a settled fact and a further study seems to be unnecessary. However, most of the studies stating differences between indigenous and non-indigenous people are based on descriptive and bivariate evidence (e.g. UDAPE and OPS, 2004; Pozo et al., 2006; PAHO, 2007). Although social exclusion and institutional mechanisms are relevant factors for racial differences in well-being, this view might not be sufficient to design policy interventions. It falls short of considering other factors which might be related to both ethnic origin and health, like poverty, urban-rural differences, geographical location and other household related characteristics - linkages that can be observed for Bolivia. Even if multivariate analyses are conducted, there are shortcomings (e.g. Larrea and Freire, 2002; Morales et al., 2004; Mayer-Foulkes and Larrea, 2005). The first is to neglect the heterogeneity of health inequality over different health outcomes. The second is related to the usage of the indigenous dummy, which masks heterogeneity within the group of native people - if one bears in mind that there are over 30 distinct indigenous groups living in Bolivia. Our study investigates several indicators on childhood diseases and vaccinations, taking the former shortcomings into consideration. The main lesson is that ethnic origin matters. However, one should go beyond indigenous origin,

Quechua, Aymara, etc. and look for factors that capture particularly characteristics of the mother like health knowledge of the mother or mother's education that might be related to the heterogeneity in health outcomes over different ethnic groups. A hypothesis, which arises from Essay 5 and would need further investigation is that these characteristics of the mother might be intermittent variables between ethnic origin and health outcomes. However, this should be investigated additionally to putting efforts into analyzing institutional mechanisms that might lead to deprivation of these groups.

To summarize: the five essays of this dissertation contribute to the understanding of the linkages between institutions, inequality and development and emphasize the role of group-based disparities related to gender and ethnicity within this triangle. They confirm the fact that institutions matter and that they influence not only the level of development but also inequality in development outcomes. The essays also show that talking about institutions in general is less useful if policy implications should be drawn. Instead one should distinguish between political and social institutions and differentiate within these types of institutions. Moreover, this dissertation contributes to a discussion about the mechanisms that relate different types of institutions with development outcomes and it highlights factors, which might influence the functioning of these mechanisms by interacting with institutions in the production of development outcomes or which might be intermittent. Concerning democracy no robust pattern about interacting factors in the production of development outcomes has been found. With regards to social institutions a first step towards identifying possible mechanisms is taken and relationships are investigated. Furthermore, learning processes or policies, which change incentive structures are considered as possibilities to change these institutions. Concerning differences in health outcomes across ethnic groups in Bolivia it can be argued that these differences are due to latent institutions that distribute power across ethnic lines. However, it is shown that variables like mother's education or health knowledge let the effect of ethnic origin vanish and further investigations could focus on their role as intermittent variables having the potential to counteract the effect of institutions.

Chapter 1

Political Institutions and Human Development:
Does Democracy Fulfill its 'Constructive' and 'Instrumental' Role?[1]

1.1   Introduction

Since Sen (1988, 1991, 1999b,a, 2003), we have been aware of the fact that development is a very encompassing and broad concept. Development can be seen as enhancing each individual's capabilities, which define the freedoms to choose the kind of life they value in accordance with individual preferences. This approach has inspired the emergence of a pluralist and integrative conception of 'human development' and its operationalization in the form of UNDP's Human Development Index. It is not only income, but also health and education and other dimensions that enable people to shape their lives in line with their desires. The aim of this paper is to discuss the contribution political institutions might make to enhancing non-income human development measured in terms of education or health. We choose education and health as both aspects are direct determinants of capabilities and both influence the freedom to choose the kind of life one wants. Education as well as health raises productivity and the ability to convert income and resources into the favored way of life (Sen, 2003). The third dimension of human development, namely income, is not of interest for this paper, since a detailed literature on the relation between democracy and economic development is already available.

Political institutions are a critical area of research as they organize social, economic and political life. Hence, an obvious question is what kinds of institutions do this job best. From

---

[1]joint work with Sebastian Vollmer

a perspective of freedom, democracy has the advantage that its beneficiaries are free to take decisions about their lives and play a part in shaping societal decisions. Therefore, democracy is also considered as an end of the development process and a piece of the puzzle of the more comprehensive picture of human development (Sen, 1999b,a, 2000a). But whether democracy indeed has a positive impact on economic and human development is not a trivial question - either from a theoretical or from an empirical perspective. With regard to theory, three major debates are centered on the instrumental value of democracy for economic development:

First, there seems to be a controversy concerning the contradictory effects of growth-enhancing property-rights protection and equalizing, market-correcting redistribution in democratic societies on growth and well-being (e.g. Mohtadi and Roe, 2003; Tavares and Wacziarg, 2001; Alesina and Rodrik, 1994; Baum and Lake, 2003). Positions that emphasize the deficiencies of democratic systems may support the Lee Hypothesis, named after the former President of Singapore, Lee Kuan Yew, which states that autocratic regimes are more efficient systems to tackle market failures, to stimulate economic growth and as a consequence to improve human development (Sen, 1999a).

A second debate revolves around causation: is democracy a cause or a consequence of the development process? Taking a historical perspective, this debate was initiated by Lipset (1959) who emphasized the modernization process including progress in education, industrialization and urban development as driving forces of democracy. Other examples of these enhancing or impeding forces are income inequality and country-specific and historical characteristics. Several authors have dedicated their work to identifying these factors and/or to filtering out the effect of democracy or democratization on development by controlling for these factors (Barro, 1999; Bourguignon and Verdier, 2000; Persson and Tabellini, 2007b; Glaeser et al., 2007; Papaioannou and Siourounis, 2008; Acemoglu et al., 2008).

Third, in addition to the historical perspective, one could also take a more contemporaneous view as there might also be factors that shape the functioning of a democratic system. It is still not obvious what the conditions are under which democracies function well and for sure there is an overlap with the factors that make democratization work. For example, Keefer and Khemani (2005) and Besley and Burgess (2002) highlight information of voters and social fragmentation, Collier (2001), Mauro (1995), Alesina et al. (1999), Miguel and Gugerty (2005) and others draw attention to ethnic fractionalization, which could disturb the provision of public goods and foster corruption, others like Keefer (2005) focus on the age of democracy.

Empirical research studies give no clear answer concerning the effect of democracy on growth. Minier (1998) and Papaioannou and Siourounis (2008) show that the efficiency ar-

gument in favor of autocratic regimes does not withstand empirical investigations. They find a positive effect of democracy on economic growth. Others, on the contrary, find a moderately negative (Tavares and Wacziarg, 2001), nonlinear or heterogeneous relationship between democracy and growth, assumed to be due to, e.g., the maturity of a democratic system, rent-seeking, or the details of democratic reforms (Barro, 1996; Persson and Tabellini, 2006, 2007b).

When studies focus on the effect of democracy on redistribution, operationalized as the provision of public goods, the size of the public sector and income inequality, the results are less ambiguous (Alesina and Rodrik, 1994; Boix, 2001; Lake and Baum, 2001; Besley and Burgess, 2002; Gradstein and Milanovic, 2004; Stasavage, 2005; Persson et al., 2000). In general, they support the view that redistribution might be higher under a democratic regime, without clearly answering the question whether this redistribution is beneficial to economic and non-income human development.

Concerning the non-income dimensions of human development, there is again uncertainty about the effects of democracy. There are only a few studies empirically investigating the links between political systems and measures for the non-income dimensions of human development. Whereas some find a positive relationship between democracy and human development measured in terms of health and education (Lake and Baum, 2001; Baum and Lake, 2003; Navia and Zweifel, 2003; Franco et al., 2004; Besley and Kudamatsu, 2006; Tsai, 2006; Safaei, 2006), others find less evidence for this influence (Ross, 2006). These research efforts are either confined to only one of the non-income dimensions of human development (Navia and Zweifel, 2003; Franco et al., 2004; Besley and Kudamatsu, 2006; Ross, 2006; Safaei, 2006) or to a cross-sectional focus leaving out developments over time (Franco et al., 2004; Tsai, 2006). Moreover, these investigations, while having in mind potential conditions influencing democracy's performance, include these requisites only as simple controls in their regression models and not in interaction with some institutional measure. Exceptions are for example the studies of Boix (2001) and Baum and Lake (2003) who build interaction terms between democracy and different levels of GDP or the Gini index to capture the distinct effects of democracy in countries with different levels of income and inequality.

Acknowledging the shortcomings of the literature, in this paper we want to extend the latter strand of research in the following ways: we want to answer the questions of whether political institutions are related to the living standard of the population and whether our empirical data support the view that democracy, besides its intrinsic importance for the development process, fulfills a constructive and instrumental role by giving people the opportunity to express, to form and aggregate their preferences and thus to steer public action in an efficient and effective manner (Sen, 1999b). To provide an answer we complement the arguments pro-

vided by Sen with theoretical implications of the median voter theory, which is an innovative
way to think about the quality and quantity of redistribution and public service provision in
political regimes. A second contribution to the literature consists in theoretically identifying
and empirically testing conditions under which democracies will display a positive effect -
given they are supposed to have one - on the provision of public goods and services that are
assumed to foster human development. Consequently, we are not interested in explaining
democratization but in investigating the potential dependence of democracy's performance
upon other factors once it is in place. We empirically test the relationship between politi-
cal institutions and the levels of education and health, using these indicators as proxies for
non-income human development. The empirical investigation is based on a panel data set in-
cluding all countries for which information is available, which allows us to consider the time
dimension in our analysis. A last contribution consists in empirically estimating interaction
effects between the conditions of democracy's performance and a democracy variable.

In section 2, we review theories of political institutions, democracy and human develop-
ment. In section 3, we examine the empirical evidence for this relationship. In section 4,
we conclude. Our results indicate that democracy is favorable for human development even
after controlling for the level of economic development. But contrary to the theoretical rea-
soning, there is no clear evidence for the factors that according to the literature are supposed
to influence democracy's performance. It seems to be democracy itself - rather independent
from the circumstances - that has a positive effect on human development. It is in particular
remarkable that democracy's performance seems not to depend on a certain level of economic
development.

1.2   The Political Economy of Democracy and Human Development

1.2.1   How Can Political Institutions Influence Human Development?

With regard to a definition and the resulting operationalization of institutions, the existing
literature leaves the impression that there is not enough precision concerning the term "insti-
tution" itself. There is a heavy use of performance indicators measuring the extent to which
certain institutional systems function, e.g. when it comes to political stability or governance
issues (Gradstein and Milanovic, 2004).[2] Such performance indicators are then often mixed
up with public policies. However, both measured performance and policies are the output of
underlying structures and procedures as well as contextual factors. These underlying (for-

---

[2]See for example the Worldwide Governance Indicators of Kaufmann et al. (2007).

mal) structures and procedures can be subsumed under the heading "political system". This is what we understand by political institutions.

According to the rational choice strand of the new institutionalism in political science or the field of new institutional economics and political economy, political institutions shape the rules which govern the political game (Hall and Taylor, 1996; Persson and Tabellini, 2000; Peters, 2005). They do not only determine, via electoral rules, the actors and preferences which can access the political arena and get heard. They also provide the means to aggregate those preferences by establishing procedures for decision-making and distributing political power (Persson, 2002). The common output of institutions and preferences is policies. Although actors and other environmental constellations may change over time, policies in general will reflect the political institutions that produced them (Peters, 2005; Persson and Tabellini, 2006). Two types of policies may be favorable to human development: policies for the protection of property rights and policies for redistribution.

Policies for the protection of property rights contribute to economic development and economic growth (Acemoglu et al., 2002). Growth increases the welfare of the population by reducing poverty at least in the longer term (Dollar and Kraay, 2002; Klasen, 2004; Kraay, 2006). Therefore, property-rights protection is a necessary condition for an increase in the overall wealth of a nation (Acemoglu et al., 2001, 2002). But whether all members of this nation can benefit from it highly depends on redistribution as well. Policies for redistribution equalize the distribution of income and welfare in a society. A trade-off between the two types of policies might occur as on the one hand property-rights protection enhances development by securing investments into physical capital but is not concerned about distributional aspects of the costs and benefits. On the other hand, redistribution fosters human capital and lowers income inequality, but might hinder investments into physical capital and disturb incentives on the labor market and moreover, it might lead to rent-seeking activities (e.g. Mohtadi and Roe, 2003; Tavares and Wacziarg, 2001; Alesina and Rodrik, 1994; Baum and Lake, 2003).

Despite the potential negative effects of policies of redistribution, we consider them as essential to achieve progress in non-income indicators of human development like health and education. This type of policy comprises broad-based programs and covers the provision of public goods and services. These policies aim at compensating for market failures and at achieving normative, social optima. Especially the poor are given access to goods and services which are not sufficiently provided by markets. The matching of society's and an individual's needs with an adequate redistribution scheme and an appropriate public provision of goods and services provides a more direct link between political institutions and human development than property-rights protection. Following this line of argumentation the following question arises: what is a political system that is appropriate to produce market-correcting

redistributive policies that are designed to match the needs of society and have the potential to advance non-income human development? The answer is democracy.[3] Democracy is conceived as a political system whose structures and procedures permit the rule of the people. Of importance are free and repeated elections and political competition, the rule of law, and political and civil liberties. These component parts frame public debate and deliberation that deal with the management of society.

Although redistribution from the rich to the poor and vice versa exists in both autocratic and democratic systems, the following theoretical arguments suggest that redistribution from the rich to the poor is more pronounced in democracies.[4] One of the best-known theoretical arguments is the model of Meltzer and Richard (1981). The median voter hypothesis states that in democratic governments the median voter is the decisive voter. The more her income falls short of the average income of all voters, the higher the tax rate, i.e. redistribution, she will decide. Therefore, government spending should be larger and social services more extensive in democratic regimes - if the majority of the voting public lives at the bottom of the income distribution and only a small part enjoy richness (Keefer and Khemani, 2005). In contrast, in authoritarian systems, the distribution of wealth does not play a decisive role. All or a substantial part of the electorate is excluded from the decision-making process, and this is precisely to avoid the redistributive consequences of democracy. As a result, the average size of the public sector and public spending remains quite small (Boix, 2001).

Another line of argumentation brought forward by Lake and Baum (2001) emphasizes the state's monopoly to use force legitimately in producing public services that mitigate market failures. In a democratic regime, these services are provided in larger quantity and at lower prices, as barriers for political competitors and costs for political participation are low compared to autocracies. In autocracies the emphasis will be more on earning rents than on providing public services, assuming that earning rents is a function of the provision of public services and restricting the supply of services will increase rents. This argumentation also supports the hypothesis that democracies will provide higher levels of public services to their citizens.

However, quantity does not imply quality. In other words, voting alone does not solve the aggregation problem resulting from different individual preferences. Thus, a second question related to the qualitative dimension of redistribution emerges: why are democratic governments more responsive to the needs of the citizenry compared to autocratic ones? According

---

[3]Democracies are considered to perform best on both dimensions: property-rights protection and redistribution. Whether the one or the other is more important depends on people's preferences and the formal and informal face of the considered democracy.

[4]See for example Gradstein and Milanovic (2004) for an empirical study finding evidence for this linkage.

to (Sen, 1999b,a), democracy - beyond its "intrinsic" value - is of eminent importance for the process of development because of the "constructive" and "instrumental" role it plays in the formation and aggregation of values, needs and preferences and their translation into well-designed policies benefiting the society. Being constituent features of a democratic system, political and civil liberties, for example those related to free speech, public debate and criticism, permit the formation of preferences and values as well as access to the relevant information so that societal needs are visible. Democratic procedures facilitate the transmission of these needs into the political arena where decision-making power is distributed amongst legitimate representatives of the society as a whole.[5]

Democracy does not only help to construct policies that are matched to the needs of its citizens, but is also instrumental and protective. Control mechanisms such as free and repeated competitive elections and the compliance with the rule of law principle reduce discretionary and corrupt behavior of representatives who hold political power. Thus, democracy provides the incentives to create responsibility and accountability that induce political-administrative leaders to listen and to act on behalf of the society they represent (Sen, 1999b,a).

In an autocratic regime a usually small ruling elite dictates "the will of the people" from above. This is frequently accompanied by a repression of the political opposition and the prohibition of free expression and opinion, thereby impeding the conceptualization of the volonté générale. The state apparatus is (mis-)used in favor of the welfare of the ruling elite. Political measures with a redistributing character that increase the welfare of the bottom quantiles of society are implemented not because of institutional structures but for ideological reasons and only to a level that will help autocrats to remain in power and to increase their own wealth (Olson, 1993; McGuire and Olson, 1996). Responsiveness, representation, accountability and the selection of competent political and administrative staff thus are uncommon in autocratic regimes (Besley and Kudamatsu, 2006).

Summarizing, democracies quantitatively and qualitatively outperform autocracies with respect to redistribution. There is no clear relation between inequality and societal needs on the one hand and redistribution on the other hand in autocracies, except for those, generally socialist ones, with a special commitment to universal welfare. In general, this leads to a lower level of human development in autocratic systems.

---

[5]The latter means that otherwise disadvantaged groups, whether they are minorities or a broad mass of poor people in a developing country, get a voice and the opportunity to be heard and represented. In cases of direct democracy or democracy at a local level, these groups even decide for themselves.

1.2.2   What Determines Public Service Provision in Democracies?

The formal existence of democracy does not guarantee that it functions in the idealized manner described above. Democratic regimes display a lot of heterogeneity regarding human development outcomes. This is due to factors that determine whether the relationships predicted by the median voter theory or Sen's theory work or not. These factors then hamper or foster the performance of democracy with regard to the satisfaction of societal needs. Problems could arise if for certain reasons - located either at the agenda setting, the policy formulation, the implementation or the evaluation phase - the allocation of public expenditures is inefficient.[6]

Our approach to explain heterogeneity in any democracy's performance follows that of Keefer and Khemani (2005) and hence differs from other studies that focus more on the preconditions for democracy and democratization (e.g. Lipset, 1959; Glaeser et al., 2007). We do not consider the question whether a country has to be prepared for democracy or whether it is democracy that lifts the country up to a certain level of development.[7] Following our theoretical reasoning, the necessary timing of the presence of the respective factors is treated here as simultaneous. Their interaction with democracy at one point in time influences the output, the policies in the form of public goods provision, and the outcome, the level of human development.

First, as redistribution and the provision of public goods depend on whether there is anything to redistribute and to invest in public goods, the performance of a democratic system will be better the higher the level of economic development is (Boix, 2001; Baum and Lake, 2003).[8] So the positive effect of democracies on public goods provision will be intensified by the level of economic development.

Secondly, if citizens are ill-informed, this may lead to insufficient participation, which would be necessary for public reasoning and the expression of 'qualified' needs. As a result, the quality of responsive government manifesting itself in policies that reflect society's demands and needs decreases. Moreover, accountability suffers from information constraints because voters cannot control politicians' behavior. Education is one of the important factors as it has a potential to alleviate information problems.[9] Education in this context is not taken

---

[6]Because poor people are highly dependent on public action as they cannot invest their own (nonexistent) private resources, they suffer the most from ineffective government in terms of redistribution and service provision (Keefer and Khemani, 2005).

[7]Hence, we follow the statement of Sen (1999a, p. 4): "A country does not have to be deemed fit for democracy; rather, it has to become fit through democracy."

[8]On the effect of income on health there is a literature concerned with the absolute income hypothesis that states that income affects individual health but at a diminishing rate (e.g. Karlsson et al., 2010).

[9]Other factors might be a well developed media sector and accountable and institutionalized parties that take over political education tasks (see Keefer and Khemani, 2005). But it can easily be argued that without a certain

as an intrinsic component of human development that we want to explain, but as a means to human development. It is not only in itself a precondition for a higher living standard because it positively affects earnings, health and so on. It is also found to be a requirement for democracies to develop and to persist as it leads to conscientious participation that may be related to an efficient and effective provision of public goods (Lipset, 1959; Glaeser et al., 2007; Keefer and Khemani, 2005).[10]

Social fragmentation can be another factor disturbing the functioning of a democratic system measured by the public goods it provides. Research has found that social fragmentation or, to be more concise, ethnic diversity leads to collective action problems, increased patronage as well as clientelism and in the end to an under-provision of public goods (Alesina et al., 1999; Alesina and Ferrara, 2005; Miguel and Gugerty, 2005). Within democratic systems, social fragmentation may pose problems because mechanisms which would hold the government accountable and responsible are undermined. In socially heterogeneous settings, governments are rewarded on the basis of identity and not on their performance (Keefer and Khemani, 2005). Moreover, social fragmentation leads to political fragmentation, which from a certain threshold value can result in increasing co-operation problems (Collier, 2001).

The last factor that is in line with the quantity-redistribution argument is income inequality, characterized by a distribution of income where the median income is smaller than the average income.[11] The following argumentation supports the idea that in a high income-inequality context democratic systems might provide more health and education services due to stronger redistributive pressures by the median voter. As a starting point it is necessary to understand how income inequality affects health and education. Income inequality reduces human development because, in more unequal societies, fewer people can afford to live a healthy life and to spend their money on education. Moreover, income inequality may lead to stress and frustration harming health, and according to e.g. Wilkinson (1992), Kawachi et al.

---

level of broad-based education, a media sector will not develop because of a lack of demand (for the role of the media see Besley and Burgess (2002). The same is supposed to hold for the institutionalization of parties and accountability issues.

[10]We leave out cultural factors as they are hard to measure. Inglehart and Welzel (2005) emphasize the people's values as being as important as socioeconomic resources and civil and political rights. According to these authors, culture provides the link between economic development and democratic freedom. Without certain values like "human autonomy" or "self-expression values", fostering the priority of self-made choices, human development might not be possible. Such values are dependent upon a certain level of socioeconomic development that we might proxy by taking the level of economic development into account. Moreover, we assume - although this is to be questioned - that the more education people have the more enlightened they are and the more freedom they demand to live the life they value.

[11]The argument that the median voter is farther away from the mean when a society is more unequal is true for right-skewed distributions. This is usually the case for national income distributions, which are quite close to log-normal distributions.

(1997) and Karlsson et al. (2010) it is found that income inequality leads to a higher mortal-
ity as social cohesion breaks down. Income inequality leads to a residential concentration of
the poor and the rich that gives rise to a segregation hindering social cohesion. Poor people
living in a poor neighborhood not only have to get along with a lack of income but also with a
worse infrastructure related to e.g. schooling or health so that their situation cannot improve,
whereas the rich invest in their neighborhood, in particular in human capital, health care and
other factors. As the share of poor people rises and segregation aggravates, the levels of
health and education in a society become worse. In segregated and polarized societies the
provision of public goods worsens. Moreover, income inequality spurs crime and violence,
affecting health directly.[12] If redistributive pressures increase, meaning that the distance of
the median voter's income from the average income becomes larger, then in democratic sys-
tems according to the median voter theory more redistribution will be demanded (Meltzer
and Richard, 1981). Whether the redistribution is in the form of education and health ser-
vices or income transfers is open, as high income inequality does not necessarily imply high
inequality in education and health (Grimm et al., 2008). In such a context the median voter
could be healthier than the average and more literate as well. Therefore, it is not obvious
why the median voter should demand more health or education services. However, redistri-
bution of any kind is expected to compensate for the negative effects of income inequality.[13]
Autocratic regimes lack such a mechanism. Moreover, democratic regimes foster the rise of
civil society organizations that preserve social cohesion and capital and take over tasks that
are insufficiently fulfilled by the state (Safaei, 2006).

1.2.3 Summary and Working Hypotheses

Summarizing the theoretical arguments above, we can state that democratic regimes in com-
parison to autocratic ones are expected to produce a higher rate of redistribution and thus lead
to higher public expenditures. Public spending priorities in democracies reflect the needs of
the society more than those in autocracies, and democratic control mechanisms will assure
the implementation of policies so that a high degree of compliance with laws, directives and
orders is reached. Hence, public action can translate into the desired human development out-
comes, for example a better health status of the population or a lower illiteracy rate. But the
performance of democracies will vary according to the specific circumstances. We assume
that the level of income, education, social fragmentation and the level of income inequality

---

[12]The mechanisms which reflect how income inequality might affect health are subsumed under the income
inequality hypothesis which states that income inequality in a society affects the health of every member of the
society (e.g. Karlsson et al., 2010).

[13]Boix (2001) states that political participation is an important condition for inequality to be translated into
redistributive pressures in a democratic regime.

all affect the level of the provision of public goods and human development in a democratic system. Therefore, the following general hypotheses can be derived:

- Democratic political systems will yield better results in human development than autocracies, independently of the level of economic development.

- The positive effect of democracies on public goods provision will be intensified by the level of economic development.

- The positive effect of democracy on human development will be higher, the higher the level of education in a society.

- Social fragmentation lowers the positive impact of democracies on human development. The more socially diverse a country is, the more difficult it is to provide broad-based services even in democracies.

- The redistribution effect of democracy compensates for the negative effect of income inequality on human development. Furthermore, the higher the level of inequality, the bigger the positive effect of democracy on human development.

1.3 Empirical Links Between Democracy and Human Development

1.3.1 Empirical Implementation

To quantify human development, we focus on the non-income components of UNDP's Human Development Index and consequently use UNDP's data on life expectancy at birth and on literacy rates. Life expectancy at birth is measured in years, whereas the literacy rate is an index value ranging from 0 to 100. We choose education and health as both aspects are direct determinants of capabilities and as they both influence the freedom to choose the kind of life one likes. Education as well as health raises productivity and the ability to convert income and resources into the favored way of life (Sen, 2003). The third dimension of human development, namely income, is not of interest for this paper, since a detailed literature on the relation between democracy and economic development is already available. Our data on political institutions is taken from the Polity IV Project of the Center for Systemic Peace at George Mason University (Marshall and Jaggers, 2005). We use the Polity2 score as our *Democracy* variable ranging from 10 (highly democratic) to -10 (highly autocratic), while a zero score indicates a state between autocracy and democracy which we consider as not

being democratic.[14] There are systems scoring around the zero point that yield traits of both autocratic and democratic systems and are therefore transitory regimes, but to facilitate the further examination we classify those regimes having a score above zero as democratic and the other ones as autocratic.

Following Besley and Kudamatsu (2006), we take the fraction of democratic years over the past five years as our measure for democracy (*Demexp*) to capture democratic experience. As an alternative measure for democracy we calculate the average Polity2 score over the past five years (*Mpol*) that allows to consider the quality of a democratic or autocratic system. While *Demexp* does not mask transitions from democracy to autocracy and vice versa like *Mpol*, it classifies all countries having a score of -10 or 0 as autocratic systems and all others as democratic hiding differences between them. Here, *Mpol* allows to differentiate within the groups of democratic and autocratic countries.

The consideration of a period of five years has the advantage to obtain a more stable value for the democracy measure used. Another reason for the five year period is that the values of life expectancy and literacy are not updated annually but roughly every five years. Nevertheless, one might argue that it is certainly arbitrary to take five years and not ten, but with this choice, we are in line with the existing literature (e.g. Besley and Kudamatsu, 2006) and our study is therefore comparable. Having different democracy measures is rather important as a check of robustness.

Other variables we expect to have an impact on human development or that describe possible conditions under which democracy affects human development are the following: *GDP* per capita PPP in constant prices[15] from the Penn World Tables 6.2; *Gini* coefficients from the WIDER dataset with improvements in terms of comparability across countries and time by Grün and Klasen (2008)[16]; a measure of ethnic fractionalization (*Fractional.*)[17] as

---

[14]According to the Polity2 measure, a system can be classified as democratic if three interdependent elements exist: 1) competitiveness of participation, institutions and procedures allow citizens to express their political preferences; 2) openness and competitiveness of executive recruitment and constraints on the chief executive, so that the executive power is institutionally constraint; 3) civil liberties. The last element as well as rule of law, system of checks and balances, freedom of the press etc. is not coded in the index as the latter are performance indicators of democratic regimes. Autocracies are defined vice versa. For more details see Marshall and Jaggers (2005: 13f.).

[15]US$, base year: 2000.

[16]Gini coefficients are not available for every year. We therefore use a simple moving average between available observations to complete the dataset. The reference category for the Gini coefficients is gross income per capita.

[17]The ethnic fractionalization measure renders the probability that two individuals selected at random from a population are members of different groups. It is calculated with data on language and origin using the following formula $FRAC_j = 1 - \sum_{i=1}^{N} s_{ij}^2$, where s_{ij} is the proportion of group $i = 1, \ldots, N$ in country j going from complete homogeneity (an index of 0) to complete heterogeneity (an index of 1). For more details see Alesina et al. (2003).

proxy for social fragmentation from Alesina et al. (2003) which is constant over time.[18] Since education is a factor assumed to influence the performance of democracy, literacy rates are also used as an explanatory variable in our panel analysis for life expectancy but are neglected in the analysis of literacy itself. As our additional control variables we consider as important whether a country experienced some conflict in the period under observation and whether a high percentage of population is suffering from HIV/AIDS. To measure war, we take data from the UCDP/PRIO intrastate conflict onset dataset, 1946-2006. We choose the variable warinci2 (*War*) that measures the incidence of intrastate war and is coded 1 in all country years with at least one active war.[19] For HIV/AIDS, we take adult (15-49 years) HIV prevalence rates (*Aids*) from the 2008 Report on the global AIDS epidemic from UNAIDS/WHO. Data coverage over time and countries lead us to the decision to create a variable that takes the value 1 when a country has a prevalence rate over 5 per cent in the year 2003. To take the heterogeneity between autocracies into account we introduce a simple *Socialism* dummy to represent autocracies with a commitment to universal welfare (Safaei, 2006). The dummy takes the value one for all Eastern European countries until 1990, Vietnam until 1980, China until 1975, and for Cuba and North Korea until today.

We suspect that democracy causes different priorities in public expenditures compared to autocracies. Therefore, increases in public expenditures on health and education can be decomposed into two components: an increase due to higher total expenditures and an increase due to different priorities in government spending. While the first source is mainly driven by economic growth, we expect democracy to be a main driver of the second source. We were unable to gather sound data for government spending for the given period. Such data would have enriched our analysis as we could have examined the channels that democracy takes to affect human development. The available data on public expenditures in health and education were not adequate for our analysis. Only for the more recent years does the Government Finance Statistics of the IMF include sufficient information concerning these issues. Thus, neither the public expenditures' path of causation nor the channel of private spending can be investigated here due to data restrictions. We must therefore rely on the theoretical argumentation that underpins our empirical analysis.

---

[18]According to Alesina et al. (2003) the assumption of stable group shares is not a problem as examples of changes in ethnic fractionalization are rare. At least over the time-horizon of 20 to 30 years, time persistence can be assumed.

[19]War is defined by more than 1000 battle deaths. As intrastate wars are more frequent than interstate wars, we decided to take the intrastate war variable.

1.3.2   Descriptive Statistics

In 1970, we have 44 democratic countries and 99 autocracies in our dataset. In 2000, these were 97 democracies and 58 autocracies. In the data on which we base our estimations and for the time span 1966 to 2000 that is used for calculating our proxy measures for democracy, there are 32 countries with a a polity2 score larger than 0 and 35 countries with a polity2 score smaller or equal to zero.[20] If the whole time span from 1966 to 2000 is considered and all countries are included that have no missing on the Polity2 score, in 66 cases there was a transition from a positive polity2 score to a zero or negative one and in 166 cases a transition from a negative or zero polity2 score to a positive one indicating a transition to democracy.

Average life expectancy was 57.39 years and an average of 62.77% of the adult population were literate in the year 1970. In the year 2000 life expectancy had increased to 64.75 years and literacy rates went up to 80.44%. In 1970, life expectancy in democratic countries was 60.6 years compared to 55 years in autocratic countries. Until 2000, the gap between democratic and autocratic countries widened as people in democracies had an average life expectancy of 67.85 and in autocracies only 58.65 years of age. Literacy rates give a similar picture with 73.3% literate persons in democratic countries compared to 54.48% in autocracies in 1970, and 85.12% literate in democracies in the year 2000 compared to 69.65% in autocratic systems.

Besides looking at simple averages it is worthwhile to take a look at the densities of life expectancy and literacy for democracies and autocracies separately (Figures 1.1 and 1.2). We use kernel density estimators for this purpose and apply boundary corrections at 0 and 100 for the literacy rate and at the minimum and maximum values for life expectancy. While in democracies, both for life expectancy and literacy the mass of the distribution tends to the right hand side, there seems to be a group of autocracies with a low level and another one with a high level of life expectancy and literacy each.

The same pattern can be observed in Tables 5.7 to 5.14 in Appendix 1 where we classified countries according to three categories: low, middle and high income; autocracy and

---

[20]The countries with a polity2 score larger than 0 are: Australia, Austria, Belgium, Botswana, Canada, Colombia, Costa Rica, Denmark, Finland, France, Germany, India, Ireland, Israel, Italy, Jamaica, Japan, Malaysia, Mauritius, Namibia, Netherlands, New Zealand, Norway, Papua New Guinea, South Africa, Sri Lanka, Sweden, Switzerland, Trinidad and Tobago, United Kingdom, United States, Venezuela. The countries with a polity2 score smaller and equal to 0 are: Afghanistan, Algeria, Bhutan, Burundi, Cameroon, Chad, China, Congo, Dem. Rep., Cuba, Egypt, Gabon, Guinea, Iraq, Jordan, Kenya, Korea, Dem. Rep., Lao PDR, Liberia, Libya, Mauritania, Morocco, Myanmar, Oman, Rwanda, Saudi Arabia, Singapore, Syrian Arab Republic, Togo, Tunisia, Yemen as well as Tajikistan, Turkmenistan, Uzbekistan, Kyrgyz Republic, and Kazakhstan taking the former status of the Russian Federation into account.

Figure 1.1: Cross-Country Distribution of Life expectancy at Birth
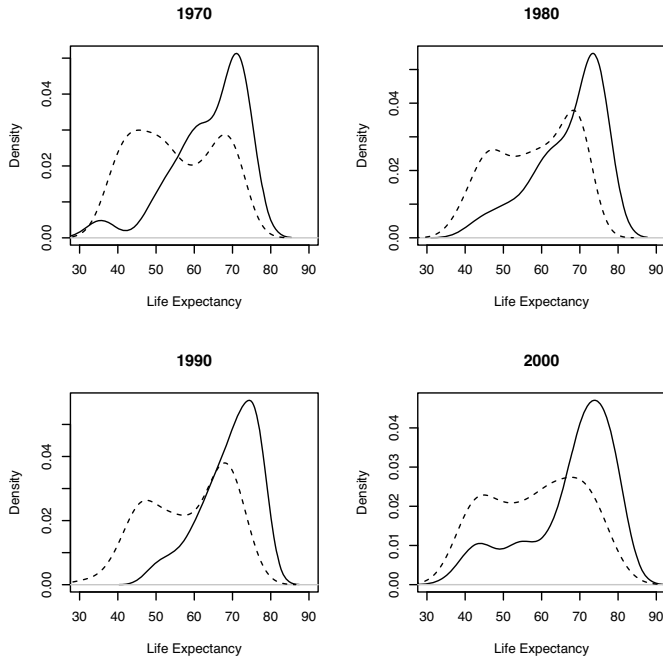
Solid line: Kernel density estimator for countries being democratic in the given year. Dashed line: Kernel density estimator for countries being autocratic in the given year. 1970: 41 democracies and 97 autocracies; 1980: 44 democracies and 107 autocracies; 1990: 67 democracies and 85 autocracies; 2000: 97 democracies and 58 autocracies.

democracy; low, middle and high life expectancy or literacy rates.[21] On average, we observe that democracies have a higher life expectancy and a higher literacy rate than autocracies. Exceptions are democracies with low life expectancies, mainly due to the HIV/AIDS tragedy in big parts of Sub-Saharan Africa. Considering the rich group of autocracies especially in 2000, it is striking that virtually all of them are oil states. This indicates, at least to some extent, that autocracies have problems catching up with the top of the income distribution, as

---

[21]To define the groups of low, middle and high life expectancy or literacy rates we computed quantiles of life expectancy and literacy. The income groups are defined according to Holzmann et al. (2008).

long as they do not control a large amount of such an important resource as oil. But what is more important for our study is the fact that although these countries show a high level of income, whether caused by natural resources or not, they display lower life expectancies and lower literacy rates than their democratic counterparts.

Figure 1.2: Cross-Country Distribution of Adult Literacy Rates
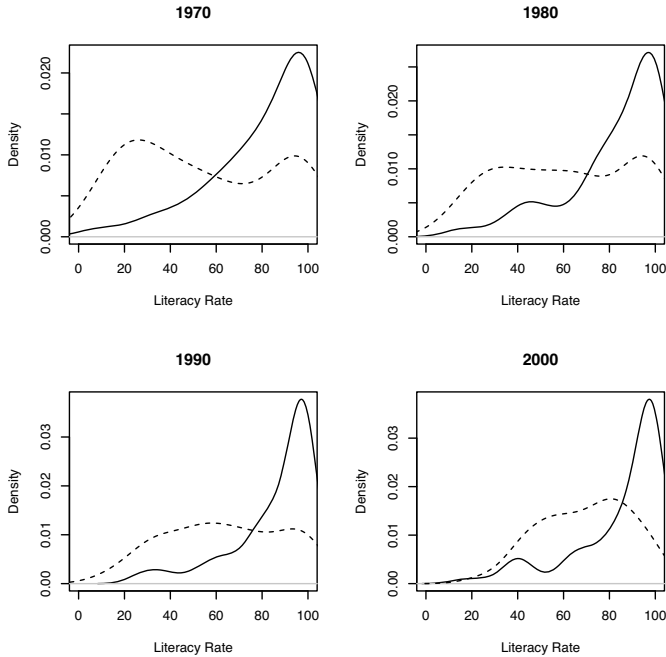
Solid line: Kernel density estimator for countries being democratic in the given year. Dashed line: Kernel density estimator for countries being autocratic in the given year. 1970: 23 democracies and 77 autocracies; 1980: 25 democracies and 87 autocracies; 1990: 44 democracies and 68 autocracies; 2000: 70 democracies and 45 autocracies.

1.3.3 Panel Analysis

The panel analysis aims at estimating the effect of democracy on life expectancy and literacy. As pre-estimation diagnostics indicate that heteroscedasticity and autocorrelation have to be dealt with, we run a cross-sectional time-series feasible generalized least squares regression

with panel specific AR(1), addressing both issues simultaneously.[22] We do the estimation without fixed effects because fixed effects generally capture institutional, political and socioeconomic country characteristics, which are usually quite time-invariant. This is reflected in Table 5.15 in Appendix 1 which shows very little variation of our democracy variables, particularly for Mpol, and also of the dependent variables life expectancy and literacy rate (i.e. little within variation) but more variation over countries (between variation). Utilization of fixed effects would disguise the impact of our democracy variables on life expectancy and the literacy rate. Moreover, one cannot assume that democracy shows effects rapidly. Democracy needs time and stability to perform well (Keefer, 2005), in particularly, with respect to social indicators like life expectancy and literacy that change only incrementally.[23]

In a simple model, we try to explain life expectancy and literacy with our measures of democracy controlling for GDP. GDP is lagged for one period to reduce the apparent problem of endogeneity. Additionally to the measures of democracy and economic development, we include the literacy rate as a proxy of the population's ability to articulate their needs in the political arena, to control politicians' activities and as a proxy of the population's priority for private spending on education and health. We also lag literacy for one period to reduce endogeneity problems. We only include education and its interaction with democracy in the model with life expectancy as our dependent variable. In line with our theoretical reasoning, we incorporate the lagged Gini coefficient to measure the effect of income inequality and ethnic fractionalization as a proxy for social fragmentation.

As pointed out, all variables describe conditions which potentially hamper or foster the functioning of democracy in terms of addressing the needs of the population. Thus, we are interested in their interaction with democracy on the one hand. On the other hand, we want to know whether they have an effect on human development independently from the political system.

Furthermore, we add a set of dummies for global regions[24] as well as year dummies to all regressions. The region dummies should capture much of the geographical, political and historical heterogeneity across the world. The inclusion of period effects allows us to capture overall upward trends in literacy and life expectancy that for example could be explained by technological improvements (Pritchett and Summers, 1996). Moreover, we control in both re-

---

[22]The Stata command xtgls is used. We assume that variance for each panel differs and that there is serial correlation where the correlation parameter is unique for each panel.

[23]Acemoglu et al. (2008) find a cross-country correlation between income and democracy only in the cross-section and attribute this to a long-term effect, i.e. positive changes in income and democracy over the past 500 years. According to this societies took divergent paths with respect to political and economic changes. This might be reflected here as well.

[24]Following the World Bank definition.

gressions for war, because it destroys lives as well as infrastructure for the provision of health and education services. Additionally, we control for HIV/AIDS in the life expectancy regressions. The AIDS dummy variable is interacted with the year dummies because HIV/AIDS was more of a problem for the more recent years in the sample and less in the earlier ones. A socialism dummy aims to capture heterogeneity across autocracies and an egalitarian tendency in those regimes.

We estimate the model for the years 1970, 1975, 1980, 1985, 1990, 1995 and 2000 (and the preceding five year periods), as both literacy rate and life expectancy are not updated annually but roughly every five years, while being interpolated in the other years. Taking observations of every fifth year is preferred to averaging the five-year data, as averaging introduces additional serial correlation that hinders inference and estimation (Acemoglu et al., 2008).

In case of life expectancy, we run separate regressions for non-OECD countries and the entire sample. For literacy, only the regression for the sub-sample of non-OECD countries makes sense as all OECD countries have an assumed constant level of literacy of exactly 99 percent in the UNDP data. The results are presented in Tables 1.1 to 1.3.

The results for the control variables are as expected in all specifications. The coefficients of the other main explanatory variables carry the expected signs and are highly significant, except for the Gini variable, which has an insignificant sign in most cases. The coefficient of the GDP per capita is positive, the literacy rate has a positive coefficient in the life expectancy regressions (remember that it is not included in the regressions where literacy is the dependent variable), and fractionalization carries a negative sign. All these results are robust to the choice of the democracy measure; they hold both for the fraction of democratic years (Demexp) and the average Polity2 score (Mpol). The coefficients of the year dummies are positive and highly significant for all years. The coefficients are continuously increasing over time and are thus capturing overall progress for human development due to for instance technology. The AIDS*time dummies are negative and highly significant for 1990, 1995 and 2000. This result displays the tragedy of HIV/AIDS and its immense impact on life expectancy in many African countries during this period. The coefficient of the War dummy is highly negative significant in the regressions with life expectancy as dependent variable and insignificant in the regressions with the literacy rate as dependent variable. The coefficient of the socialism dummy is positive and highly significant whenever included.

Table 1.1: Panel Analysis for All Countries (Dependent Variable: Life Expectancy at Birth)

	Demexp		Mpol	
Democracy	**1.238***	**1.112***	**0.089***	**0.101***
	(0.136)	**(0.169)**	**(0.010)**	**(0.011)**
log GDP(-1)	4.140***	3.982***	3.813***	3.803***
	(0.121)	(0.165)	(0.145)	(0.135)
Gini(-1)	0.878	1.185	1.155	1.065
	(0.719)	(0.775)	(0.716)	(0.597)
Literacy (-1)	0.199***	0.204***	0.205***	0.202***
	(0.007)	(0.008)	(0.008)	(0.008)
Fractional.	-2.100***	-2.474***	-1.919***	-2.407***
	(0.420)	(0.543)	(0.461)	(0.493)
War	-0.635***	-0.644***	-0.651***	-0.543***
	(0.163)	(0.165)	(0.158)	(0.154)
Socialism	2.236***	2.039***	2.197***	2.370***
	(0.600)	(0.580)	(0.562)	(0.552)
Aids*1975	0.24	0.192	0.605	0.559
	(0.599)	(0.618)	(0.575)	(0.617)
Aids*1980	0.42	0.362	0.763	0.759
	(0.601)	(0.623)	(0.580)	(0.625)
Aids*1985	0.064	-0.012	0.312	0.295
	(0.616)	(0.642)	(0.598)	(0.645)
Aids*1990	-2.362***	-2.421***	-2.261***	-2.180**
	(0.652)	(0.686)	(0.638)	(0.690)
Aids*1995	-8.044***	-8.128***	-8.116***	-8.067***
	(0.701)	(0.727)	(0.693)	(0.740)
Aids*2000	-15.595***	-15.681***	-15.749***	-15.566***
	(0.776)	(0.812)	(0.777)	(0.839)
Year 1980	0.838***	0.826***	0.822***	0.808***
	(0.066)	(0.062)	(0.065)	(0.068)
Year 1985	1.897***	1.845***	1.861***	1.809***
	(0.088)	(0.084)	(0.088)	(0.090)
Year 1990	2.374***	2.304***	2.355***	2.244***
	(0.103)	(0.099)	(0.103)	(0.104)
Year 1995	2.908***	2.823***	2.894***	2.715***
	(0.115)	(0.109)	(0.117)	(0.111)
Year 2000	3.332***	3.251***	3.316***	3.169***
	(0.127)	(0.123)	(0.131)	(0.119)
Dem.*GDP		-0.132		0.037*
		(0.265)		(0.017)
Dem.*Literacy		-0.008		-0.001*
		(0.009)		(0.001)
Dem.*Gini		2.251		0.270**
		(1.412)		(0.088)
Dem.*Fract.		-2.674***		-0.142**
		(0.762)		(0.050)
constant	60.442***	60.687***	60.758***	60.412***
	(0.200)	(0.228)	(0.224)	(0.220)
N	621	621	621	621

* p<0.05, ** p<0.01, *** p<0.001; dummies for global regions included and jointly significant

Table 1.2: Panel Analysis for Non-OECD Countries (Dependent Variable: Life Expectancy at Birth)

	Demexp		Mpol	
Democracy	**1.159***	**1.202***	**0.078***	**0.098***
	(0.166)	**(0.205)**	**(0.012)**	**(0.016)**
log GDP(-1)	3.334***	3.339***	3.304***	3.495***
	(0.214)	(0.227)	(0.206)	(0.194)
Gini(-1)	-0.785	0.098	-0.683	0.859
	(0.820)	(0.893)	(0.855)	(0.888)
Literacy (-1)	0.203***	0.198***	0.194***	0.190***
	(0.009)	(0.010)	(0.009)	(0.010)
Fractional.	-1.708*	-3.379***	-2.632***	-3.161***
	(0.745)	(0.725)	(0.785)	(0.801)
War	-0.797***	-0.786***	-0.903***	-0.779***
	(0.159)	(0.165)	(0.128)	(0.172)
Socialism	5.057***	4.985***	4.970***	5.229***
	(0.831)	(0.853)	(0.852)	(0.890)
Aids*1975	0.932	0.657	1.359*	1.07
	(0.664)	(0.675)	(0.635)	(0.668)
Aids*1980	0.828	0.563	1.315*	0.964
	(0.665)	(0.680)	(0.639)	(0.677)
Aids*1985	0.164	-0.076	0.613	0.215
	(0.677)	(0.695)	(0.656)	(0.697)
Aids*1990	-2.477***	-2.586***	-2.086***	-2.426**
	(0.709)	(0.732)	(0.692)	(0.739)
Aids*1995	-8.181***	-8.200***	-7.804***	-8.031***
	(0.760)	(0.776)	(0.748)	(0.788)
Aids*2000	-15.401***	-15.481***	-15.138***	-15.319***
	(0.830)	(0.854)	(0.826)	(0.885)
Year 1980	1.186***	1.200***	1.172***	1.214***
	(0.115)	(0.121)	(0.116)	(0.123)
Year 1985	2.476***	2.507***	2.509***	2.582***
	(0.158)	(0.164)	(0.159)	(0.170)
Year 1990	3.195***	3.163***	3.247***	3.252***
	(0.192)	(0.197)	(0.193)	(0.205)
Year 1995	3.606***	3.542***	3.648***	3.615***
	(0.224)	(0.229)	(0.226)	(0.240)
Year 2000	3.880***	3.849***	4.009***	3.965***
	(0.240)	(0.244)	(0.242)	(0.252)
Dem.*GDP		-0.329		0.003
		(0.288)		(0.020)
Dem.*Literacy		-0.002		-0.000
		(0.010)		(0.001)
Dem.*Gini		2.377		0.091
		(1.612)		(0.116)
Dem.*Fract.		-3.057**		-0.146
		(0.968)		(0.076)
constant	58.833***	59.052***	59.674***	59.194***
	(0.325)	(0.326)	(0.323)	(0.344)
N	469	469	469	469

* p<0.05, ** p<0.01, *** p<0.001; dummies for global regions included and jointly significant

Table 1.3: Panel Analysis for Non-OECD Countries (Dependent Variable: Adult Literacy Rate)

	Demexp			Mpol
Democracy	1.643***	1.169**	0.300***	0.055
	(0.374)	(0.414)	(0.018)	(0.036)
log GDP(-1)	11.685***	10.358***	11.499***	11.098***
	(0.451)	(0.494)	(0.471)	(0.498)
Gini(-1)	5.010**	-5.292**	6.463**	-1.452
	(1.570)	(1.828)	(2.379)	(1.970)
Fractional.	-8.857***	-8.748***	-3.516**	-10.025***
	(1.695)	(1.903)	(1.338)	(2.103)
Socialism	8.627***	6.947**	10.897***	7.046**
	(2.237)	(2.283)	(2.175)	(2.255)
War	0.221	-0.311	-0.581*	-0.06
	(0.350)	(0.351)	(0.232)	(0.369)
Year 1980	2.896***	2.917***	3.033***	2.843***
	(0.293)	(0.286)	(0.316)	(0.286)
Year 1985	6.613***	6.724***	6.611***	6.649***
	(0.393)	(0.385)	(0.422)	(0.385)
Year 1990	9.384***	9.618***	9.316***	9.437***
	(0.469)	(0.465)	(0.502)	(0.464)
Year 1995	11.841***	12.214***	11.031***	11.871***
	(0.550)	(0.541)	(0.580)	(0.540)
Year 2000	13.861***	14.185***	12.857***	13.890***
	(0.563)	(0.552)	(0.580)	(0.557)
Dem.*GDP		-0.935*		-0.074
		(0.466)		(0.038)
Dem.*Fract.		0.867		0.106
		(1.765)		(0.144)
Dem.*Gini		-10.606***		-0.451
		(2.846)		(0.232)
constant	81.373***	81.681***	81.667***	82.503***
	(0.699)	(0.589)	(0.615)	(0.773)
N	526	526	526	526

* p<0.05, ** p<0.01, *** p<0.001; dummies for global regions included and jointly significant

There is a strong positive and highly significant correlation between our measures of human development and democracy in nearly all specifications (we will discuss the one exception below). The fraction of democratic years (Demexp) and the institutional maturity of a system measured by the mean of the polity2 score (Mpol) both are positively related to life expectancy at birth.

When it comes to the interaction effects of democracy with GDP per capita, ethnic fractionalization, inequality and literacy respectively the results are rather ambiguous. The interaction of GDP and democracy sometimes carries a positive sign and sometimes a negative sign depending on the measure of democracy and the countries included in the sample. In fact, it is insignificant in most cases. We conclude that there is no robust evidence for this in-

teraction and thus the democracy's performance seems to not depend on the level of economic development. A similar argument holds true for the interaction of inequality and democracy. In the life expectancy regression, its coefficient is only positive and significant when Mpol is used as measure of democracy. In the literacy regression, the Gini interaction effect is only significant for one of the two democracy measures (Demexp) and thus not fully reliable. Contrary to the median voter prediction, it carries a negative sign in this case. The interaction of democracy and literacy is only significant for Mpol and not for Demexp. The interaction of democracy and ethnic fractionalization is significant in the life expectancy regressions for the full sample; it carries the expected negative sign. For the sample of non-OECD countries, it is only significant when Demexp is used as measure of democracy, both for literacy and life expectancy. Hence, there is more support for this interaction effect in the data than for the others, indicating that social fragmentation might disturb democracy's performance in a country.

Overall, there is only weak evidence for any of these interactions. The specifications excluding interaction effects are therefore the more reliable ones. This might also explain why there is no significant effect of democracy on literacy in the model including Mpol and all interaction effects. Summarizing, it can be said that a democracy's association with life expectancy and literacy is positive and robust but does not depend on the circumstances.

1.4   Conclusion

We believe that our study has its associated merits explaining the linkage between democracy and human development. In our theoretical section, we clarified the causal channels of democracy influencing human development. In contrast to earlier studies, which put their focus on property rights, we emphasized the importance of the effects of redistribution and of public goods provision in a democracy. The statistical association between democracy and human development is investigated descriptively and analytically. Extending existing literature, we not only measure the association between democracy and human development, but we theoretically and empirically analyze conditions that are assumed to be important for the functioning of democracy in terms of improving the level of human development.

Empirically, the results show a strong and robust correlation between democracy and human development measured by life expectancy at birth and the literacy rate, even if one controls for the level of economic development and other important variables. Besides, the effect is observed even if autocorrelation of the error terms is taken into account. Since the control of autocorrelation also remedies the omitted variable problem - if the correlation of omitted variables with the right hand side variables is low - we can be sure that the results are indeed robust. The results show that people living in democratic systems do better than

people in autocracies and, relying on the theoretical reasoning, a population's well-being is influenced by the political system. Both the stability of a democratic system as well as its institutional maturity are relevant.

However, the observed effect might be traced back more on the cross-sectional variation than the variation over time. This implies that causality is difficult to establish and one can be less certain about the effect of other social and political factors, which are very well proxied by democracy and that do not change over time. Future studies should incorporate social capital as well as the degree of decentralization of the political-administrative system. Conducting a historical examination that begins at the time when democratic systems (in a modern sense) evolve would give more reliable results. In addition, it would certainly be an improvement of our analysis to empirically identify and model the channels that democracy takes before it affects human development, for example via public expenditures. Unfortunately, the data for this endeavor have not been available. Theoretical expectations about the precise conditions interacting with democracy in the creation of a healthy and literate society have not been met. The interaction of democracy and its other presumed conditions of functioning turned out to be insignificant or not robust to different democracy measures or samples. One could therefore conclude that the functioning of democracy - in terms of non-income human development improvements - is rather independent of GDP per capita, inequality, education and also ethnic fractionalization. But the missing robustness of our interaction effects does not permit any inferences.

GDP per capita, education and ethnic fractionalization influence non-income human development levels directly. A high level of economic development and education is related to a high level of non-income human development. High social fragmentation, on the contrary, is associated with lower levels of non-income human development. Income inequality has rather ambiguous results and turns out to be insignificant in most cases, a result that weakens the income inequality hypothesis according to which income inequality worsens well-being in a society.

To sum up: our empirical analysis cannot establish causality. However, based on the theoretical reasoning, the statistical associations suggest, that what is important is democracy itself and only to a smaller extent the circumstances under which it occurs. First, living in a democracy is associated with better health and education, independently from the level of economic development in a country. Secondly, even if the picture here is more ambiguous, the positive association between democratic systems and human development seems to be rather independent of the circumstances. This stands in contrast to what the theoretical literature has told us. However, it can be considered good news for promoting democracy in poor, fragmented or uneducated societies.

Since income inequality did not play a major role in our estimations we found no supporting evidence for the median voter theory. This might be due to different degrees of inequality aversion in a country, although the region dummies in the regression analysis controlled for cultural factors that might capture differences in inequality aversion. Nevertheless, as democracy is positively associated with the well-being of a population, the main question of this paper deserves an affirmative answer. We thus cautiously support Sen's argument that democracy fulfils its "constructive" and "instrumental" role.

Chapter 2

The Institutional Basis of Gender Inequality: The Social Institutions and Gender Index (SIGI)[1]

2.1   Introduction

Despite considerable progress in recent decades, gender inequality in the manifold dimensions of well-being remains pervasive in many developing countries. This is an intrinsic issue of equity as the affected women are deprived of their basic freedoms (Sen, 1999b). But going beyond this intrinsic feature of gender inequality, there is considerable evidence that it implies high costs for society in the form of lower human capital, worse governance, and lower growth (e.g. World Bank, 2001; Klasen, 2002; Klasen and Lamanna, 2009). The intrinsic and instrumental value of gender equality has been recognized and incorporated in the development agenda, for example in Millennium Development Goal 3 "Promote gender equality and empower women" as well as the Convention on the Elimination of Discrimination against Women.

To measure the extent of this problem at the cross-country level several gender-related indices have been proposed, e.g. the Gender-Related Development Index (GDI) and the Gender Empowerment Measure (GEM) (United Nations Development Programme, 1995), the Global Gender Gap Index from the World Economic Forum (Lopez-Claros and Zahidi, 2005), the Gender Equity Index developed by Social Watch (2005) or the African Gender Status Index proposed by the Economic Commission for Africa (2004). These measures focus on gender inequality in well-being or in agency and they are typically outcome-focused (Klasen, 2006, 2007).

---

[1]joint work with Boris Branisa and Stephan Klasen

Focusing only on outcomes neglects the question of the origins of these inequalities and their great heterogeneity across space and time. Gender inequality is the result of human behavior, and how people behave and interact is influenced by institutions. Thus to understand gender inequality in outcomes, one needs to study the institutional basis of gender inequality.

There are several approaches to institutions. According to North (1990, p. 3 ff.) "institutions are the rules of the game in a society", they are "humanly devised constraints that shape human interaction". From an economics perspective, institutions are conceived as the result of collective choices in a society to achieve gains from cooperation by reducing uncertainty, collective action dilemmas and transaction costs. A sociological or cultural perspective, which is complementary to the rational choice one, relates institutions to culture. Institutions in this sense frame meanings and beliefs. People try to satisfy norms rather than to act individually within the rules of the game, i.e. institutions do not canalize preferences of actors, they influence the preferences and shape the role models and identities of the actors themselves. Legitimacy and appropriateness as well as cultural authority, power in a society and community dynamics might be more relevant in shaping such institutions that become taken for granted without continuously being evaluated against efficiency considerations (Hall and Taylor, 1996, and references therein).

There is a particular type of institutions that is relevant for gender inequality, *social institutions related to gender inequality*. These institutions are more embedded in the cultural-sociological account although efficiency issues may also be important. We conceive these social institutions as long-lasting norms, values and codes of conduct that find expression in traditions, customs and cultural practices, informal and formal laws. They underlie gender roles and the distribution of power between men and women in the family, in the market and in social and political life. Consequently, they shape the social and economic opportunities of men and women, their autonomy in taking decisions (Dyson and Moore, 1983; Abadian, 1996; Hindin, 2000; Bloom et al., 2001) or their capabilities to live the life they value (Sen, 1999b). That is why they might affect important development outcomes and contribute to outcome gender inequalities (De Soysa and Jütting, 2007).

Three measures proxy in one way or another social institutions, which determine how women are treated in society: the Women's Political Rights index (WOPOL), the Women's Economic Rights index (WECON), and the Women's Social Rights index (WOSOC) of the CIRI Human Rights Data Project.[2] These indices take a human rights perspective and measure on a yearly basis whether a number of internationally recognized rights for women are included in law and whether government enforces them. From the three indices, WOSOC is

---

[2]Information is available on the webpage of the project http://ciri.binghamton.edu/.

the most encompassing measure covering social relations (Bjørnskov et al., 2009). However, it does not allow one to differentiate between different dimensions of social institutions. For example, it is important to distinguish between what happens within the family and what happens in public and social life. Furthermore, other shortcomings of all three indices are that they also cover outcomes of institutions, and they can only take four values from 0 (no rights) to 3 (legally guaranteed and enforced rights) which makes it difficult to compare and rank countries as there are many ties, i.e. equal scores, in the data.

In this paper we propose new composite measures that proxy social institutions related to gender inequality in non-OECD countries which are based on variables of the OECD Gender, Institutions and Development (GID) database (Morrisson and Jütting, 2005; Jütting et al., 2008). These are the Social Institutions and Gender Index (SIGI) as a multidimensional measure of the deprivation of women and its five one-dimensional subindices Family code, Civil liberties, Physical integrity, Son preference and Ownership rights.

In general, the construction of composite measures requires several decisions, for example about the weighting scheme and the method of aggregation (e.g. Nardo et al., 2005). The subindices as one-dimensional measures are built using the method of polychoric principal component analysis to extract the common information of the variables corresponding to a subindex (Kolenikov and Angeles, 2009). When we combine the subindices to construct the SIGI, we use a reasonable methodology to capture the multidimensional deprivation of women caused by social institutions. The formula of the SIGI is inspired by the Foster-Greer-Thorbecke poverty measures (Foster et al., 1984) and offers a new way of aggregating gender inequality in several dimensions measured by the subindices. It is transparent and easy to understand, it penalizes high inequality in each dimension and allows only for partial compensation between dimensions.

The SIGI and the subindices are useful tools to compare the societal situation of women in over 100 non-OECD countries from a new perspective, allowing the identification of problematic countries and dimensions of social institutions that deserve attention by policy makers and need to be scrutinized in detail. Empirical results show that the SIGI provides additional information to that of other well-known gender-related indices. Moreover, regression analysis shows that the SIGI is related to indices that measure outcome gender inequality, even if one controls for region, religion and the level of economic development.

This paper is organized as follows. In section 2.2, we describe the OECD GID Database. Then, in sections 2.3 and 2.4 we focus on the construction of the subindices and of the SIGI. In section 3.5.2, we present empirical results by country, interesting regional patterns and a comparison between the SIGI and other gender-related measures. Furthermore, using regression analysis we illustrate the relevance of the SIGI for explaining outcome gender

inequality. The last section concludes with a discussion of the strengths and weaknesses of the proposed measures.

2.2   The OECD Gender, Institutions and Development (GID) Database

As input for the composite measures we use variables from the OECD GID database (Morrisson and Jütting, 2005; Jütting et al., 2008). This is a cross-country database covering about 120 countries with more than 20 variables measuring social institutions related to gender inequality.[3] These variables proxy social institutions through prevalence rates, legal indicators or indicators of social practices. We assume that the concept social institutions related to gender inequality is multidimensional. Following previous work done by the OECD (Jütting et al., 2008) we choose twelve variables that are assumed to measure each one of four dimensions of social institutions.

The *Family code* dimension refers to the private sphere with institutions that influence the decision-making power of women in the household. Family code is measured by the following four variables. *Parental authority* measures whether women have the right to be the legal guardian of a child during marriage, and whether women have custody rights over a child after divorce. *Inheritance* is based on formal inheritance rights of spouses. *Early marriage* measures the percentage of girls between 15 and 19 years of age who are/were ever married. *Polygamy* measures the acceptance of polygamy in the population. Countries where this information is not available are assigned scores based on the legality of polygamy.[4]

The public sphere is measured by the *Civil liberties* dimension that captures the freedom of social participation of women and includes the following two variables. *Freedom of movement* indicates the freedom of women to move outside the home. *Freedom of dress* is based on the obligation of women to use a veil or burqa to cover parts of their body in public.

The *Physical integrity* dimension comprises different indicators on violence against women. The variable *violence against women* indicates the existence of laws against domestic violence, sexual assault or rape, and sexual harassment. *Female genital mutilation* is the percentage of women who have undergone female genital mutilation. *Missing women* measures gender bias in mortality. Countries were coded based on estimates of gender bias in mortality

---

[3]The data are available at the web-pages http://www.wikigender.org and http://www.oecd.org/dev/gender/gid.

[4]Acceptance of polygamy in the population might proxy actual practices better than the formal indicator legality of polygamy and, moreover, laws might be changed faster than practices. Therefore, the acceptance variable is the first choice for the subindex Family code. The reason for using legality when acceptance is missing is to increase the number of countries.

for a sample of countries (Klasen and Wink, 2003) and on sex ratios of young people and adults.

The *Ownership rights* dimension covers the economic sphere of social institutions prox-ied by the access of women to several types of property. *Women's access to land* indicates whether women are allowed to own land. *Women's access to bank loans* measures whether women are allowed to access credits. *Women's access to property other than land* covers mainly access to real property such as houses, but also any other property.

Concerning the *missing women* variable in the *Physical integrity* dimension, it could be argued that it reflects another dimension of gender inequality. Missing women is an extreme manifestation of son preference under scarce resources. 100 million women are not alive who should be alive if women were not discriminated against (Sen, 1992; Klasen and Wink, 2003). The other components of *Physical integrity*, *violence against women* and *female gen-ital mutilation*, measure particularly the treatment of women which is not only motivated by economic considerations. In the next section, we check with statistical methods if *miss-ing women* measures another dimension as the variables *violence against women* and *female genital mutilation*.

These twelve variables are between 0 and 1. The value 0 means no or very low inequality and the value 1 indicates high inequality. Three of the variables (*early marriage, female genital mutilation and violence against women*) are continuous. The other indicators measure social institutions on an ordinal categorical scale. The chosen variables cover around 120 non-OECD countries from all regions in the world except North America.[5] The choice of the variables is also guided by the availability of information so that as many countries as possible can be ranked by the SIGI. Within our sample 102 countries have information for all twelve variables.

2.3   Construction of the Subindices

The objective of the subindices is to provide a summary measure for each dimension of social institutions related to gender inequality. In every subindex we want to combine variables that are assumed to belong to one dimension. The first step is to check the statistical association between the variables. The second step consists in aggregating the variables with a reasonable weighting scheme.

---

[5]The OECD Gender, Institutions and Development Database does not contain variables that capture relevant social institutions related to gender inequality in OECD countries.

2.3.1   Measuring the Association Between Categorical Variables

To check the association between variables, and as most of them are ordinal, we use Kendall Tau b and Multiple Joint Correspondence Analysis (Greenacre, 2007; Nenadić, 2007). Kendall Tau b is a rank correlation coefficient. These measures are useful when the data are ordinal and thus the conditions for using Pearson's correlation coefficient are not fulfilled. For each variable, the values are ordered and ranked. Then the correspondence between the rankings is measured.[6]

Taking into account tied pairs, the formula for Kendall Tau b is

$$\tau_b = \frac{C - D}{\sqrt{\frac{n(n-1)}{2 - T_x}\frac{n(n-1)}{2 - T_y}}}, \tag{2.1}$$

where C is the number of concordant pairs, D is the number of discordant pairs, n is the number of observations, $\frac{n(n-1)}{2}$ is the number of all pairs, T_x is the number of pairs tied on the variable x and T_y is the number of pairs tied on the variable y. The notation is taken from Agresti (1984).

As a second method to check the association between variables we examine the graphics produced by Multiple Joint Correspondence Analysis (MJCA) (Greenacre, 2007; Nenadić, 2007), after having discretized the three continuous variables. Correspondence Analysis is a method for analyzing and representing the structure of contingency tables graphically. We use MJCA to find out whether variables seem to measure the same.[7]

The results for Kendall Tau b (Tables 5.17- 5.21) are reported in Appendix 2. A significant positive value of Kendall Tau b is a sign for a positive association between two variables. This is the case for all variables belonging to one dimension, except *missing women* in the

---

[6]For calculating Kendall Tau, one counts the number of concordant and discordant pairs of two rankings, builds the difference and divides this difference by the total number of pairs. A value of 1 means total correspondence of rankings, i.e. the rankings are the same. A value of -1 indicates reverse rankings or a negative association between rankings. A value of 0 means independence of rankings. Kendall Tau b is a variant of Kendall tau that corrects for ties, which are frequent in the case of discrete data (Agresti, 1984, chap. 9). We consider Kendall Tau b to be the appropriate measure of rank correlation to find out whether our data are related.

[7]Correspondence Analysis is an exploratory and descriptive method to analyze contingency tables. Instead of calculating a correlation coefficient to capture the association of variables, the correspondence of conditional and marginal distributions of either rows or columns - also called row or column profiles - is measured using a χ^2-statistic, that captures the distance between them. These row or column profiles then are plotted in a low-dimensional space, so that the distances between the points reflect the dissimilarities between the profiles. Multiple Joint Correspondence Analysis is an extended procedure for the analysis of more than two variables and considers the cross-tabulations of the variables against each other in a so-called Burt matrix but with modified diagonal sub-tables. This facilitates to figure out whether variables are associated. This is the case when they have similar deviations from homogeneity, and therefore get a similar position in a profile space (Greenacre, 2007; Nenadić, 2007).

subindex *Physical integrity*. The graphs produced with MJCA (Figures 5.1- 5.5) are also in Appendix 2.[8] The results of MJCA confirm that within every dimension all the variables seem to measure the same dimension, with the exception of *missing women* in the dimension *Physical integrity*. These results support the argumentation in section 2.2.

We decide to use the variable *missing women* as a fifth subindex called *Son preference*. The artificially higher female mortality is one of the most important and cruel aspects of gender inequality and should not be neglected, as over 100 million women that should be alive are missing (Sen, 1992; Klasen and Wink, 2003). Missing women is the "starkest manifestation of the lack of gender equality" (Duflo, 2005).

2.3.2   Aggregating Variables to Build a Subindex

The five subindices *Family code*, *Civil liberties*, *Son preference*, *Physical integrity* and *Ownership rights* use the twelve variables as input that were mentioned in the previous section to measure each one dimension of social institutions related to gender inequality. In the case of Son preference, the subindex takes the value of the variable missing women. In all other cases, the computation of the subindex values involves two steps.

In a first step, the method of polychoric principal component analysis is used to extract the common information of the variables corresponding to a subindex. Principal component analysis (PCA) is a method of dimensionality reduction that is valid for normally distributed variables (Jolliffe, 1986). This assumption is violated in this case, as the data include variables that are ordinal, and hence the Pearson correlation coefficient is not appropriate. Following Kolenikov and Angeles (2004, 2009) we use polychoric PCA, which relies on polychoric and polyserial correlations. These correlations are estimated with maximum likelihood, assuming that there are latent normally distributed variables that underly the ordinal categorical data. We use the First Principal Component (FPC) as a proxy for the common information contained by the variables corresponding to the subindices. The first principal component is the weighted sum of the standardized original variables that captures as much of the variance in the data as possible.[9] The standardization of the original variables is done as follows. In the case of continuous variables, one subtracts the mean and then divides by the standard deviation. In the case of ordinal categorical variables, the standardization uses results of an ordered probit model. The weight that each variable gets in these linear combi-

---

[8]The graphs produced with MJCA can be interpreted in the following way. In most cases, one of the axes represents whether there is inequality and the other axe represents the extent of inequality. If one connects the values of a variable one obtains a graphical pattern. If this is similar to the pattern obtained for another variable, then both variables are associated.

[9]The proportion of explained variance by the first principal component is 70% for *Family code*, 93% for *Civil liberties*, 60% for *Physical integrity* and 87% for *Ownership rights*.

nations is obtained by analyzing the correlation structure in the data. The weights are shown in Table 2.1.

Table 2.1: Weights from Polychoric PCA

	Weights
Family code	
Parental authority	0.5212
Inheritance	0.5404
Early marriage	0.3877
Polygamy	0.5348
Civil liberties	
Freedom of movement	0.7071
Obligation to wear a veil	0.7071
Physical integrity	
Female genital mutilation	0.7071
Violence against women	0.7071
Ownership rights	
Women´s access to land	0.5811
Women´s access to loans	0.5665
Women´s access to other property	0.5843

In a second step, the subindex value is obtained rescaling the FPC so that it ranges from 0 to 1 to ease interpretation. A country with the best possible performance (no inequality) is assigned the value 0 and a country with the worst possible performance (highest inequality) the value 1. Hence, the subindex values of all countries are between 0 and 1. Using the score of the FPC the subindex is calculated using the following transformation. Country *X* corresponds to a country of interest, Country *Worst* corresponds to a country with worst possible performance and Country *Best* is a country with best possible performance.

$$
\begin{aligned}
\text{Subindex(Country X)} \quad = \quad & \frac{\text{FPC(Country X)}}{\text{FPC(Country Worst)} - \text{FPC(Country Best)}} \\
- \quad & \frac{\text{FPC(Country Best)}}{\text{FPC(Country Worst)} - \text{FPC(Country Best)}}
\end{aligned}
\tag{2.2}
$$

To check whether the subindices are empirically non-redundant, so that each of them provides additional information, we conduct an empirical analysis of the statistical association between them. In the case of well-being measures, McGillivray and White (1993) suggest using two explicit thresholds to separate redundancy from non-redundancy, that is a correlation coefficient of 0.90 and 0.70. Based on this suggestion we use the threshold 0.80. In

Table 2.2 we present Kendall tau b as a measure of the statistical association between the five subindices. In all cases, the subindices are positively correlated, showing that they all measure social institutions related to gender inequality. It must be noted, however, that the correlation is not always statistically significant. Kendall tau b is lower than 0.80 in all cases, which means that each subindex measures a distinct aspect of social institutions related to gender inequality.

Table 2.2: Kendall Tau b Between Subindices

		Family code	Civil liberties	Physical integrity	Son preference	Ownership rights
Family code	Kendall tau b	1				
	Number obs.	112				
Civil liberties	Kendall tau b	0.3844	1			
	Number obs.	112	123			
	p-value	0.0000				
Physical integrity	Kendall tau b	0.4367	0.2648	1		
	Number obs.	103	113	114		
	p-value	0.0000	0.0005			
Son preference	Kendall tau b	0.1603	0.4264	0.0272	1	
	Number obs.	112	122	114	123	
	p-value	0.0317	0.0000	0.7220		
Ownership rights	Kendall tau b	0.5484	0.3047	0.3937	0.1039	1
	Number obs.	111	121	112	121	122
	p-value	0.0000	0.0001	0.0000	0.181	

2.4 The Social Institutions and Gender Index (SIGI)

With the subindices described in the last section as input, we build a multidimensional composite index named Social Institutions and Gender Index (SIGI) which reflects the deprivation of women caused by social institutions related to gender inequality. The proposed index is transparent and easy to understand. As in the case of the variables and of the subindices, the index value 0 corresponds to no inequality and the value 1 to complete inequality.

The SIGI is an unweighted average of a non-linear function of the subindices. We use equal weights for the subindices, as we see no reason for valuing one of the dimensions more or less than the others.[10] The non-linear function arises because we assume that inequality

---

[10]Empirically, even in the case of equal weights the ranking produced by a composite index is influenced by the different variances of its components. The component that has the highest variance has the largest influence

in gender-related social institutions leads to deprivation experienced by the affected women, and that deprivation increases more than proportionally when inequality increases. Thus, high inequality is penalized in every dimension. The non-linearity also means that the SIGI does not allow for total compensation among subindices, but permits partial compensation. Partial compensation implies that high inequality in one dimension, i.e. subindex, can only be partially compensated with low inequality on another dimension.[11]

For our specific five subindices, the value of the index the SIGI is then calculated as follows.

$$
\begin{aligned}
\text{SIGI} \;=\; & \frac{1}{5}\,(\text{Subindex Family Code})^2 + \frac{1}{5}\,(\text{Subindex Civil Liberties})^2 \\
+ \; & \frac{1}{5}\,(\text{Subindex Physical Integrity})^2 + \frac{1}{5}\,(\text{Subindex Son preference})^2 \\
+ \; & \frac{1}{5}\,(\text{Subindex Ownership Rights})^2
\end{aligned}
$$

Using a more general notation, the formula for the SIGI $I(X)$, where X is the vector containing the values of the subindices x_i with $i = 1, ..., n$, is derived from the following considerations. For any subindex x_i, we interpret the value 0 as the goal of no inequality to be achieved in every dimension. We define a deprivation function $\phi(x_i, 0)$, with $\phi(x_i, 0) > 0$ if $x_i > 0$ and $\phi(x_i, 0) = 0$ if $x_i = 0$ (e.g. Subramanian, 2007). Higher values of x_i should lead to a penalization in $I(X)$ that should increase with the distance x_i to zero. In our case the deprivation function is the square of the distance to 0 so that deprivation increases more than proportionally as inequality increases.

$$
SIGI = I(X) = \frac{1}{n}\sum_{i=1}^{n}\phi(x_i, 0) = \frac{1}{n}\sum_{i=1}^{n}(x_i - 0)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i)^2.
$$

The formula is inspired by the Foster-Greer-Thorbecke (FGT) poverty measures (Foster et al., 1984). The general FGT formula is defined for $y_i \leq z$ as:

$$
FGT(Y, \alpha, z) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z - y_i}{z}\right)^{\alpha},
$$

where Y is the vector containing all incomes, y_i with $i = 1, ..., n$ is the income of individual i, z is the poverty line, and $\alpha > 0$ is a penalization parameter.

---

on the composite index. In the case of the SIGI the variances of the five components are reasonably close to each other, *Ownership rights* having the largest and *Physical integrity* having the lowest variance.

[11]Other approaches have also been proposed in the literature, e.g. the non-compensatory approach by Munda and Nardo (2005a,b).

To compute the SIGI, the value 2 is chosen for α as the square function has the advantage of easy interpretation. With $\alpha = 2$ the *transfer principle* is satisfied (Foster et al., 1984). In the context of poverty this principle means that a transfer from a person below the poverty line to a person less poor will raise poverty if the set of poor remains unchanged. In the case of the SIGI, the transfer principle means that an increase in inequality in one dimension and a decrease of inequality in another dimension of the same magnitude will raise the SIGI.

Some differences between the SIGI and the FGT measures must be highlighted. In the case of the SIGI, we are aggregating across dimensions and not over individuals. Moreover, in contrast to the income case, a lower value of x_i is preferred, and the normalization achieved when dividing by the poverty line z is not necessary as $0 \le x_i \le 1$, $i = 1, \ldots, n$.

The SIGI fulfills several properties. For a formal presentation of the properties and the proofs, see Appendix 2.

- *Support and range*: The value of the index can be computed for any values of the subindices, and it is always between 0 and 1.

- *Anonymity*: Neither the name of the country nor the name of the subindex have an impact on the value of the index.

- *Unanimity or Pareto Optimality*: If a country has values for every subindex that are lower than or equal to those of another country, then the index value for the first country is lower than or equal to the one for the second country.

- *Monotonicity*: If one country has a lower value for the index than a second country, and a third country has the same values for the subindices as the first country, except for one subindex which is lower, then the third country has a lower index value than the second country.

- *Penalization of dispersion*: For two countries with the same average value of the subindices, the country with the lowest dispersion of the subindices gets a lower value for the index.

- *Compensation*: Although the SIGI is not conceived for changes over time this property is more intuitively understood in the following way. If a country experiences an increase in inequality by a given amount on a subindex, then the country can only have the same value of the index as before, if there is a decrease in inequality on another subindex that is higher in absolute value than the increase.

To highlight the effects of partial compensation as compared to total compensation we computed the statistical association between the SIGI and a simple arithmetic average of the five subindices that allows for total compensation and compared the country rankings of both measures in Appendix 2.[12] The Pearson correlation coefficient between the SIGI and the simple arithmetic average of the five subindices is 0.96 and statistically significant showing a high correlation between both measures. However, when we compare the ranks of the SIGI with those obtained using a simple arithmetic average of the five subindices in Table 5.22 in Appendix 2, we observe that there are noticeable differences in the rankings of the 102 included countries. Examples are China and Nepal. China ranks in position 55 using the simple average, but worsens to place 83 in the SIGI ranking. Nepal has place 84 when the simple average is used, and improves to rank 65 in the SIGI ranking. For China, this is due to the high value on the subindex *Son preference*, which in the SIGI case cannot be fully compensated with relatively low values for the other subindices. For Nepal we observe the opposite case as all subindices have values reflecting moderate inequality.

2.5   Results

2.5.1   Country Rankings and Regional Patterns

In Table 5.23 in Appendix 2, the results for the SIGI and its five subindices are presented. Among the 102 countries considered by the SIGI Paraguay, Croatia, Kazakhstan, Argentina and Costa Rica have the lowest levels of gender inequality related to social institutions. Sudan is the country that occupies the last position, followed by Afghanistan, Sierra Leone, Mali and Yemen, which means that gender inequality in social institutions is a major problem there.[13]

Rankings according to the subindices are as follows. For *Family code* 112 countries can be ranked. Best performers are China, Jamaica, Croatia, Belarus and Kazakhstan. Worst performers are Mali, Chad, Afghanistan, Mozambique and Zambia. In the dimension *Civil liberties* 123 countries are ranked. Among them 83 share place 1 in the ranking. Sudan, Saudi Arabia, Afghanistan, Yemen and Iran occupy the last five positions of high inequality. 114 countries can be compared with the subindex *Physical Integrity*. Hong Kong, Bangladesh, Chinese Taipei, Ecuador, El Salvador, Paraguay and Philippines are at the top of the ranking while Mali, Somalia, Sudan, Egypt and Sierra Leone are at the bottom. In the dimension *Son*

---

[12]We cannot compare the SIGI with the results of the non-compensatory index as proposed by Munda and Nardo (2005a,b). The algorithm used for calculating non-compensatory indices compares pairwise each country for each subindex. However, as our dataset includes many countries with equal values on several subindices, the numerical algorithm cannot provide a ranking.

[13]The subindices are computed for countries that have no missing values on the relevant input variables. In the case of the SIGI only countries that have values for every subindex are considered.

preference 88 out of 123 countries rank at the top as they do not have problems with missing women. The countries that rank worst are China, Afghanistan, Papua New Guinea, Pakistan, India and Bhutan. Finally, 122 countries are ranked with the subindex *Ownership rights*. 42 countries share position 1 as they have no inequality in this dimension. On the other hand, the four worst performing countries are Sudan, Sierra Leone, Chad and the Democratic Republic of Congo.

Table 2.3: Regional Pattern of the Composite Index and Subindices

	ECA	LAC	EAP	SA	SSA	MENA	Total
SIGI							
Quintile 1	6	10	4	0	1	0	**21**
Quintile 2	6	8	5	0	0	1	**20**
Quintile 3	1	1	2	1	14	2	**21**
Quintile 4	0	0	1	2	13	4	**20**
Quintile 5	0	0	1	4	10	5	**20**
Total	**13**	**19**	**13**	**7**	**38**	**12**	**102**
Family Code							
Quintile 1	7	11	4	0	1	0	**23**
Quintile 2	5	8	6	1	0	2	**22**
Quintile 3	1	1	4	3	9	5	**23**
Quintile 4	0	0	0	0	15	7	**22**
Quintile 5	0	0	0	3	16	3	**22**
Total	**13**	**20**	**14**	**7**	**41**	**17**	**112**
Civil Liberties							
Quintile 1, 2, 3	17	22	14	0	27	3	**83**
Quintile 4	0	0	1	3	12	3	**19**
Quintile 5	0	0	2	4	3	12	**21**
Total	**17**	**22**	**17**	**7**	**42**	**18**	**123**
Physical Integrity							
Quintile 1	5	13	5	3	4	2	**32**
Quintile 2	4	4	1	0	3	2	**14**
Quintile 3	7	5	7	3	6	4	**32**
Quintile 4	0	0	3	1	13	2	**19**
Quintile 5	0	0	0	0	14	3	**17**
Total	**16**	**22**	**16**	**7**	**40**	**13**	**114**
Missing Women							
Quintile 1, 2, 3	15	21	10	1	38	3	**88**
Quintile 4	0	1	4	0	4	3	**12**
Quintile 5	1	0	3	6	1	12	**23**
Total	**16**	**22**	**17**	**7**	**43**	**18**	**123**
Ownership Rights							
Quintile 1	12	12	11	1	2	4	**42**
Quintile 2	2	4	2	0	1	1	**10**
Quintile 3	2	3	2	1	8	7	**23**
Quintile 4	1	1	2	4	18	6	**32**
Quintile 5	0	0	0	1	14	0	**15**
Total	**17**	**20**	**17**	**7**	**43**	**18**	**122**

ECA stands for Europe and Central Asia, LAC for Latin America and the Caribbean, EAP for East Asia and Pacific, SSA for Sub-Saharan Africa, and MENA for Middle East and North Africa.

To find out whether apparent regional patterns in social institutions related to gender inequality are systematic, we divide the countries in quintiles following the scores of the SIGI and its subindices (Table 2.3). The first quintile includes countries with lowest inequality, and the fifth quintile countries with highest inequality.

For the SIGI, no country of Europe and Central Asia (ECA) or Latin America and the Caribbean (LAC) is found in the two quintiles reflecting social institutions related to high gender inequality. In contrast, most countries in South Asia (SA), Sub-Saharan Africa (SSA), and Middle East and North Africa (MENA) rank in these two quintiles. It is interesting to note that in the most problematic regions two countries rank in the first two quintiles. These are Mauritius (SSA) and Tunisia (MENA). East Asia and Pacific (EAP) has countries in all five quintiles with Philippines, Thailand, Hong Kong and Singapore in the first quintile and China in the fifth quintile.

Going on with the subindices the patterns are similar to the one of the SIGI. As more information is available for the subindices, the number of countries covered by every subindex is different and higher than for the SIGI. In the following some interesting facts are highlighted, especially those countries whose scores are different than the average in the region.

- *Family code*: No country in ECA, LAC or EAP shows high inequality. SA, MENA and SSA remain problematic with countries with social institutions related to high gender inequality. Exceptions are Bhutan in SA, Mauritius in SSA, and Tunisia and Israel in MENA.

- *Civil liberties*: Only three groups of countries using the quintile analysis can be generated with the first group including the first three quintiles. In SSA over one-half of the countries are now in the first group. Also in MENA there are some countries with good scores (Israel, Morocco and Tunisia). No country in SA is found in the first three quintiles of low and moderate inequality.

- *Physical integrity*: Most problematic regions are SSA and MENA. Exceptions in these regions are Botswana, Mauritius, South Africa and Tanzania (SSA), and Morocco and Tunisia (MENA).

- *Son preference*: Again only three groups of countries can be built by quintile analysis, with the first group including the first three quintiles. As in the case of Civil liberties most of the countries in SSA do not show problems. Missing women is mainly an issue in SA and MENA. But in both regions there are countries that rank in the first group. These are Sri Lanka in SA, and Israel, Lebanon and Occupied Palestinian Territory in MENA.

- *Ownership rights*: Most problematic regions are SA, SSA and MENA. Nevertheless, there are cases in these regions that rank in the first quintile. These are Egypt, Israel, Kuwait and Tunisia (MENA), Bhutan (SA), and Eritrea and Mauritius (SSA).

2.5.2 Simple Correlation with other Gender-related Indices

The SIGI is an important measure to understand gender inequality as it measures institutions that influence the basic functioning of society and explain gender inequality in outcomes. From this perspective, the SIGI has an added value to other gender-related measures irrespective from an empirical redundancy perspective, i.e. whether it provides additional information as compared to other measures.

Nevertheless, one can check whether the index is empirically redundant by computing the statistical association between the SIGI and other well-known gender-related indices. Relying on McGillivray and White (1993) we use a correlation coefficient of 0.80 in absolute value as the threshold to separate redundancy from non-redundancy.

We calculate Pearson correlation coefficient and Kendall Tau b as a measure of rank correlation between the SIGI and each of the following indices: the Gender-related Development Index (GDI) and the Gender Empowerment Measure (GEM) from United Nations Development Programme (2006), the Global Gender Gap Index (GGG) from Hausmann et al. (2007) and the Women's Social Rights Index.[14] As the GDI and the GEM have been criticized in the literature (e.g. Klasen, 2006; Schüler, 2006), we also do the analysis for two alternative measures, the Gender Gap Index Capped (GGI) and a revised Gender Empowerment Measure (GEM2) based on income shares proposed by Klasen and Schüler (2009).[15] For all the indices considered both measures of statistical association are lower than 0.80 in absolute value and statistically significant (Table 2.4). We conclude that the SIGI is related to these gender measures but is non-redundant. The comparison of the country rankings of the SIGI and these other measures can be found in Table 5.24 in Appendix 2.

---

[14]Data obtained from http://ciri.binghamton.edu/.

[15]The Gender Gap Index Capped (GGI) is a geometric mean of the ratios of female to male achievements in the dimensions health, education and labor force participation. "Capped" means that every component is capped at one before calculating the geometric mean. This is necessary as a better relative performance of women, e.g. in the dimension health can be due to a risky behavior of men that should not be rewarded. GGI can be more directly interpreted as a measure of gender inequality while the GDI measures human development penalizing gender inequality. The GEM has three components, political representation, representation in senior positions in the economy, and power over economic resources. The most problematic component is power over economic resources proxied by earned incomes. This component measures female and male earned incomes using income levels adjusted for gender gaps but not the gender gaps themselves. The revised version GEM2 uses income shares of males and females.

Table 2.4: Statistical Association Between the SIGI and Other Gender-related Measures

GDI	Kendall tau b	-0.501	Pearson Corr. Coeff.	-0.5852
Number obs. 79	p-value	0.0000	p-value	0.0000
GGI (capped)	Kendall tau b	-0.5088	Pearson Corr. Coeff.	-0.7169
Number obs. 85	p-value	0.0000	p-value	0.0000
GEM	Kendall tau b	-0.425	Pearson Corr. Coeff.	-0.7024
Number obs. 33	p-value	0.0005	p-value	0.0000
GEM (revised)	Kendall tau b	-0.4402	Pearson Corr. Coeff.	-0.7507
Number obs. 33	p-value	0.0003	p-value	0.0000
GGG	Kendall tau b	-0.4741	Pearson Corr. Coeff.	-0.7295
Number obs. 73	p-value	0.0000	p-value	0.0000
WOSOC	Kendall tau b	-0.4861	Pearson Corr. Coeff.	-0.5266
Number obs. 99	p-value	0.0000	p-value	0.0000

Data for the Gender-related development Index (GDI) and the Gender Empowerment Measure (GEM) are from United Nations Development Programme (2006) and are based on the year 2004. The Gender Gap Index (GGI) capped and the revised Gender Empowerment Measure (GEM revised) are taken from Klasen and Schüler (2009) based on the year 2004. Data for the Global Gender Gap Index (GGG) are from Hausmann et al. (2007). The Women's Social Rights Index (WOSOC) data correspond to the year 2007 and are obtained from http://ciri.binghamton.edu/. The p-values correspond to the null hypothesis that the SIGI and the corresponding measure are independent.

2.5.3   Regression Analysis

The SIGI is aimed to measure the institutional basis of gender inequality. To explore whether the SIGI is associated with gender inequality in outcomes controlling for other factors we run linear regressions with two well-known indices of gender inequality as dependent variables and the SIGI as a regressor. We choose the Global Gender Gap Index (GGG) as the first response variable because it is an encompassing measure reflecting gaps in outcome variables related to basic rights such as health, economic participation and political empowerment. The second response variable is the ratio of GDI to HDI being a composite measure of gender inequality in the dimensions health, education and income. As the GDI is not really a measure of gender inequality, but measures human development penalizing gender inequality, UNDP recommends using the ratio of GDI to HDI.[16]

In both regressions we control for the level of economic development using the log of per capita GDP in constant prices (US$, PPP, base year: 2005) (World Bank, 2008); for religion using a Muslim majority and a Christian majority dummy, the left-out category being countries that have neither a majority of Muslim nor a majority of Christian population (Central Intelligence Agency, 2009); and for geography and other unexplained heterogeneity that

---

[16]http://hdr.undp.org/en/statistics/indices/gdi_gem/, date of access: April 16, 2010

might go together with region using region dummies, the left-out category being Sub-Saharan Africa. As the number of observations is lower than 100, we use HC3 robust standard errors proposed by Davidson and MacKinnon (1993) to account for possible heteroscedasticity in our data.

Table 2.5: Linear Regression with Dependent Variables GGG and Ratio GDI to HDI

	GGG	Ratio GDI to HDI
SIGI	-0.282***	-0.053***
	(0.090)	(0.017)
Log GDP	0.014*	0.004
	(0.008)	(0.003)
SA	-0.006	-0.001
	(0.032)	(0.008)
ECA	-0.012	0.007
	(0.018)	(0.005)
LAC	-0.040**	-0.000
	(0.017)	(0.005)
MENA	-0.044	0.000
	(0.028)	(0.011)
EAP	0.004	0.009**
	(0.023)	(0.005)
Muslim	-0.001	-0.002
	(0.018)	(0.006)
Christian	0.026	0.002
	(0.017)	(0.005)
constant	0.567***	0.959***
	(0.064)	(0.020)
Number of obs.	72	78
Adjusted R2	0.615	0.431
Prob > F	0.000	0.000

Note: *** p<0.01, ** p<0.05, * p<0.1
HC3 robust standard errors in brackets.

The regression results are presented in Table 2.5. The regression with GGG as dependent variable includes 72 countries and the coefficient of determination $adjustedR^2$ is 0.62. The SIGI is negatively associated with GGG and significant at the 1% level. The second regression with the ratio of GDI to HDI as dependent variable includes 78 countries and the corresponding $adjustedR^2$ is 0.43. The SIGI is again negatively associated with the response variable and this association is statistically significant at the 1% level. The results suggest that gender inequality in well-being and empowerment is strongly associated with social institutions that shape gender roles.

Even if we include control variables in the regressions we cannot rule out omitted variable bias, but as we consider that social institutions related to gender inequality are relatively stable and long-lasting, we consider that endogeneity does not pose a major problem. To

check that our findings are not driven by observations that have large residuals and/or high leverage, we also run robust regressions which yield similar results.[17]

2.6   Conclusion

In this paper we present composite indices that offer a new approach to gender inequality, which has been neglected in the literature and by other gender measures, which focus mainly on well-being and agency. Instead of measuring gender inequality in education, health, economic or political participation and other dimensions, the proposed measures proxy the underlying social institutions that are mirrored by societal practices and legal norms that might produce inequalities between women and men in developing countries.

Based on 12 variables of the OECD Gender, Institutions and Development (GID) Database (Morrisson and Jütting, 2005; Jütting et al., 2008) we construct five subindices each capturing one dimension of social institutions related to gender inequality: *Family code*, *Civil liberties*, *Physical integrity*, *Son preference* and *Ownership rights*. The Social Institutions and Gender Index (SIGI) combines the subindices into a multidimensional index of deprivation of women caused by social institutions related to gender inequality. With these measures over 100 developing countries can be compared and ranked.

When constructing composite indices one is always confronted with decisions and trade-offs concerning for example the choice and treatment of the variables included, the weighting scheme and the aggregation method. We try to be transparent in our choices. As the subindices are intended each to proxy one dimension of social institutions, we use the method of polychoric PCA to extract the common element of the included variables (Kolenikov and Angeles, 2009). The methodology for constructing the multidimensional SIGI is based on the assumption that in each dimension deprivation of women increases more than proportionally when inequality increases, and that each dimension should be weighted equally. The formula of the SIGI is inspired by the FGT poverty measures (Foster et al., 1984) and has the advantage of penalizing high inequality in each dimension and only allowing for partial compensation among the five dimensions. We consider that the formula to compute the SIGI is easy to understand and to communicate.

However, some limitations of the subindices and the SIGI must be noted. First, a composite index depends on the quality of the data used as input. Social institutions related to gender inequality are hard to measure and the work accomplished by the OECD in building the GID database is an important step forward. It is worthwhile to continue this endeavor and

---

[17]Results are available upon request. The type of robust regression we perform uses iteratively reweighted least squares and is described in Hamilton (1992). A regression is run with ordinary least squares, then case weights based on absolute residuals are calculated, and a new regression is performed using these weights. The iterations continue as long as the maximum change in weights remains above a specified value.

invest more resources in the measurement of social institutions related to gender inequality. This includes data coverage, coding schemes and the refinement of indicators. It would be useful to exploit data available, for example from Demographic and Health Surveys (DHS)[18], that specifically address the perception that women have of violence against women, and to finance surveys in countries where data is not available.

Secondly, by aggregating variables and subindices, one evitable loses some information. Figures and rankings according to the SIGI and the subindices should not substitute a careful investigation of the variables from the database. Furthermore, to understand the situation in a given country additional qualitative information could be valuable.

Thirdly, one should keep in mind that OECD countries are not included in our sample, as social institutions related to gender inequality in these countries are not well captured by the 12 variables used for building the composite measures. This does not mean that this phenomenon is not relevant for OECD countries, but that further research is required to develop appropriate measures.

Nonetheless, the SIGI and its subindices offer a new perspective to understand gender inequality. Empirical results show that the SIGI is statistically non-redundant and adds new information to other well-known gender-related measures. The SIGI and the five subindices can help policy-makers to detect in which developing countries and in which dimensions of social institutions problems need to be addressed. For example, according to the SIGI scores, regions with highest inequality are South Asia, Sub-Saharan Africa, and the Middle East and North Africa. The composite measures can be valuable instruments to generate public discussion. Moreover, the SIGI and its subindices have the potential to influence current development thinking as they highlight social institutions that affect overall development. As is shown in the literature (e.g. Klasen, 2002; Klasen and Lamanna, 2009), gender inequality in education negatively affects overall development. Economic research investigating these outcome inequalities should consider social institutions related to gender inequality as possible explanatory factors. Results from regression analysis show that the SIGI is related to gender inequality in well-being and empowerment, even after controlling for region, religion and the level of economic development.

---

[18]Information is available on the webpage http://www.measuredhs.com/.

Chapter 3

Why We Should All Care About Social Institutions Related to Gender Inequality[1]

3.1 Introduction

Institutions are a major factor explaining development outcomes. They guide human behavior and shape human interaction (North, 1990). Institutions are humanly devised to reduce uncertainty and transaction cost, they are rooted in culture and history and sometimes they are taken for granted and become beliefs (Hall and Taylor, 1996; De Soysa and Jütting, 2007). This study centers on a special type of institutions and their explanatory value for development outcomes: social institutions related to gender inequality.

It is an established fact that gender inequalities come at a cost. Besides the consequences that the affected women experience because they are deprived of their basic freedoms (Sen, 1999b), gender inequalities affect the whole society. They can lead to ill-health, low human capital, bad governance and lower economic growth (e.g. World Bank, 2001; Klasen, 2002). Gender inequalities can be observed in outcomes like education, health and economic and political participation, but they are rooted in gender roles that evolve from institutions that shape everyday life and form role models that people try to fulfill and satisfy. We refer to these long-lasting norms, values and codes of conduct as social institutions related to gender inequality.

We investigate the impact of these social institutions related to gender inequality on development outcomes, controlling for relevant determinants such as religion, political system, geography and the level of economic development. As development outcomes we choose indicators from the fields of education, demographics, health and governance. In particular,

---

[1] joint work with Boris Branisa and Stephan Klasen

we use female secondary schooling, fertility rates, child mortality and governance in the form of rule of law and voice and accountability. We choose these indicators as they are related to economic development and allow us to find out whether social institutions related to gender inequality hinder progress in reaching the Millennium Development Goals.[2]

Most of the studies that have a similar research focus are conducted at the household level and proxy social institutions related to gender with measures of the autonomy or status of women (e.g. Abadian, 1996; Hindin, 2000). At the cross-country level data are scarce and therefore only a few studies are available that center on the development impact of gender-relevant social institutions (e.g. Morrisson and Jütting, 2005; Jütting et al., 2008).

Using the *Social Institutions and Gender Index* (SIGI) and its five subindices *Family code*, *Civil liberties*, *Physical integrity*, *Son preference* and *Ownership rights* proposed in Essay 2, we investigate whether social institutions related to gender inequality are associated with the chosen development outcomes at the cross-country level.[3] These indices cover between 102 and 123 developing countries and are built out of twelve variables of the OECD Gender, Institutions and Development Database that proxy social institutions through prevalence rates, indicators of social practices and legal indicators.(Morrisson and Jütting, 2005; Jütting et al., 2008).[4] The five subindices of the SIGI each measure one dimension of social institutions related to gender inequality.[5] The *Family code* subindex captures institutions that directly influence the decision-making power of women in the household. It is composed of four variables that measure whether women have the right to be the legal guardian of a child during marriage and whether women have custody rights over a child after divorce, whether there are formal inheritance rights for wives, the percentage of girls between 15 and 19 years of age who are/have been married, and the acceptance of polygamy in the population.[6] The *Civil liberties* subindex covers the freedom of social participation of women and combines two variables, freedom of movement of women and freedom of dress, i.e. whether there is

---

[2]In particular, goal 3 "Promote gender equality and empower women", goal 4 "Reduce child mortality" and goal 5 "Improve maternal health" are relevant here, although the other goals can be at least indirectly linked to our chosen indicators.

[3]As discussed in Essay 2, an alternative measure of social institutions would be the Women's Social Rights index (WOSOC) of the CIRI Human Rights Data Project (http://ciri.binghamton.edu/), which measures from a human rights perspective the type of institutions we are interested in. We prefer to work with the SIGI and its subindices and not with WOSOC as the latter also covers outcomes of these institutions and does not allow one to differentiate between dimensions of social institutions, e.g. between what happens within the family and what happens in public life. Moreover, WOSOC can only take four values, from 0 to 3, which makes it difficult to compare countries as there are many ties, meaning equal scores, in the data.

[4]The data are available at the web-pages http://www.wikigender.org and http://www.oecd.org/dev/gender/gid.

[5]To extract the common information of the variables used to construct one subindex the method of polychoric principal component analysis is used (Kolenikov and Angeles, 2009).

[6]Countries where this information is not available are assigned scores based on the legality of polygamy.

an obligation for women to use a veil or burqa to cover parts of their body in public. The *Physical integrity* dimension comprises two indicators of violence against women, the existence of laws against domestic and sexual violence and the percentage of women who have undergone female genital mutilation. The subindex *Son preference* measures the economic valuation of women and is based on a 'missing women' variable that measures an extreme form of preferring boys over girls based on information about the female population that has died as a result of gender inequality. The last subindex *Ownership rights* covers the access of women to several types of property: land, credit and property other than land. The values of the SIGI and of all the subindices are between 0 and 1. The value 0 means no or very low inequality and the value 1 indicates high inequality.

The SIGI combines the five subindices into a multidimensional measure of deprivation of women in a country. The underlying methodology of construction is inspired by the Foster-Greer-Thorbecke poverty measures (Foster et al., 1984). It leads to penalization of high inequality in each dimension and allows for only partial compensation between dimensions. The value of the SIGI is calculated as follows:

$$
\begin{aligned}
\text{SIGI} \ = \ & \frac{1}{5} \left(\text{Subindex Family Code} \right)^2 + \frac{1}{5} \left(\text{Subindex Civil Liberties} \right)^2 \\
& + \ \frac{1}{5} \left(\text{Subindex Physical Integrity} \right)^2 + \frac{1}{5} \left(\text{Subindex Son preference} \right)^2 \\
& + \ \frac{1}{5} \left(\text{Subindex Ownership Rights} \right)^2
\end{aligned}
$$

The main shortcoming of these indices is that they cover only developing countries. This is due to the fact that the variables used as input do not measure relevant social institutions related to gender inequalities in OECD countries. Further research is required to develop appropriate measures for developed countries. Nevertheless, these social institutions indicators are innovative measures of the social, economic and political valuation of women that focus on the roots of gender inequalities and add information to other existing measures of gender inequality in well-being and empowerment.[7] The ranking of countries according to the SIGI and its subindices is presented in Appendix 2 belonging to Essay 2.

We proceed as follows. First, we look for relevant theories linking - at least implicitly - social institutions related to gender inequality with development outcomes such as health,

---

[7]Examples are the Gender-Related Development Index (GDI) and the Gender Empowerment Measure (GEM) from United Nations Development Programme (1995), the Global Gender Gap Index from the World Economic Forum (Lopez-Claros and Zahidi, 2005), the Gender Equity Index developed by Social Watch (Social Watch, 2005), and the African Gender Status Index proposed by the Economic Commission for Africa (Economic Commission for Africa, 2004).

demographics, education and the governance of a society. We refer to bargaining household models (e.g. Manser and Brown, 1980; McElroy and Horney, 1981; Lundberg and Pollak, 1993) and models considering the costs and returns of children (e.g. Becker, 1981; King and Hill, 1993; Hill and King, 1995) as well as to contributions from several disciplines on governance and democracy. These contributions focus on differences in behavior between men and women, and on women's movements as a countervailing power to personal rule (e.g. Swamy et al., 2001; Tripp, 2001). Secondly, we run several linear regressions with the outcome indicators as dependent variables and the SIGI and its subindices as the main explanatory variables. Our results show that social institutions related to gender inequality matter; higher inequality in social institutions is associated with lower development outcomes.[8]

The rest of the paper is organized as follows. In section 3.2 we review existing theory on household decision-making and incorporate social institutions into the models, deriving hypotheses on their impact on female education, fertility and child mortality. In section 3.3, we formulate hypotheses on the impact of social institutions on rule of law, and voice and accountability based on the literature on governance, democracy and gender. Data is described in section 3.4. The empirical estimation and the results are presented in section 3.5. Section 3.6 concludes.

3.2   Social Institutions and Household Decisions

In this section, we review the existing literature about the potentials effects of social institutions related to gender inequality on development outcomes. It is beyond the scope of this study to develop a formal model that incorporates social institutions and specifies the exact functional relationships. Instead, we use the non-unitary approach to the household and the Net Present Value which give hints on how social institutions operate at the household level. These approaches provide the necessary micro-foundation for the empirical analysis which can only be conducted at the macro-level because of the available data.

Non-unitary household models show that household decisions are the result of the distribution of bargaining power in the household. Common to the non-unitary models, initiated by Manser and Brown (1980) and McElroy and Horney (1981), is a game-theoretic approach to the household. Husband and wife have their own utility function, $U^h(c^h)$ for the husband and $U^w(c^w)$ for the wife, that depend each on the consumption of private goods c.[9] They bargain over the allocation of resources to maximize their utility. In the case they do not reach agreement they receive a payoff which corresponds to an individual 'threat point', $P^h(\mathbf{S}, Z)$

---

[8]In a related paper, Jütting and Morrisson (2009) follow the same econometric procedure we use here and study the impact of the SIGI and its subindices on gender inequality on labor market outcomes.

[9]Certainly, there are public goods in the household that both husband and wife consume within the marriage.

and $P^w(\mathbf{S},Z)$ which comprises the utilities associated with non-agreement.[10] $\mathbf{S}$ and Z are defined below. The implication of non-unitary models is that household members do not simply pool resources and that inequality in power may cause inequality in outcomes (Kanbur, 2003; Pollak, 2003, 2007; Lundberg and Pollak, 2008).[11] Empirical evidence supports this (e.g. Thomas, 1997; Schultz, 1990; Haddad and Hoddinott, 1994; Rasul, 2008).

If husband and wife have to take decisions about their sons and daughters which will affect the future then time needs to be considered. The Net Present Value (NPV) allows to take into account present and future costs and returns to investments. To simplify the illustration we ignore that bargaining takes place and name the decision-maker 'parents'. The maximization of utility in a multi-period model leads parents to consider the costs and returns of the investment in their children (e.g. King and Hill, 1993). This private calculation of parents at period $t = 0$ can then be represented with the NPV of the investment in a child, with $NPV = \sum_{t=0}^{T} \frac{R(\mathbf{S},Z)_t - K(\mathbf{S},Z)_t}{(1+r)^t}$ where T is the number of time periods considered, R represents the returns, K the costs of investments in a child, and r represents the discount rate. Like the threat point P in the non-unitary models, R and K are functions of $\mathbf{S}$ and Z that will be explained below. If the NPV is positive parents decide to invest in a child. Gender inequality in the investments in boys and girls arises if the NPV of boys is larger then the one of girls.[12]

Finally, let us explain $\mathbf{S}$ and Z. $\mathbf{S}$ can be defined as 'extrahousehold environmental parameters' (McElroy, 1990) or 'gender-specific environmental parameters' (Folbre, 1997) that influence the threat point in the non-unitary household models and the NPV of a child. We consider that $\mathbf{S}$ can be best described as *social institutions related to gender inequality*. Z represents all other influential factors besides $\mathbf{S}$.

3.2.1 Social Institutions and Female Education

The following examples illustrate how social institutions related to gender inequality affect the private costs and returns of educational investments.[13] Social institutions related to gender

---

[10]The threat point may be external to the marriage. In this case it corresponds to the individual's utility outside the family in case of divorce, as it is modeled in the divorce threat models of Manser and Brown (1980) and McElroy and Horney (1981). In the separate spheres bargaining models of Lundberg and Pollak (1993) the threat point is internal to the marriage and is the utility associated with a non-cooperative equilibrium within marriage given by traditional gender roles and social norms, where the spouses receive benefits due to the joint consumption of public goods.

[11]Using Nash-Bargaining a solution to these non-unitary models can be found. Husband and wife maximize the Nash product function $N = [U^h(c^h - P^h(\mathbf{S},Z)][U^w(c^w - P^w(\mathbf{S},Z)]$, that is subject to a pooled budget constraint. The result is the demand function $c^i = f^i(p,y,\mathbf{S},Z)$ with p for prices, y for total household income and $i = w, h$ (Lundberg and Pollak, 2008).

[12]See Pasqua (2005) who considers both perspectives, the non-unitary approach to the household and the cost and returns approach in the case of education of girls.

[13]It must be noted that the private NPV of investments in the education of children does not correspond to the social NPV. Social returns to education, especially female education, are often higher than the private ones.

inequality influence the costs of education as they shape a gendered division of labor and the opportunity costs of educating girls. Opportunity costs include income from child labor and are higher for girls when they are expected to do housework, to care for their younger siblings or to work in agriculture (Hill and King, 1995; Lahiri and Self, 2007). Social institutions related to gender inequality also affect the returns to education. The returns are generally lower for girls than for boys because girls and women are discriminated on the labor market in the form of entry restrictions and wage gaps. Thus, boys are expected to be economically more productive. Furthermore, parents often expect only low returns from female education because the daughter marries and leaves the house implying that the family loses her labor force. As a consequence sons become the building block of their parents' old-age security (Hill and King, 1995; Pasqua, 2005; Song et al., 2006).[14]

The costs and returns perspective does not rule out that the distribution of decision-making power in the household matters. The non-unitary household approach can be used to explain low female education (Pasqua, 2005). Several empirical studies show that when women dispose of more resources, investments in the education of girls are higher (e.g. Schultz, 2004; Emerson and Souza, 2007).

- *Hypothesis 1: Social institutions that deprive women of their autonomy and bargaining power in the household or that increase the private costs and reduce the private returns to investments into female education are associated with lower female education than in a more egalitarian environment.*

3.2.2   Social Institutions and Fertility and Child Mortality Rates

Social institutions related to gender inequality that influence female decision-making power in the household and the *NPV* of the investment in girls in comparison to boys are also relevant for fertility levels and child mortality.

Concerning fertility, one can use the non-unitary household approach and argue that the net utility of a woman associated with getting a child might differ from that of a man. If one assumes that man and woman derive the same satisfaction of having a child, the net utility a woman derives is lower than the one of the man as she bears most of the costs of having children. These costs are related to the discomfort and health risks related to pregnancy, and

---

There is evidence that society benefits from female education as it contributes to overall development and drives economic growth (Hill and King, 1995; Klasen, 2002; Braunstein, 2007; Klasen and Lamanna, 2009). The resulting investment in female education will then often be sub-optimal.

[14]In addition to all of these considerations, social institutions related to gender inequality might affect the supply of schooling which might influence the decision to send girls to school if school environments are hostile to the needs of girls (e.g. no female teachers available, long distances to school or prices in favor of boys) (Hill and King, 1995; Alderman et al., 1996; Pasqua, 2005; Lahiri and Self, 2007).

the income losses associated with time spent on child care. This might explain why women want less children than men, but cannot achieve their objectives as social institutions restrict their power in limiting the number of children born. Empirical studies support the hypothesis that reduced female bargaining power leads to shorter time spans between births, a lower use of contraceptives and higher fertility levels (Abadian, 1996; Hindin, 2000; Saleem and Bobak, 2005; Seebens, 2008).

The perspective of the *NPV* provides a second explanation for higher fertility. In the absence of well-functioning insurance markets and pension systems, parents in developing countries may need more children to feel secure. Depending on the costs of a child and the returns to the investment in a child parents will consider to get more children. As it was explained in the previous subsection on female education, social institutions related to gender inequality affect the *NPV* of investments in children. If these social institutions lower income earning opportunities for girls, the *NPV* of investments in girls will be lower than the *NPV* of investments in boys. Hence, sons yield the promise of more economic security as compared to daughters. As long as parents cannot perfectly control the sex of their offspring, they will bear more children to increase the chance of having more sons (Abadian, 1996; Kazianga and Klonner, 2009).

To explain higher child mortality levels with social institutions that disadvantage women one has to bear in mind that mothers are usually the primary caregivers of children. Within the non-unitary framework, if mothers have only limited power in the household, they are constrained in the use of health care or in the access to food and other goods necessary for children. Thus, they cannot take care of their children as they would without those restrictions. This might lead to worse child health and higher child mortality rates (Thomas, 1997; Bloom et al., 2001; Smith et al., 2002; Maitra, 2004).

From the *NPV* perspective it might be rational for parents to invest more in the health and nutrition of boys than in girls who as a consequence could suffer more heavily from health problems and experience higher mortality rates than boys. It is possible that this behavior increases overall child mortality rates. In addition, the limited education that women typically receive in patriarchal societies as a result of past *NPV* calculations of their parents might also lead to worse child health and to higher child mortality figures (Schultz, 2002; Shroff et al., 2009).

- *Hypothesis 2: Social institutions that deprive women of their autonomy and bargaining power in the household or that increase the private costs and reduce the private returns of investments into girls are associated with higher fertility levels than in an egalitarian environment.*

- *Hypothesis 3: Social institutions that deprive women of their autonomy and bargaining power in the household or that increase the private costs and reduce the private returns of investments into girls are associated with higher child mortality than in an egalitarian environment.*

3.3   Social Institutions and the Society: Governance

In societies where social institutions limit the rights of women, and where women's place is restricted to the private sphere, they have no or less say in the public and political domain. What is the impact of social institutions related to gender inequality on governance? We use Kaufmann et al. (2008, p. 7)'s definition of governance "as 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."

There are at least two approaches that allow to link social institutions with governance. First, there exist psychological and sociological explanations that state that women are less egoistic than men. Women are more risk-averse, they tend to follow the rules and they are more community-oriented than men (Dollar et al., 2001; Swamy et al., 2001). Countries in which women have more power will have a political system that is more rule oriented, responsive and accountable. Second, women's movements, being the answer to the exclusion of women from power, play an important role in increasing the quality of political systems by challenging e.g. personal rule (Waylen, 1993; Tripp, 2001). This argumentation suggests that countries with social institutions that hinder women to organize and to express their interests might lack an important oppositional force and therefore have a bad quality of governance.

- *Hypothesis 4: Social institutions related to high gender inequality inhibit the building blocks of good governance. In societies with social institutions favoring gender inequality political systems will be less responsive and less open to the citizens, so that voice and accountability will be reduced.*

- *Hypothesis 5: Social institutions related to high gender inequality inhibit the building blocks of good governance. In societies with social institutions favoring gender inequality there might be more personal rule in the political system as well as inequality in justice and legal systems, so that the rule of law will be weakened.*

3.4   Data

Our investigation uses macro-data at the country level. Table 5.29 in Appendix 3 gives an overview over the variables used for our estimations, the definitions and the data sources. Descriptive statistics of the variables used are presented in Table 5.26 in Appendix 3. As main regressors we use the SIGI and its five subindices *Family code*, *Civil liberties*, *Physical integrity*, *Son Preference* and *Ownership rights* in our estimations to check their explanatory value for the development outcomes female education, fertility, child mortality and governance.

First, we are interested in the impact of social institutions on female education, fertility and child mortality. As dependent variables we use *total fertility rates* from World Bank (2009) and *child mortality rates* from World Bank (2008). To measure education we choose *female gross secondary school enrollment rates* because this enables important functionings and empowers women. Furthermore, we assume that parents take into account that basic education of both boys and girls is necessary for fulfilling tasks related to the household. Data for secondary school enrollment are from World Bank (2009).

Second, we want to estimate the association between governance and our social institutions measures. We use the Governance Indicators developed by Kaufmann et al. (2008) and choose two of them to capture equality before the law, justice, tolerance and security as well as responsiveness, political openness and accountability in the political system. The *rule of law* index measures the extent to which contracts are enforced and property rights are ensured and the extent to which people trust in the state and respect the rules of the society. The *voice and accountability* index proxies civil and political liberties like freedom of expression, freedom of association, free media and the extent of active and passive political participation of citizens.

In all regressions we control for the level of economic development, religion, region and the political system in a country. The specific variables we use are:

- the log of per capita GDP in constant prices (US\$, PPP, base year: 2005) to control for the level of economic development (*log GDP*);

- a *Muslim* majority and a *Christian* majority dummy to control for the impact of religion, the left-out category being countries that have neither a majority of Muslim nor a majority of Christian population (*Muslim* and *Christian*);

- region dummies to capture geography and other unexplained heterogeneity that might go together with region, the left-out category being Sub-Saharan Africa (*SA* for South

Asia, *ECA* for Europe and Central Asia, *LAC* for Latin America and Caribbean, *EAP* for East Asia and Pacific);

- two political institutions variables, the *electoral democracy* variable and the *civil liberties index* from Freedom House (2008) that together measure liberal democracy which is assumed to be related to responsiveness to the needs of the public, political openness and tolerance in a country.[15]

We use different additional control variables in each regression following suggestions in the literature. In the fertility and child mortality regressions, we additionally control for

- *female literacy rates* to measure the ability of women to control their reproductive behavior, to care for themselves and their children (e.g. Basu, 2002; Hatt and Waters, 2006);

- a dummy proxying for high *HIV/AIDS* prevalence rates to control for extreme health problems especially in Sub-Saharan Africa due to AIDS (e.g. Foster and Williamson, 2000).

The Governance regressions exclude as control variables the civil liberties index from Freedom House as this index is used to build the voice and accountability index that we choose as dependent variable. We keep the electoral democracy variable because it does not pose a problem. We additionally include as control variables

- the share of *literate adult population* to control for the population's ability to be informed, to express their needs and to hold politicians' accountable (Keefer and Khemani, 2005);

- *ethnic fractionalization* as it might disturb governance through identity politics, patronage and distribution conflicts (e.g. Collier, 2001; Tripp, 2001);

- a measure of trade *openness* as openness increases the incentives to build 'good' institutions to attract trading partners, to join trading agreements etc. (e.g. Al-Marhubi, 2005).

Social institutions, i.e. normative frameworks, change only slowly and incrementally. As the social institutions indicators are not expected to change much over time we have to decide which year or time span should be covered by the other variables. For our response variables

---

[15]We multiply the civil liberties index by -1 to facilitate interpretation.

we choose to take the average of the existing values over five or six years (2000-2005, 2001-2005). For the control variables we take the averages of the existing values over ten years (1996-2005).[16] The averages provide information that is more stable than using a particular year. Using a longer time span for the control variables than for the response variables allows to capture possible time delays until effects can be observed. Nevertheless, we acknowledge that the choice of the time spans is arbitrary.

3.5 Empirical Estimation and Results

3.5.1 Empirical Estimation

We empirically test with linear regressions whether the composite measures reflecting social institutions related to gender inequality s_i are associated with each of the response variables y_i, representing the chosen development outcomes. We estimate regressions in the form

$$y_i = \alpha + \beta s_i + \text{control variables}_i + \varepsilon_i \tag{3.1}$$

using information at the country level. We are mainly interested in testing the null hypothesis that the coefficient β is zero at a statistical significance level of $\alpha = 5\%$. If the null hypothesis is rejected, it is reasonable to infer that the measure proxying social institutions related to gender inequality does matter for the given response variable, as predicted in the hypotheses from sections 3.2 and 3.3.

The general procedure used for each of the response variables consists of two steps. First, we start examining the effect of SIGI. We begin our estimation with a simple linear regression with SIGI as the only regressor s_i. We then run a multiple linear regression adding the main group of control variables that consists of the level of economic development, region dummies, religion dummies and the political system variables. If SIGI is significant in this regression, we continue and, if applicable, estimate the complete model with all identified control variables to confirm whether SIGI remains significant.

As SIGI is a rather broad measure to rank and compare countries and policy implications are difficult to derive from it, in a second step we focus on the subindices to get a more precise idea about what kind of social institutions might be related to the chosen development outcomes. We estimate the same multiple linear regression(s) described above using the five subindices s_i one at a time instead of SIGI to explore which dimension of social institutions related to gender inequality seems to be the most relevant. In the corresponding regression

---

[16]The ethnic fractionalization variable is constant over time as changes in the ethnic composition of a country at least over 20 and 30 years are rare (Alesina et al., 2003).

tables we only report the specification with the subindex or subindices that are statistically significant. It must be noted that we keep and show even those control variables that are not statistically significant in the regression, as we want to stress that the social institutions indices are associated with the development outcomes even if we include these control variables.

All regressions are estimated with Ordinary Least Squares (OLS). Regression diagnostics not reported here suggest that heteroscedasticity is a possible issue in our data and that there are influential observations that could drive our results. Concerning the first issue, it is known that if the model is well specified, the OLS estimator of the regression parameters remains unbiased in the presence of heteroscedasticity, but the estimator of the covariance matrix of the parameter estimates can be biased and inconsistent making inference about the estimated regression parameters problematic. Violations of homoscedasticity can lead to hypothesis tests that are not valid and confidence intervals that are either too narrow or too wide. To deal with heteroscedasticity, we use 'heteroscedasticity-consistent' (HC) standard errors. This means that while the parameters are still estimated with OLS, alternative methods of estimating the standard errors that do not assume homoscedasticity are applied. As the samples we use contain less than 150 observations, we use HC3 robust standard errors proposed by Davidson and MacKinnon (1993), which are better in the case of small samples. These are the standard errors that are presented in the regression Tables 3.1-3.5. Simulation studies by Long and Ervin (2000) have shown that HC standard error estimates tend to maintain test size closer to the nominal alpha level in the presence of heteroscedasticity than OLS standard error estimates that assume homoscedasticity. These authors recommend the use of HC3 robust standard errors, especially for sample sizes less than 250, as they can keep the test size at the nominal level regardless of the presence or absence of heteroscedasticity, with only a minor loss of power associated when the errors are indeed homoscedastic.[17]

In addition to this, we also use bootstrap with 1000 replications to compute a Bias-corrected and accelerated (BCa) 95% confidence interval of the regression coefficients computed with OLS (Efron and Tibshirani, 1993). One of the main advantages of bootstrapping methods is that no assumptions about the sampling distribution or about the statistic are needed. The results are not reported here, but are available upon request, and confirm that all the coefficients that are significant at the 5% level in Tables 3.1-3.5 remain significant when using Bca 95% confidence intervals around them.

---

[17]Certainly, heteroscedasticity-consistent standard errors are not a panacea for inferential problems under heteroscedasticity. As pointed out by some authors, there are limitations and trade-offs in these estimators (e.g. Kauermann and Carroll, 2001; Wilcox, 2001).

To deal with the second issue and check whether influential observations drive the results, we take the estimates of a regression obtained with OLS with standard variance estimator to detect the observations with unusual influence or leverage based on Cook's distance. Cook's distance is a commonly used estimate of the influence of a data point when doing least squares regression. We exclude countries from the sample if the value of Cook's distance is larger than $4/n$, with n being the number of observations, and re-estimate each regression on the re-stricted sample with HC3 robust standard errors. In all the cases we confirm that even after we exclude influential observations, the results remain basically unchanged.[18] The regressions are not reported here, but are available upon request.

We consider that the model specification is reasonable. However, possible endogeneity of our main regressors s_i (the SIGI and its subindices) should be taken into account when interpreting the coefficients of s_i as they would be biased and inconsistent in this case. Endogeneity is given if s_i is correlated with the disturbance ε_i in equation 3.1. There are three sources of endogeneity: omitted variables, measurement error and simultaneity (Wooldridge, 2002). We have included control variables to minimize omitted variable bias, although it is impossible to completely rule out this problem. Concerning measurement error, we regard the SIGI and the subindices as adequate proxies of social institutions related to gender inequality. It is not very plausible that there are errors in measurement that are related to the unobserved social institutions. The last source, simultaneity, arises when s_i is determined simultaneously with y_i. We consider that social institutions related to gender inequality s_i are relatively stable and long-lasting. Therefore, we think it is unlikely that the response variables y_i influence s_i.[19]

3.5.2   Results

Before we run the regressions it is necessary to check first the correlation between the subindices to rule out redundancy, and secondly between the subindices and the control variables to check whether the social institutions indices are proxies for these control variables. The Pearson correlation coefficient between the subindices is always positive, but not always significant. The correlation coefficients are always lower than 0.6, with the exception of the

---

[18]As an alternative procedure we use robust regression with iteratively reweighted least squares as described in Hamilton (1992), and confirm that results are similar.

[19]Social institutions are hard to measure. Therefore, sometimes one has to rely on legal indicators to proxy them, although we acknowledge that this could pose problems as there is for example an international mechanism, the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW), that aims at changing social institutions through legal measures. However, the impact of CEDAW on national legislation depends on the willingness of governments to sign and ratify it without reservation and on its willingness and ability to enact the new laws. Given the constituting function of social institutions for a society this could be difficult and depends on many factors.

correlation between the subindices *Family Code* and *Ownership rights*, which is equal to 0.74 (Table 5.27).[20] Table 5.28 shows that the absolute value of the Pearson correlation coefficient between the social institutions indicators and the control variables is always lower than 0.6, except for the SIGI and the subindices *Family code* and *Ownership rights* and the two variables capturing literacy of the whole population and of the female population.

Table 3.1: Linear Regressions with Dependent Variable Female Secondary Schooling

Specification with SIGI	(1)	(2)	Specification with Subindex	(3)
SIGI	**-141.77*****	**-10.91**	**Subindex family code**	**-39.10****
	(37.31)	(36.37)		(11.64)
log GDP		12.69***	log GDP	11.46***
		(3.39)		(2.61)
Muslim		-2.21	Muslim	3.43
		(5.47)		(4.84)
Christian		5.31	Christian	4.18
		(5.48)		(4.33)
SA		16.05	SA	12.3
		(8.75)		(8.44)
ECA		40.26***	ECA	28.25***
		(8.98)		(6.95)
LAC		18.33*	LAC	8.64
		(9.07)		(7.41)
MENA		33.86**	MENA	29.67**
		(12.50)		(9.69)
EAP		24.73**	EAP	14.36*
		(8.26)		(6.53)
Electoral democracy		8.11	Electoral democracy	6.19
		(7.67)		(6.84)
FH civil liberties		1.95	FH civil liberties	2.72
		(3.56)		(2.89)
constant	74.75***	-56.71	constant	-27.87
	(4.12)	(37.27)		(30.56)
Number of obs.	94	91	Number of obs.	99
Adj. R-Square	0.28	0.75	Adj. R-Square	0.78
Prob>F	0.0003	0.0000	Prob>F	0.0000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
HC3 robust standard error in brackets.
Regression (2) and (3) with controls for economic development, geography, religion and political system. In this case, this specification corresponds to the complete specification.

Regression results using *female secondary education* as dependent variable are presented in Table 3.1. Regression (1) with SIGI as the only regressor yields a negative and statistically significant association. Higher levels of inequality are associated with lower levels of female secondary education. The association vanishes in regression (2) if one includes the level of economic development, religion, region and the political system as control variables. Using

---

[20]Table 2.2 of Essay 2 shows Kendall Tau b between the five subindices and confirms that they are positively correlated, albeit not perfectly.

the subindex *Family code* instead of SIGI as the main regressor in regression (3) shows a different picture. The subindex is statistically significant even if the control variables are included. The adjusted coefficient of determination R^2 is 0.78. Hence, we find no evidence against Hypothesis 1 that states that social institutions related to high gender inequality are negatively associated with female education.[21]

Table 3.2: Linear Regressions with Dependent Variable Fertility

Specification with SIGI	(1)	(2)	Specification with Subindex	(3)	(4)
SIGI	**8.25***	1.73	**Subindex family code**	**1.89****	**2.03****
	(2.31)	(2.61)		(0.70)	(0.70)
log GDP		-0.71***	log GDP	-0.60***	-0.43***
		(0.16)		(0.12)	(0.12)
Muslim		0.52	Muslim	0.34	0.18
		(0.27)		(0.27)	(0.27)
Christian		0.25	Christian	0.24	0.46
		(0.26)		(0.25)	(0.26)
SA		-1.89***	SA	-1.73***	-1.88***
		(0.37)		(0.41)	(0.38)
ECA		-2.44***	ECA	-2.08***	-1.59***
		(0.48)		(0.38)	(0.43)
LAC		-0.96*	LAC	-0.68	-0.57
		(0.47)		(0.36)	(0.40)
MENA		-1.42*	MENA	-1.07*	-1.23*
		(0.63)		(0.50)	(0.48)
EAP		-1.74***	EAP	-1.37***	-1.20**
		(0.42)		(0.39)	(0.38)
Electoral democracy		-0.2	Electoral democracy	0.02	-0.03
		(0.31)		(0.29)	(0.30)
FH civil liberties		-0.02	FH civil liberties	-0.11	-0.14
		(0.17)		(0.13)	(0.13)
			Literacy female		-1.62**
					(0.60)
			Aids		-0.51
					(0.30)
constant	2.55***	9.76***	constant	7.89***	7.47***
	(0.25)	(1.82)		(1.30)	(1.29)
Number of obs.	100	97	Number of obs.	106	99
Adj. R-Square	0.31	0.82	Adj. R-Square	0.80	0.84
Prob>F	0.0006	0.0000	Prob>F	0.0000	0.0000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
HC3 robust standard error in brackets.
Regression (2) and (3) with minimum of controls for economic development, geography, religion and political system. Regression (4) with complete specification for fertility.

Results obtained using *total fertility rate* and *child mortality* as response variables are shown in Tables 3.2 and 3.3. In both cases, the simple linear regression (1) using SIGI as the

---

[21]Regressions not reported here, but available upon request, using primary gross completion rates obtained from World Bank (2008) instead of female secondary schooling as the dependent variable yield similar results.

only regressor shows a positive and significant statistical association between SIGI and the response variable. Higher levels of inequality are associated with higher levels of fertility and with higher levels of child mortality. However, once control variables related to the level of economic development, religion, region and the political system in a country are included in regression (2), SIGI is not longer statistically significant. This is not the case when we use the subindex *Family code* as the main regressor, as it is significant in regression (3) which uses the same control variables, and even in regression (4) which adds two additional regressors: the share of literate adult female population and a dummy reflecting high adult HIV/AIDS prevalence. In regression (4) the obtained adjusted R^2 is 0.84 for fertility and 0.82 for child mortality. Hence, we cannot reject Hypotheses 2 and 3, suggesting that social institutions related to high gender inequality are associated with higher fertility levels and higher child mortality.[22] As the subindex *Family code* is the relevant social institutions measure in our empirical estimations it seems that social institutions that deprive women of their autonomy and bargaining power in the family and that might restrict women's possibilities outside the family do matter for female education, fertility and child mortality.

Table 3.4 shows the results obtained for the dependent variable *voice and accountability*. Regression (1) with SIGI as the only regressor shows a negative and statistically significant association: higher levels of gender inequality are associated with lower levels of voice and accountability. This association remains significant in regression (2) where we add the level of economic development, religion, region and the political system[23] as control variables, and in the complete specification shown in regression (3) where we additionally include the proportion of seats held by women in national parliaments, the literacy rate of the population, a measure of openness of the economy, and a measure of ethnic fractionalization. In regression (3), we obtain an adjusted R^2 of 0.69. We explore which dimension of social institutions related to gender inequality is behind this result and find that it is the subindex *Civil liberties*. The specifications with the subindex *Civil liberties* in regressions (4) and (5) show that this subindex is negatively associated with voice and accountability and that this association is statistically significant even with the control variables. In regression (5) the adjusted R^2 is 0.69. Hypothesis 4 cannot be rejected with this evidence suggesting that social institutions related to gender inequality inhibit the building blocks of good governance in the form of voice and accountability. The subindex *Civil liberties* is the relevant social institutions measure in our empirical estimations. The freedom of women to participate in public life seems

---

[22]Regressions not shown here, but available upon request, confirm the results concerning mortality rates when infant mortality rates taken from World Bank (2008) are used instead of child mortality rates.

[23]Recall that in the governance regressions we only include the electoral democracy variable of Freedom House (2008) as the civil liberties index is included in the chosen governance indicators which are now the response variables.

to increase the quality of governance of a society. Relating back to theory, this could be due to the behavior of women as they tend to be more socially oriented than men and are a group that cross-cuts cleavages in general.

Table 3.3: Linear Regressions with Dependent Variable Child Mortality

Specification with SIGI	(1)	(2)	Specification with Subindex	(3)	(4)
SIGI	318.56**	50.42	**Subindex family code**	80.14**	77.23*
	(108.81)	(150.58)		(25.85)	(31.50)
log GDP		-22.55**	log GDP	-20.24***	-13.82**
		(7.35)		(5.34)	(5.09)
Muslim		26.61	Muslim	14.23	5.74
		(14.13)		(13.13)	(14.50)
Christian		7.49	Christian	9.47	14.27
		(11.72)		(10.31)	(10.81)
SA		-68.33***	SA	-61.30***	-71.03***
		(18.87)		(17.05)	(16.33)
ECA		-85.65***	ECA	-66.13***	-53.16*
		(23.82)		(16.75)	(20.65)
LAC		-66.65**	LAC	-50.69***	-50.23**
		(23.84)		(14.88)	(18.89)
MENA		-97.73***	MENA	-86.25***	-93.71***
		(26.90)		(21.71)	(23.48)
EAP		-73.44***	EAP	-59.37***	-55.65**
		(17.23)		(15.02)	(17.85)
Electoral democracy		-0.79	Electoral democracy	7.05	1.75
		(15.86)		(15.96)	(14.80)
FH civil liberties		-4.54	FH civil liberties	-8.33	-8.32
		(7.86)		(6.65)	(6.44)
			Literacy female		-62.77**
					(21.39)
			Aids		-19.02
					(14.56)
constant	43.38***	272.39**	constant	209.47**	209.34**
	(10.80)	(93.09)		(66.26)	(63.27)
Number of obs.	99	97	Number of obs.	106	99
Adj. R-Square	0.28	0.79	Adj. R-Square	0.79	0.82
Prob>F	0.0043	0.0000	Prob>F	0.0000	0.0000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
HC3 robust standard error in brackets.
Regression (2) and (3) with controls for economic development, geography, religion and
political system. Regression (4) with complete specification for child mortality.

Table 3.4: Linear Regressions with Dependent Variable Voice and Accountability

	Specification with SIGI			Specification with Subindex	
	(1)	(2)	(3)	(4)	(5)
SIGI / **Subindex civil liberties**	**-2.60***	**-1.42***	**-1.59***	**-0.61***	**-0.65***
	(0.50)	(0.48)	(0.54)	(0.23)	(0.23)
log GDP		0.27***	0.30***	0.31**	0.27***
		(0.06)	(0.06)	(0.05)	(0.06)
Muslim		0.18	0.15	0.16	0.21
		(0.13)	(0.14)	(0.13)	(0.14)
Christian		-0.03	-0.04	-0.05	-0.08
		(0.12)	(0.13)	(0.12)	(0.12)
SA		-0.27	-0.28	-0.12	-0.04
		(0.20)	(0.21)	(0.18)	(0.20)
ECA		-0.64***	-0.56*	-0.52***	-0.57**
		(0.14)	(0.22)	(0.13)	(0.22)
LAC		-0.40*	-0.41*	-0.32*	-0.31
		(0.17)	(0.18)	(0.15)	(0.16)
MENA		-0.45	-0.47	-0.27	-0.23
		(0.23)	(0.25)	(0.19)	(0.24)
EAP		-0.30*	-0.21	-0.14	-0.21
		(0.14)	(0.21)	(0.13)	(0.18)
Electoral democracy		1.10**	1.07***	1.13***	1.14***
		(0.12)	(0.11)	(0.10)	(0.10)
Parliament			0.01		0.01
			(0.01)		(0.01)
Literacy population			-0.31		0.24
			(0.42)		(0.37)
Openness			-0.07		0.23
			(0.36)		(0.22)
Ethnic			-0.07		0.01
			(0.25)		(0.23)
constant	-0.23*	-2.80***	-2.77***	-3.28***	-3.37***
	(0.10)	(0.45)	(0.47)	(0.41)	(0.39)
Number of obs.	102	97	95	112	108
Adj. R-Square	0.18	0.69	0.69	0.68	0.69
Prob>F	0.0000	0.0000	0.0000	0.0000	0.0000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
HC3 robust standard error in brackets.
Regression (2) and (4) with controls for economic development, geography, religion and political system.
Regressions (3) and (5) with complete specification for governance/voice and accountability.

Table 3.5: Linear Regressions with Dependent Variable Rule of Law

	Specification with SIGI			Specification with Subindices			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	SIGI			Subindex ownership		Subindex civil	
SIGI / Subindex ownership / Subindex civil	-1.73***	-1.88***	-1.33*	-0.89***	-0.71***	-0.75**	-0.63*
	(0.49)	(0.53)	(0.60)	(0.20)	(0.23)	(0.24)	(0.25)
log GDP		0.41***	0.36***	0.37***	0.30***	0.47***	0.36***
		(0.08)	(0.07)	(0.08)	(0.07)	(0.08)	(0.07)
Muslim		0	-0.04	-0.03	-0.02	0.04	0.11
		(0.16)	(0.16)	(0.13)	(0.14)	(0.14)	(0.14)
Christian		-0.18	-0.18	-0.11	-0.14	-0.22	-0.22
		(0.15)	(0.14)	(0.14)	(0.13)	(0.14)	(0.13)
SA		0.18	0.26	0.11	0.21	0.37	0.44
		(0.22)	(0.24)	(0.17)	(0.20)	(0.22)	(0.26)
ECA		-0.84***	-0.67*	-0.93***	-0.83***	-0.71***	-0.74**
		(0.18)	(0.27)	(0.16)	(0.22)	(0.15)	(0.22)
LAC		-0.74***	-0.54*	-0.78***	-0.61**	-0.58***	-0.51**
		(0.19)	(0.21)	(0.19)	(0.19)	(0.17)	(0.18)
MENA		-0.14	0.17	-0.09	0.18	0.10	0.30
		(0.27)	(0.32)	(0.25)	(0.29)	(0.24)	(0.28)
EAP		-0.31	-0.28	-0.35*	-0.36	-0.12	-0.23
		(0.16)	(0.23)	(0.15)	(0.20)	(0.15)	(0.20)
Electoral democracy		0.33*	0.40**	0.38**	0.44***	0.38**	0.46***
		(0.14)	(0.13)	(0.11)	(0.11)	(0.13)	(0.12)
Parliament			0.01		0.01		0.01
			(0.01)		(0.01)		(0.01)
Literacy population			-0.29		-0.03		0.20
			(0.42)		(0.38)		(0.36)
Openness			0.69**		0.71**		0.73**
			(0.33)		(0.27)		(0.23)
Ethnic			-0.07		-0.12		-0.13
			(0.32)		(0.28)		(0.27)
constant	-0.35***	-3.37***	-3.32***	-3.06***	-2.94***	-4.05***	-3.83***
	(0.10)	(0.58)	(0.52)	(0.56)	(0.53)	(0.52)	(0.46)
Number of obs.	102	97	95	112	108	112	108
Adj. R-Square	0.09	0.49	0.51	0.53	0.56	0.52	0.56
Prob>F	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

$* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$
HC3 robust standard error in brackets.
Regression (2), (4) and (6) with controls for economic development, geography, religion and political system.
Regressions (3), (5) and (7) with complete specification for governance/rule of law.

Results for the other component of governance, *rule of law*, are shown in Table 3.5, providing evidence for Hypothesis 5. Regression (1) shows a negative and statistically significant association between SIGI and rule of law: higher levels of inequality are associated with lower levels of rule of law. This association remains significant in regression (2) where we add the level of economic development, religion, region and the political system as control variables, and in the complete specification in regression (3) where we additionally include the proportion of seats held by women in national parliaments, the literacy rate of the population, a measure of openness of the economy, and a measure of ethnic fractionalization. In this last regression, we obtain an adjusted R^2 of 0.51. Again, we are interested in exploring which dimension of social institutions related to gender inequality is the relevant one for rule of law finding that two subindices matter: *Ownership rights* and *Civil liberties*.[24] The specifications with the subindices yield similar results to those of the SIGI and are presented in regressions (4) and (5) for *Ownership rights* and (6) and (7) for *Civil liberties*. For both subindices the adjusted R^2 obtained for the complete specification is 0.56. As postulated in Hypothesis 5, social institutions related to gender inequality seem to matter for governance inhibiting the rule of law, e.g. through personal rule and inequality in justice. Assuming that women's attitudes are different from those of men and that they challenge injustice, women's power in a society contributes to improve rule of law. The two subindices proxy where this power comes from, with *Ownership rights* measuring economic power through access to property and *Civil liberties* measuring the freedom to participate in and to shape public life.

A reasonable question is whether the social institutions indicators are capturing different religions. In the regressions reported here, we control for religion using a Christian and a Muslim dummy. As the results show, at least one subindex is significant when we control for religion. One could argue that what matters is how religion is practiced in the considered regions, and that the SIGI and the subindices might capture regional practice of religion. Therefore, we re-estimate all regressions including interactions between the religion and region dummies. The results for the SIGI and the subindices remain unchanged suggesting that they capture something different than religion and the regional practice of it.[25]

3.6   Conclusion

This study presents several answers to the question why we should care about social institutions related to gender inequality beyond the intrinsic value of gender equality. We derive hypotheses from existing theories and empirically test them with linear regression at the

---

[24]As shown in Table 5.27 the Pearson Correlation coefficient between the subindices Ownership rights and Civil liberties is 0.36.

[25]The results are available upon request.

cross-country level using the newly created Social Institutions and Gender Index (SIGI) and its subindices. Our results show that social institutions related to gender inequality are associated with lower female secondary education, higher fertility rates, higher child mortality and lower levels of governance measured as voice and accountability and rule of law. We find that apart from geography, political system, the level of economic development and religion, one has to consider social institutions related to gender inequality to better account for differences in important development outcomes.

The empirical estimation follows a two-step procedure for each outcome measure. First, the focus is to examine the explanatory value of the SIGI. In the specifications including all control variables, the SIGI is significant in the regressions for the measures of governance like voice and accountability and rule of law. If one interprets the SIGI as a summary measure of lack of power of women in all spheres of society then it seems that when women have more power, governance is better.[26] In the case of female secondary schooling, fertility rate and child mortality the SIGI turns out to be insignificant in the complete specifications.

Secondly, as the SIGI is a broad measure of social institutions related to gender inequality, we investigate which particular dimension of social institutions is significantly related to the chosen development outcomes, using the complete specifications. The subindex *Family code* is negatively associated with female education and positively with fertility and child mortality. These results suggest that social institutions that deprive women of their autonomy and bargaining power in the family do matter for female education, fertility and child mortality. The subindex *Civil liberties* is the dimension of social institutions that is significantly related to the governance component voice and accountability. The freedom of women to participate in public life seems to increase the quality of governance of a society as women tend to be more socially oriented than men and are a group that cross-cuts cleavages in general. The rule of law component of governance is negatively related to the subindices *Civil liberties* and *Ownership rights*. The two subindices proxy where this power comes from, with *Ownership rights* measuring access to property and *Civil liberties* measuring the freedom to participate in public life. Assuming that women's attitudes are different from those of men and that they challenge personal rule, women's power in a society is a relevant factor in increasing the rule of law.

Although the subindices *Family code*, *Ownership rights* and *Civil liberties* are the relevant dimensions of social institutions related to gender inequality for the response variables considered in this study, this does not mean that the other two subindices *Son preference* and *Physical integrity* are not important intrinsically or instrumentally for other outcomes.

---

[26]The association between two composite measures like the SIGI and the governance indicators has to be interpreted carefully.

Case studies investigating the mechanisms between social institutions and the outcome variables are necessary. Our study has the limitations of any cross-sectional regression analysis as we cannot rule out omitted variable bias. Causality can never be derived from regression analysis with cross-sectional data unless valid instruments are found. Concerning the results of the subindices, these should be considered exploratory and need to be confirmed with further research, which should also include the elaboration of appropriate theories linking social institutions related to gender inequality with each of the development outcomes used in this study.

Social institutions are long-lasting and deep-seated in people's minds. Changing them is a difficult task and requires approaches tailored to the particular needs and the socio-economic context (Jütting and Morrisson, 2005). The state can certainly help attenuate the effects of social institutions through specific policies. It may set incentives to counteract social institutions, e.g. in the form of laws to fight against discriminatory practices or through the implementation of programs favoring girls and women. Micro-credit programs or subsidies targeted at mothers are good examples here. Nevertheless, changing social institutions needs more than that. It needs a thorough understanding of the power relations in a country and people that are willing to become reform drivers and initiate learning processes that should be complemented by deliberation and public discussion at all levels of society. Be it through internal or external forces, women need help to empower themselves. That is what Sen calls 'agency of women' (Sen, 1999b).

Chapter 4

Reexamining the Link Between Gender and Corruption: The Role of Social Institutions[1]

4.1 Introduction

Is there a link between gender inequality and corruption in a society? The studies of Swamy et al. (2001) and Dollar et al. (2001) suggest that countries with greater representation of women in political and economic life tend to have lower levels of corruption. How can this relationship be explained?

This could be attributed to behavioral differences between men and women. As mentioned by Dollar et al. (2001), there are experimental studies and studies using survey data that find that, on average, women are less selfish and might have higher moral and ethical standards than men (e.g. Eagly and Crowley, 1986; Glover et al., 1997; Eckel and Grossman, 1998; Rivas, 2008).[2] If one accepts that women are less selfish and align their actions on higher moral standards than men, having women in important political and economic positions might lead to less corruption in a country.

An alternative explanation is put forward by Swamy et al. (2001), who argue that the negative relationship between women's participation and corruption could be due to self-selection. Only a few women reach powerful positions, and these women possibly gain access to these positions as they are from the "better" part of the women's distribution. >From a historical perspective, Goetz (2007) claims that it is gendered access to political positions that explains

---

[1]joint work with Boris Branisa

[2]There are empirical studies that challenge the finding that women are the "fairer sex" (e.g. Andreoni and Vesterlund, 2001; Alhassan-Alolo, 2007; Alatas et al., 2009). Another investigation highlights that when women are in a powerful position, they take decisions that are closely related to women's needs (Chattopadhyay and Duflo, 2004).

why women seem to be less corrupt than men. Excluded from male patronage networks, women are restricted in their opportunities for corrupt behavior. As they are newcomers or only few in the political or business sphere, women lack familiarity with the rules of illicit exchange to their own benefit. They try to assert their position by acting honestly and trust-worthily. This all leads to fewer corrupt activities by women, but as time passes and more women get access to power this effect might vanish.

It can also be argued that the observed relationship between women's representation and corruption is spurious. Swamy et al. (2001) and Dollar et al. (2001) warn that even if one controls for other factors in the regression, the observed relationship at the cross-country level could be due to some unobserved variable which influences both female representation and corruption. For example, according to Sung (2003) it might be the political system in the form of liberal democratic institutions that influences both. Sung (2003) argues that institutions of *liberal democracy* increase women's participation in government through values like equality, pluralism, fairness and tolerance. Competitive elections, an independent judiciary and a free press, which are elementary to a liberal democratic system, guarantee transparency and hold government officials accountable, thereby reducing corruption. Therefore, the negative effect of women's representation in government on corruption is spurious and vanishes when one includes a measure of democracy in the regression, which is empirically confirmed by Sung (2003). Swamy et al. (2001) draw attention to the "level of discrimination against women" as another possible omitted variable that drives both female participation and corruption. They claim that in countries that are more corrupt there is more discrimination against women and argue that in countries where traditions and clientelism prevail, there is a preference for men in power.

In this paper, we focus on the effect of discrimination against women on corruption in a society as we have a new measure of society's attitude towards gender inequality to empiri-cally test this relationship. Swamy et al. (2001) do not explain how this relationship operates, but several studies deal with this issue in a direct or indirect way (Tripp, 2001; Inglehart et al., 2002; Rizzo et al., 2007). The authors of these studies claim that society's attitude towards women influences how a political system functions and that it affects the positions women take in this system. Assuming that the level of corruption depends on the functioning of the political system, one could argue that society's attitude towards gender inequality has an impact on corruption.

The study of Tripp (2001) focuses on women's movements as a countervailing force to prevailing practices of corruption in Eastern and Southern Africa.[3] Political reforms at the

---

[3]Waylen (1993) makes a similar point for Latin America.

beginning of the 1990s, including free and competitive elections, a multi-party system and freedom of expression and association were not enough to give women access to powerful positions and to curtail the practices of patronage and clientelism. Women could enter the system, but they were excluded from male-dominated networks and therefore from the benefits of clientelism. However, political reforms allowed the formation of social forces. The disadvantaged women organized in autonomous movements, which were broad-based, multi-ethnic and multi-religious. These movements crosscut cleavages and started to demand transparency and the removal of clientelistic networks.

A similar perspective is adopted by Inglehart et al. (2002) and Rizzo et al. (2007) who state that when a society favors gender equality, there is more tolerance in general, more personal freedom and individual autonomy. The absence of these values inhibits political reforms towards a democratic system. The study of Inglehart et al. (2002) finds that gender equality is the most important part of "self-expression values" appearing in post-industrial societies which directly contribute to both democratization and to a greater representation of women in politics. Focusing on Arab and non-Arab Muslim countries, Rizzo et al. (2007) shows that even if democratic political institutions like elections, political parties or checks and balances are put in place, gender inequality can prevent these institutions from functioning well.

We empirically test on a sample of developing countries the relationship between social institutions related to gender inequality and the level of corruption, and contribute to the literature discussed above. We focus on public corruption, which refers to the misuse of public office for private gain. It comprises grand corruption, which refers to activities of top officials and big companies, and petty corruption, which refers to the activities of people at the lower end of hierarchies (Pardo, 2004). To proxy society's attitude towards gender inequality or what Swamy et al. (2001) call "level of discrimination against women" we introduce social institutions related to gender inequality into the analysis. These are long-lasting norms, traditions and codes of conduct that shape gender roles and influence the opportunities of women and men in a society. As suggested by e.g. De Soysa and Jütting (2007) and Essay 3, these guiding principles of human behavior affect development outcomes and should not be neglected in the study of a society. We measure social institutions related to gender inequality with the subindex Civil liberties proposed in Essay 2, which is based on variables from the OECD Gender, Institutions and Development Database (Jütting et al., 2008). This subindex captures society's attitude with regard to gender roles based on the freedom of women to participate in social life.

Our aim is to investigate whether society's attitude towards gender inequality matters for corruption once one takes into account the representation of women in parliament and

business as well as the political system of a country. The hypothesis is that in a society where women's participation in social life is restricted, there is a higher level of corruption.

Even after controlling for democracy and political and economic participation of women, as well as for other factors, we find a robust and significant relationship between the subindex Civil liberties and the level of corruption. We show that social institutions related to gender inequality are an important factor for the study of corruption. In societies where women are deprived of their freedoms to participate in social life, corruption is higher. As should be clear from the various existing theories the exact causal mechanism behind this relationship is not obvious and it cannot be established in this study since we conduct a cross-sectional analysis. This implies that one needs to carefully investigate the context, as tackling corruption might require more than pushing democratic reforms and increasing female representation in political and economic positions. The rest of the paper is organized as follows. Section 4.2 describes the data used, the empirical estimation and the main results, which are discussed in Section 4.3.

4.2   Empirical Estimation and Results

4.2.1   Data

The definition of all variables and descriptive statistics are presented in Tables 5.29, 5.30 and 5.31 in Appendix 4. Measuring corruption is a complex task as it has many faces. There is public corruption, which refers to the misuse of public office for private gain, and corruption that comprises the collusion between firms or misuse of corporate assets (Svensson, 2005). Other authors differentiate between grand and petty corruption. Grand corruption refers to activities of top-officials and big companies. Petty corruption refers to the activities of people at the lower end of hierarchies (Pardo, 2004).

We use two different measures of public corruption in our estimations comprising grand and petty corruption. The first measure is the Corruption Perception Index (*CPI*) of Transparency International.[4] The CPI measures the perception of corruption in a country. It is based on various data sources, business surveys and expert panels about perceptions of corruption, and is a comprehensive measure that covers the different forms of grand and petty corruption in business, politics and administration. It is continuous and ranges from 0 meaning high corruption to 10 meaning low corruption (Lambsdorff, 2006).

The second indicator is the Corruption in Government Index from the International Country Risk Guide (*ICRG*) provided by the Political Risk Services.[5] The ICRG index assesses

---

[4]Data are available at http://www.transparency.org/policy_research/surveys_indices/cpi.
[5]Data are available at http://www.prsgroup.com/.

the political risk associated with corruption and focuses in particular on those types of corruption that lead to instability in the political system as they distort the economic and financial environment, put foreign investments into risk and reduce the efficiency of government and business because people come to power not because of their ability but through patronage and clientelistic practices.[6] Hence, this measure gives the extent of political risk of instability that is assumed to increase with corruption. Therefore, it is only under certain conditions an indicator of the level of corruption. Whether the political risk of instability caused by corruption coincides with the level of corruption depends on the degree of tolerance towards corruption (Lambsdorff, 2006). The ICRG corruption index goes from 0 to 6 with 0 meaning high risk and 6 indicating low risk. Pearson correlation coefficient between both corruption measures is significant and is 0.58 indicating that both measures seem to capture different aspects of corruption.

The subindex Civil liberties (*Subindex Civil lib.*) is one of five composite indices (the others being subindex Family code, subindex Son preference, subindex Physical integrity, subindex Ownership rights) that measure social institutions related to gender inequality (see Essay 2). These social institutions are conceived as long-lasting norms, traditions and codes-of conduct that find expression in traditions, customs and cultural practices, informal and formal laws and guide people's behavior and interaction. They shape gender roles and therefore the social and economic opportunities of men and women. We use the subindex Civil liberties in this study as it covers those social institutions that directly shape the opportunities of women to participate in social life. Hence, it reflects better their opportunities to gain power in politics and economics than the other subindices related to gender inequality. Indeed, we find that the subindex Civil liberties is the only subindex that is significant in the regression analysis. It is built out of two variables of the OECD Gender, Institutions and Development Database (Morrisson and Jütting, 2005; Jütting et al., 2008), which are freedom of movement and freedom of dress. The variables measure whether women are allowed to go outside the house and whether they are obliged to use a veil or burqa to cover parts of their body in public. Both variables are ordinal taking the values 0, 0.5 and 1 with 0 indicating no restrictions and 1 indicating high restrictions on women.[7] They are proxies of civil liberties in a sense that when women are restrained to leave the house it is difficult to imagine

---

[6]http://www.prsgroup.com/ICRG_Methodology.aspx#PolRiskRating

[7]The variable dress code takes the value 0 if there are less than 50% of women that are obliged to follow a certain dress code, 0.5 if there are more than 50% of women forced to follow a certain dress code and 1 if all women are obliged to follow a certain dress code, or if it is punishable by law not to follow it. The variable freedom of movement is 0 if there are no restrictions of women's movement outside the home, 0.5 if (some) women can leave home sometimes, but with restrictions, and 1 if women can never leave home without restrictions (i.e. they need a male companion, etc.)

that they can actively participate in social, political and economic life. Wearing a veil might be a form of self-determination and expression, and different traditions, styles and customs are connected to it. However, *forced* veiling is incompatible with agency, as it might be a sign of subordination in a society and might hinder interactions with other human beings - either as women cannot interact because they wear a veil or they can only interact if they wear a veil (Macdonald, 2006; Milallos, 2007). The subindex is the rescaled weighted sum of the two variables with the weights obtained from polychoric principal component analysis (Kolenikov and Angeles, 2009). The subindex goes from 0 (no gender inequality) to 1 (high gender inequality). As the subindex Civil liberties does not cover developed (OECD) countries, the subsequent empirical analysis focuses on developing countries. The list of countries covered by the subindex Civil liberties can be found in Table 5.32 in Appendix 4.

Table 4.1: Variation of the Subindex Civil Liberties Over Religion

Subindex civil liberties	No christian/ Muslim majority	Christian majority	Muslim majority	Total
0	22	46	15	83
0.298	5	8	1	14
0.301	1	0	4	5
0.599	1	0	15	16
0.781	0	0	2	2
0.818	0	0	1	1
1	0	0	2	2
Total	29	54	40	123

The variables that are contained in the subindex could be considered as proxies for religion and therefore one could think that the subindex Civil liberties might be a proxy for religion as well. When investigating the variation of the subindex over religion, one observes that there is more variation within Muslim majority countries than in countries with either Christian majority or countries without Christian or Muslim majority (Table 4.1).[8] To further examine whether the subindex measures Muslim religion, we plot the subindex Civil liberties against the percentage of Muslim population in a country (Figure 4.1). It is true that countries having less than 50% Muslim population tend to have lower values on the subindex Civil liberties with the exception of India which scores 0.6 with about 15% of Muslim population. For countries with more than 50% Muslim population the subindex shows more variation. Noticeably, there are several countries that have more than 70% of Muslim population and

---

[8]The variable freedom of movement varies over all three religious categories, while the variable freedom of dress has almost no variation in countries having a Christian majority or countries without Christian or Muslim majority, except for India and Sri Lanka.

Figure 4.1: Scatter Plot: Subindex Civil Liberties Against Percentage of Muslim Population
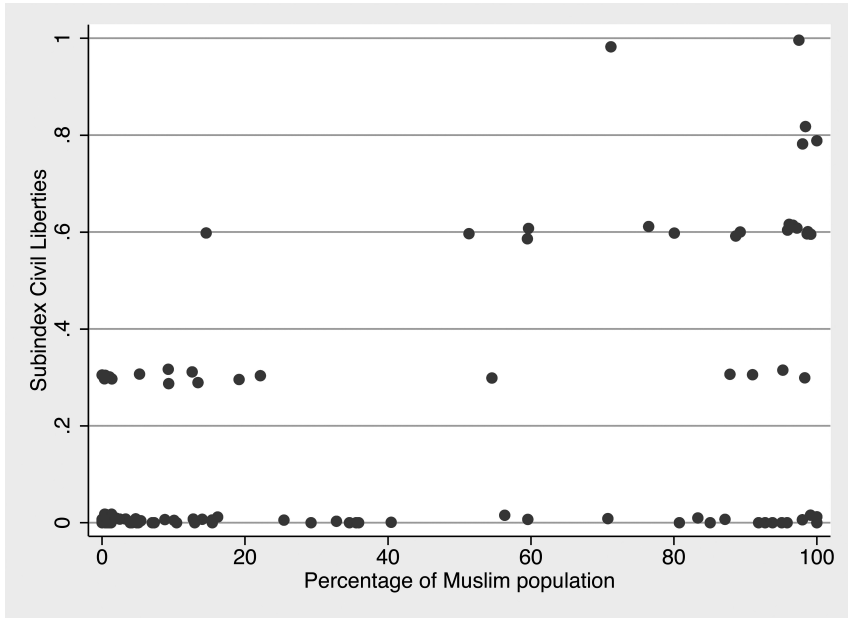
the value 0 on the subindex Civil liberties.[9] Consequently, there is no perfect correspondence between the subindex and the percentage of Muslim population. Nevertheless, in the regressions we include a Muslim and a Christian dummy (*Muslim* and *Christian*) to control for the impact of religion, the left-out category being countries that have neither a majority of Muslim nor a majority of Christian population.[10]

To account for female representation, which is highlighted by e.g. Swamy et al. (2001) and Dollar et al. (2001), we include three measures of female representation. We take data from World Bank (2009) on the proportion of female legislators (*Parliament*), the female

---

[9]Albania, Azerbaijan, Gambia, Guinea, Kyrgyz Republic, Mali, Morocco, Niger, Senegal, Sierra Leone, Tajikistan, Tunisia, Turkmenistan, Uzbekistan

[10]As Muslim religion is related to the subindex we also use the percentage of Muslim population instead of the two religion dummies in the regressions. The results are unchanged.

share in professional, technical, administrative and managerial positions (*Managers*),[11] and
women's share of labor force (*Labor force*).

To capture democracy we choose the Electoral Democracy index (*Electoral democ.*) of
Freedom House (2008) that takes the value 1 if there are competitive, universal, free and
secret elections and a multiparty system. An alternative measure is the Polity2 index of
the Polity IV Project that we use to check the robustness of the results as *Polity2* measures
more closely liberal democracy (Marshall and Jaggers, 2009).[12] Unfortunately, it covers
fewer countries than the Electoral democracy index.[13] Dollar et al. (2001), Swamy et al.
(2001) and Sung (2003) use either the Civil Liberties index[14], the Political Rights index or
the Freedom of the Press index of the Freedom House project as regressors in their empirical
analysis to measure or to refine the measurement of democracy. It needs to be stressed that
these measures are not without methodological problems as they include questions about
bribing and other forms of corrupt behavior and are therefore by construction correlated with
corruption. The Civil Liberties index includes questions on corruption that restrains free and
independent media. The Political Rights index includes questions related to corruption in
government. The Freedom of the Press index includes questions on the impact of corruption
and bribery on content of the press. Moreover, Sung (2003) uses a rule of law index that is
also problematic as rule of law is closely related to the prevalence of corruption. Therefore,
from all Freedom House measures only the Electoral Democracy index is included in our
regressions to account for democracy.

As additional controls we include:

- the log of GDP per capita in constant prices to control for the level of economic devel-
  opment as combatting corruption might be costly, and as poorer people might tend to
  engage more in corrupt activities (*log GDP*) [15] (Swamy et al., 2001);

---

[11]Both indicators have been criticized (Bardhan and Klasen, 1999; Dijkstra, 2002). In some countries, for
example communist ones, parliaments lack power and the representation of women in these parliaments does
not reflect actual power of women. Moreover, female representation in parliament measures representation
only at the national level and ignores women's participation at other levels of the state and in civil society. A
similar problem is attached to the representation of women in senior economic positions that measures only
formal sectors. In addition, this indicator does not fluctuate much over years. However, given that there is a
lack of data available for women's representation at the local and societal level as well as for informal economic
participation and to be comparable to other studies, we use both measures.

[12]Current data for the Polity IV Project can be found at
http://www.systemicpeace.org/polity/polity4.htm.

[13]We use averages over ten years to capture stability of democracy. For the 121 countries for which both
Electoral democracy and Polity2 are available, the Pearson Correlation Coefficient between them is 0.90 and
significant.

[14]The Civil liberties index from Freedom House (2008) measures civil liberties in general and is not to be
mixed up with the subindex Civil liberties related to gender inequality.

[15]US$, PPP, base year: 2005.

- region dummies to capture geography and other unexplained regional heterogeneity, with Sub-Saharan Africa as the reference category (*SA* for South Asia, *ECA* for Europe and Central Asia, *LAC* for Latin America and Caribbean, *EAP* for East Asia and Pacific);

- ethnic fractionalization as it might increase corruption through clientelistic networks, identity politics and patronage along ethnic lines (e.g. Tripp, 2001) (*Ethnic frac.*);

- literacy rates to control for the knowledge of the population about laws against corruption, and as higher education might come along with less tolerance towards corruption (Swamy et al., 2001) (*Literacy pop.*);

- a measure of trade openness as trade barriers increase the incentives for corrupt behavior between individuals and customs officials (Ades and Tella, 1997; Gatti, 2004) (*Openness*);

- a dummy indicating whether a country has never been a colony (*Not colony*) and a dummy measuring whether a country was a British colony (*British colony*) based on the Correlates of War 2 Project (2003) as corruption might also be linked to the history of colonialism (Swamy et al., 2001).

The subindex Civil liberties reflects the information available around the year 2000 and is not expected to change rapidly over time as social institutions are long-lasting and change only slowly and incrementally. For this reason, we use averages of the existing values over time in the case of all other variables to minimize the loss of observations due to missing values and to obtain a more stable value for the indicators used. For the corruption indicators representing our response variables we take averages over the years 2001 to 2005 for the CPI and in the case of the ICRG over the period 2000-2004. For the other regressors we use averages over ten years (1996-2005), with the exception of ethnic fractionalization as changes in the ethnic composition of a country in less than 20 years are rare (Alesina et al., 2003). Concerning the two democracy variables, choosing averages over ten years has the advantage of capturing the stability of a democratic system, which has been highlighted by Treisman (2007) as important for corruption. In addition, having a difference of five years between response variable and the regressors might help to alleviate endogeneity and capture delays until possible effects can be observed.

4.2.2 Empirical Estimation

We empirically test with multiple linear regressions whether the subindex Civil liberties s_i, which measures the freedom of social participation of women, is correlated with a response

variable y_i capturing the level of corruption, after controlling for other factors that have been described in the literature as possible determinants of corruption.[16] As was discussed previously, we consider that social institutions related to gender inequality are relatively stable and long lasting. Therefore, we assume that they do not depend on the response variable for the period considered.[17]

We run regressions as

$$y_i = \alpha + \beta s_i + \text{control variables}_i + \varepsilon_i \qquad (4.1)$$

using information at the country level. We are mainly interested in testing the null hypothesis that coefficient β is zero at a statistical significance level of 10%. The control variables included to attenuate omitted variable bias are described in Table 5.29 in the Appendix. We acknowledge, however, that it is impossible to entirely rule out this problem.

To reproduce the findings from the literature, we first run a regression without the subindex Civil liberties to focus on the effects of democracy and representation of women, which have been largely discussed. In a second step, we add to the regressions the subindex Civil liberties as a measure of society's attitude towards gender inequality, as it can be argued that it is a variable that has been omitted in the previous regressions (Swamy et al., 2001). We run each specification for the two measures of corruption and use each time one of the two alternative measures of democracy. At the end, we present four regressions for each corruption indicator.

Preliminary regressions not reported here suggest that heteroscedasticity is a possible issue in our data and that there are influential observations that could drive the results. If our model is well specified, the OLS estimator of the regression parameters remains unbiased in the presence of heteroscedasticity, but the estimator of the covariance matrix of the parameter estimates can be biased and inconsistent, making inference about the estimated regression parameters problematic. Violations of homoscedasticity can lead to hypothesis tests that are not valid and confidence intervals that are either too narrow or too wide. To deal with heteroscedasticity, we run the regressions with OLS and 'heteroscedasticity-consistent' (HC) standard errors. As our sample sizes are less than 150, we use HC3 robust standard errors proposed by Davidson and MacKinnon (1993), which are better with small samples.[18]

---

[16]Before conducting the multiple linear regression analysis, we account for the importance of GDP for corruption. We first run a simple linear regression of each corruption measure on log GDP. We then compute the estimated residuals from this regression and use them as the dependent variable in a new simple linear regression where the subindex Civil liberties is the only regressor. For both CPI and ICRG we obtain a negative and significant coefficient for the subindex Civil liberties which suggests that the subindex is able to account for something that goes beyond GDP when explaining corruption.

[17]In general, social institutions, i.e. normative frameworks, change only slowly and incrementally.

[18]Simulation studies by Long and Ervin (2000) have shown that HC standard error estimates tend to maintain test size closer to the nominal alpha level in the presence of heteroscedasticity than OLS standard error

For all the regressions, we check whether the results concerning the subindex civil liberties are stable in three ways. First, it is clear that in the multiple regressions, the estimate of the effect of our main variable, the subindex civil liberties, depends on the values of the other explanatory variables included (Mukherjee et al., 1998). We also try a simpler model to confirm that the estimated coefficient of the subindex civil liberties is negative and statistically significant. In this smaller model and based on the arguments presented before, we include as additional regressors the variables capturing the representation of women in society, a measure of democracy, the log GDP, religion dummies and regional dummies. This has the advantage that less parameters have to be estimated with the available observations.

Secondly, we use bootstrap with 1000 replications to compute a Bias-corrected and accelerated (Bca) 90% confidence interval of the regression coefficients computed with OLS to confirm that the value zero is not contained in the confidence interval around β (Efron and Tibshirani, 1993). One of the main advantages of bootstrapping methods is that one does not make any assumptions about the sampling distribution or about the statistic. Third, we detect observations with high influence or leverage based on the first estimates (OLS with standard variance estimator) using Cook's distance. Cook's distance is a commonly used estimate of the influence of a data point when doing least squares regression, and it measures the effect of deleting a given observation. We exclude the countries identified as outliers from the sample if the value of Cook's distance is larger than $4/n$, with n being the number of observations, and re-estimate equation 4.1 on the restricted sample using HC3 robust standard errors.

One should consider that possible endogeneity of the regressor s_i (the subindex Civil liberties), meaning that s_i is correlated with the error term ε_i in the regression, might lead to an estimated coefficient of s_i that is biased. Endogeneity might arise due to omitted variables, measurement error and simultaneity (Wooldridge, 2002). The control variables included in the regression aim at minimizing omitted variable bias, albeit one cannot rule out this problem. We do not find it plausible that there are measurement errors in s_i which are related to the unobserved 'true' social institutions. Simultaneity could arise if s_i is determined simultaneously with the dependent variable y_i. As was discussed previously, social institutions related to gender inequality s_i are relatively stable and long-lasting. Hence, it is unlikely that the response variable y_i influences s_i.

---

estimates that assume homoscedasticity. These authors recommend the use of HC3 robust standard errors, especially for sample sizes less than 250, as they can keep the test size at the nominal level regardless of the presence or absence of heteroscedasticity, with only a minor loss of power associated when the errors are indeed homoscedastic. We acknowledge that heteroscedasticity-consistent standard errors are not a panacea for inferential problems under heteroscedasticity. As pointed out by some authors, there are limitations and trade-offs in these estimators (e.g. Kauermann and Carroll, 2001; Wilcox, 2001).

4.2.3   Results

Results for the CPI as the first measure of corruption are presented in Table 4.2. Specifications (1) and (2) do not include the subindex Civil liberties. In both specifications, none of the democracy variables Electoral democracy and Polity2 are significant. >From the three measures of representation of women only Parliament is significant and positively related to corruption in specification (1) where Electoral democracy is the measure of democracy. Of the control variables only GDP has a significant and positive coefficient. In specifications (3) and (4) the subindex Civil liberties is added as a new regressor to the former specifications. Its coefficient is negative and significant in both. Both democracy variables as well as the measures for participation of women in the economy are not significant. Only Parliament carries a positive and significant coefficient when Electoral democracy is used (specification (3)). In the same specification (3) two control variables besides log GDP become significant: British colony and the regional dummy for ECA. For all four specifications the adjusted R square is around 0.5.

Table 4.3 shows the results when ICRG is used as the measure of corruption. For all 4 specifications (1-4), none of the variables reflecting representation of women and none of the democracy measures is significant. Interestingly, log GDP is also insignificant in all specifications, whereas it is always significant when the CPI is used as measure of corruption. Openness is the only control variable which is significant in all specifications. Important for the results of this paper, the subindex Civil liberties is significant in specifications (3) and (4), and adding it to the corresponding regressions yields values for adjusted R-square that are noticeably larger than without it. It must be noted, however, that the obtained values for adjusted R-square for the regressions with the ICRG are lower than for the CPI (between 0.2 and 0.3 for the ICRG and around 0.5 for the CPI), suggesting that the model is not able to explain much of the variation of the political risk of instability due to corruption.

Table 4.2: Linear Regressions With Dependent Variable CPI

	Model 1	Model 2	Model 3	Model 4
Representation of women				
Parliament	0.031*	0.033	0.032*	0.037
	(0.018)	(0.023)	(0.018)	(0.023)
Managers	0.025	0.022	0.011	0.006
	(0.029)	(0.032)	(0.031)	(0.034)
Labor force	0.007	0.009	0.001	0.004
	(0.009)	(0.010)	(0.010)	(0.011)
Democracy				
Electoral democ.	0.339		0.263	
	(0.234)		(0.231)	
Polity2		0.039		0.032
		(0.025)		(0.023)
Social inst. related to gender ineq.				
Subindex Civil lib.			-1.730***	-1.624*
			(0.593)	(0.866)
log GDP	0.710***	0.738***	0.766***	0.821***
	(0.197)	(0.212)	(0.193)	(0.209)
Muslim	-0.367	-0.271	0.049	0.107
	(0.319)	(0.394)	(0.305)	(0.363)
Christian	-0.392	-0.240	-0.280	-0.131
	(0.288)	(0.341)	(0.283)	(0.329)
Ethnic frac.	-0.334	-0.364	-0.267	-0.124
	(0.628)	(0.824)	(0.595)	(0.809)
Literacy pop.	-0.928	-1.122	-0.470	-0.831
	(1.070)	(1.193)	(1.009)	(1.091)
Openness	1.457	1.752	1.199	1.455
	(1.106)	(1.435)	(1.063)	(1.378)
Not colony	0.135	0.146	0.331	0.197
	(0.315)	(0.410)	(0.300)	(0.362)
British colony	0.478	0.313	0.611**	0.407
	(0.298)	(0.391)	(0.298)	(0.387)
constant	-3.305**	-3.455*	-3.364**	-3.809*
	(1.634)	(1.964)	(1.687)	(2.108)
Number of obs.	103	86	103	86
R2	0.576	0.580	0.613	0.607
Adjusted R2	0.491	0.474	0.530	0.501
Prob > F	0.000	0.000	0.000	0.000

HC3 robust standard errors in brackets.
Regional dummies included in all estimations.
$*p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.3: Linear Regressions With Dependent Variable ICRG

	Model 1	Model 2	Model 3	Model 4
Representation of women				
Parliament	0.015	0.012	0.016	0.016
	(0.018)	(0.020)	(0.014)	(0.017)
Managers	0.025	0.025	0.010	0.011
	(0.021)	(0.021)	(0.017)	(0.019)
Labor force	-0.003	-0.000	-0.009	-0.006
	(0.007)	(0.008)	(0.007)	(0.008)
Democracy				
Electoral democ.	0.273		0.221	
	(0.234)		(0.223)	
Polity2		0.029		0.027
		(0.025)		(0.025)
Social inst. related to gender ineq.				
Subindex Civil lib.			**-1.488***	**-1.260****
			(0.425)	**(0.604)**
log GDP	0.122	0.081	0.153	0.123
	(0.149)	(0.182)	(0.135)	(0.166)
Muslim	-0.337	-0.229	0.076	0.070
	(0.293)	(0.316)	(0.261)	(0.315)
Christian	-0.351	-0.321	-0.300	-0.289
	(0.272)	(0.338)	(0.257)	(0.333)
Ethnic frac.	0.507	0.349	0.655	0.652
	(0.427)	(0.465)	(0.410)	(0.496)
Literacy pop.	-0.165	0.118	0.404	0.436
	(0.930)	(0.988)	(0.769)	(0.873)
Openness	1.277**	1.523**	0.991*	1.274**
	(0.625)	(0.650)	(0.588)	(0.596)
Not colony	0.033	0.122	0.255	0.177
	(0.237)	(0.304)	(0.308)	(0.396)
British colony	-0.022	-0.055	0.131	0.067
	(0.228)	(0.289)	(0.210)	(0.293)
constant	0.474	0.529	0.461	0.351
	(1.082)	(1.193)	(0.924)	(1.094)
Number of obs.	86	72	86	72
R2	0.361	0.423	0.462	0.482
Adjusted R2	0.201	0.241	0.318	0.306
Prob > F	0.005	0.001	0.000	0.001

HC3 robust standard errors in brackets.
Regional dummies included in all estimations.
$*p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Using a simpler model does not change the results for the subindex Civil liberties and the variables measuring representation of women and democracy. These findings do also withstand the two other robustness checks. First, we confirm with Bias-corrected and accelerated (Bca) confidence intervals that in all cases the value zero is not contained in the 90% confidence interval around the regression coefficient of the subindex Civil liberties. Secondly, excluding outliers (6 to 7 countries) and re-running specifications (3) and (4) for both corruption measures, the subindex Civil liberties remains significant in all estimations. It is

worth mentioning that for every restricted sample, the adjusted R-square is higher than in the corresponding complete sample.[19]

Summarizing the results, when we do not include the subindex Civil liberties we find that from all variables for representation of women only Parliament is significant in the case of the CPI as long as Electoral democracy is used as measure of democracy. If one uses Polity2 instead, Parliament becomes insignificant. None of the democracy measures turns out to be significant. When we include the subindex Civil liberties, the results for representation of women and the democracy variables stay unchanged. Neither representation of women, except Parliament in the case of CPI when Electoral democracy is used, nor the democracy variables are significantly related to corruption. The main result concerning the subindex Civil liberties is that even after controlling for democracy and for measures of political and economic participation of women as well as for other factors, we find a robust and significant relationship between the subindex Civil liberties, which reflects society's attitude towards gender inequality, and the level of corruption. Social institutions favoring gender inequality are associated with higher levels of corruption.

4.3 Conclusion

The literature investigating the link between gender and corruption finds that there is a relationship between female representation in political and economic life and the level of corruption in a country. However, some studies warn that the observed relationship may be due to omitted variable bias. A possible variable that might influence both participation of women and corruption, is liberal democracy (e.g. Sung, 2003). We introduce a further omitted variable that has either been neglected in the literature or not been adequately dealt with because of insufficient data. Swamy et al. (2001) refer to this as the "level of discrimination against women" and proxy it with the gaps in educational attainment and life expectancy between men and women. We use the subindex civil liberties, which we consider a better proxy of the "level of discrimination against women" as it captures social institutions that restrict women in their freedom to participate in the public and reflect society's attitude towards gender inequality. The subindex measures underlying institutions and not outcomes of these institutions as do the variables used by Swamy et al. (2001).

When we replicate the findings of the literature for our sample of developing countries without the social institutions indicator, the results support the hypothesis of Sung (2003) and others that, when liberal democracy (in our case measured with Polity2) is considered in the regression, the representation of women in political and economic life is insignificant. However, Sung's hypothesis is weakened by the fact that there is no statistically significant

---

[19]Results for all the robustness checks are not reported here, but are available upon request.

association between democracy and corruption. Consequently, our statistical results support neither Sung's arguments nor the arguments put forward by Swamy et al. (2001) and Dollar et al. (2001) that representation of women is negatively related to corruption.[20] These results make it difficult to interpret social institutions related to gender inequality as an omitted variable when one investigates the relationship between representation of women in society, democracy and corruption.[21]

Once we include the subindex Civil liberties as a regressor, we find that after controlling for representation of women in political and economic life and for democracy, it has a robust negative and significant effect on corruption. Consequently, the main finding of this study is that in countries where social institutions inhibit the freedom of women to participate in social life, the level of corruption is higher.

Admittedly, one has to be cautious with these results. Interpretations for these findings in the light of the theories discussed are difficult, and country or regional studies are needed. Measurement is another relevant issue as the concepts of social institutions, democracy, participation of women and corruption are all hard to operationalize. Finally, it cannot be ruled out that another factor, which has been neglected from the analysis, shapes the results.

Nevertheless, we derive one policy implication from this study, which should be mainly targeted at developing countries. In a context where social institutions deprive women of the freedom to participate in social life, neither political reforms towards democracy nor the representation of women in political and economic positions might be enough to reduce corruption. How women are treated in a society is not only important for them, but has major implications for the functioning of the whole society.

---

[20]Once again, our sample includes only developing countries, while the other studies include developed countries as well.

[21]We have estimated with multivariate regressions, not reported here, whether there is (1) a relationship between democracy and the subindex Civil liberties and (2) a relationship between representation of women in society and the subindex Civil liberties in our sample of developing countries, but did not find significant results.

Chapter 5

Health Inequality in Bolivia: The Role of Indigenous Heterogeneity[1]

5.1   Introduction

Improving child health is a priority issue in achieving the Millennium Development Goals (MDGs). Goal 4 is exclusively dedicated to the reduction of child mortality. In developing countries, child health status is often linked to ethnic origin. Especially in Latin America, ethnic and racial divisions exist all over the continent and it is widely acknowledged that people of indigenous origin are still a socially disadvantaged group suffering more from marginalization, poverty and health problems than the non-indigenous population (Hall et al., 2006; Stephens et al., 2006). Although at the end of 2004 the General Assembly of the United Nations proclaimed a Second International Decade of the World's Indigenous People, starting in 2005, as a response to the problems that indigenous people face, the MDGs are too general and fail to incorporate the indigenous face of poverty and health problems (Telles, 2007).

Health inequity plays a major role for indigenous people (Braveman and Tarimo, 2002; Stephens et al., 2006). It refers to social inequality in health that arises because of social disadvantages associated with characteristics like gender, ethnicity, geographical location, economic, political resources, etc. Health inequity harms the affected people, and has a damaging effect on the welfare of a country as it contributes to the spread of diseases not only among the disadvantaged but also among the more privileged groups. Possible cost savings through preventive measures are not realized. Health inequity also means that the labor productivity of parts of the society decreases. In general, it is an impeding factor for development (Braveman and Tarimo, 2002). This is one reason why the World Health Organization has set

---

[1]joint work with Elena Gross

itself the target of making inequities in health visible so as to overcome existing inequalities (WHO, 2008).

This study analyzes inequality in child health between indigenous and non-indigenous people in Bolivia to explore whether and how indigeneity should be dealt with in order to achieve improvements in child health and attain the MDGs. We focus on Bolivia because the indigenous population constitutes more than half of the total. According to the last census (2001) there were about 3.9 million indigenous people living in Bolivia, which corresponds to 62% of the total population (Layton and Patrinos, 2006; Pozo et al., 2006).[2]

The existence of gaps in health between the indigenous and non-indigenous population of Bolivia are confirmed by the literature (e.g. UDAPE and OPS, 2004; Pozo et al., 2006; PAHO, 2007). Overall infant mortality, incidence of child diseases like measles and rubella, diarrheal diseases and malnourishment are higher for the indigenous groups; adult indigenous persons' health status is lagging behind the health status of the population with Spanish ancestors; and native people are disadvantaged in access to medical care (UDAPE and OPS, 2004; Pozo et al., 2006; PAHO, 2007). However, these studies mostly conduct a descriptive and bivariate analysis which has the shortcoming that other factors that might be related to both ethnic origin and health such as poverty, urban-rural differences and geographical location and other household-related characteristics are not taken into account.

If one wants to analyze indigeneity and health in Bolivia, it cannot be completely ruled out that native origin might be only a proxy for other characteristics like wealth, geographical setting, or living in an urban or rural area, which all have the potential to cause inequalities in health. In Bolivia, these characteristics are strongly associated with being indigenous (see Tables 5.1 and 5.2). Using the DHS data, we find that the indigenous population makes up over 70% of the total rural population, whilst almost 80% of the non-indigenous population resides in urban areas. Urban-rural differences have to be considered in particular because of better infrastructure and provision of public services in urban areas. The urban population may therefore have advantages in access to health facilities and services. Moreover, sanitation and water services as well as education and social networks that might contribute to better health are expected to be of higher quality in less sparsely populated urban areas (Heaton and Forste, 2003).

---

[2]According to our estimations using the Demographic and Health Survey (DHS) data in 2003, they make up 46%.

Table 5.1: Population Shares by Indigenous Groups

DHS 2003	Bolivia	Urban	Rural	High Plains	Valleys	Lowlands	Poor
Non-indigenous (a)	54.54%	68.70%	28.92%	39.31%	44.02%	85.61%	23.94%
Indigenous (a)	45.46%	31.29%	71.08%	60.68%	55.98%	14.38%	76.06%
Quechua	27.15%	15.00%	49.13%	21.77%	51.51%	9.56%	53.28%
Aymara	17.26%	15.83%	19.83%	38.75%	4.22%	1.76%	19.61%
Other groups (b)	1.05%	0.46%	2.12%	0.16%	0.25%	3.06%	3.17%
Guarani	0.75%	0.36%	1.45%	0.01%	0.16%	2.34%	2.29%
other	0.30%	0.10%	0.67%	0.15%	0.09%	0.72%	0.88%

Read as "percentage of the Bolivian/urban/rural etc. population".
(a) **"Indigenous"** consists of Quechua, Aymara, Guarani for the DHS 2003.
(b) The category **"Other groups"** comprises Guarani and other in the DHS data 2003.
Source: DHS 2003, own estimations

Table 5.2: Regional Shares of Ethnic Groups

DHS 2003	Non-indigenous	Indigenous (a)	Quechua	Aymara	Other groups (b)	Guarani	other
Urban	81.13%	44.34%	35.59%	59.09%	28.27%	31.29%	20.78%
Rural	18.87%	55.66%	64.41%	40.91%	71.73%	68.71%	79.22%
High Plains	28.75%	53.24%	31.98%	89.55%	6.20%	0.47%	20.43%
Valleys	24.47%	37.33%	57.53%	7.41%	7.10%	6.52%	8.54%
Lowlands	46.78%	9.43%	10.50%	3.05%	86.70%	93.01%	71.02%
Poor	10.06%	32.84%	38.01%	22.89%	54.77%	54.83%	54.62%

Read as "percentage of the non-indigenous/indigenous/Quechua etc. population". Urban-rural as well as high plains-valleys-lowlands add up to 100%.
(a) **"Indigenous"** consists of Quechua, Aymara, Guarani and for the DHS 2003.
(b) The category **"Other groups"** comprises Guarani and other in the DHS data 2003.
Source: DHS 2003, own estimations

Considering the regional distribution of these population groups, one observes a pattern of location and settlement. This is important as climatic conditions, agricultural production and food availability differ by region so that nutritional patterns, diseases and access to health care vary across location (e.g. Pérez-Cueto et al., 2009). Bolivia consists of three regions corresponding to three distinct ecozones that differ according to health conditions and opportunities for production. The semi-arid high plains of the Andes in the western part of the country are characterized by cool temperatures and frost, infertile soils and irregular rainfalls that limit farming activities to raising livestock for wool production and the cultivation of crops like potatoes and cereals that can withstand the conditions. Mining is another major activity as there are still deposits of e.g. tin, zinc and silver. The high altitude can have a negative influence on health status as low oxygen concentration and low atmospheric pressure, cold and radiation might negatively affect children's growth (Morales et al., 2004). The fertile valleys in the east-southern Andes have more moderate to semi-tropical temperatures, making traditional agriculture in the form of dairy farming and the cultivation of crops easier. The eastern lowlands are mainly tropical except for the semi-arid region of the Chaco and provide fertile grounds for commercial agriculture and cattle ranches. Moreover, oil and natural gas deposits exist in this region (Liberato et al., 2006; The PRS Group, 2008).[3] Indigenous people are concentrated in the high plains with about 50% of all indigenous people followed by the valley region with about 37%. In the high plains they account for more than 60% and in the valleys for about 56% of the total population. About 50% of the non-indigenous population has its residence in the more prosperous but less settled region of the lowlands and they account for over 80% of the population there.

Indigeneity can also be used as a proxy for poverty (Stephens et al., 2006). A lower socioeconomic level is associated with a higher risk of infections and diseases due to bad nourishment and with lower access to health services and treatment (Marmot, 2005; PAHO, 2007).[4] In 2002, poverty rates reached 73.9% among the indigenous population of Bolivia whereas only 52.5% were poor among the non-indigenous population (Pozo et al., 2006). Additionally, Table 5.3 presents the distribution of some household and maternal characteristics over ethnic origin using variables contained in the DHS. The figures suggest that ethnic origin might foremost capture differences in years of education of the mother or in mother's knowledge about health.

---

[3]For a good summary of the geographical conditions see
http://www.fao.org/ag/AGP/AGPC/doc/Counprof/Bolivia/bolivia.htm, date of access May, 2010.
[4]20% of the poorest quintile have access to health services in Bolivia. In the second poorest quintile 45 % of the population have access. Only Guatemala and Peru rank lower (PAHO, 2007).

Table 5.3: Distribution of Maternal and Household Characteristics by Ethnic Origin

Characteristics of the mother			Characteristics of the household		
Mother's	indigenous	4.36	Child is	indigenous	50.13%
education	non-indig.	8.71	female	non-indig.	48.54%
	Quechua	3.85		Quechua	50.38%
	Aymara	5.15		Aymara	49.71%
	other ind.	4.67		other ind.	50.46%
Mother age	indigenous	25.31	Female	indigenous	15.27%
at birth	non-indig.	24.03	household	non-indig.	18.69%
	Quechua	25.21	head	Quechua	16.92%
	Aymara	25.57		Aymara	12.89%
	other ind.	23.58		other ind.	11.76%
Mother knows	indigenous	86.38%	Number of	indigenous	1.02
modern	non-indig.	98.66%	children under	non-indig.	0.82
contraceptive	Quechua	87.00%	five in the	Quechua	1.04
method	Aymara	85.23%	household	Aymara	0.97
	other ind.	89.23%		other ind.	1.21
Mother has	indigenous	50.48%	Number of	indigenous	0.59
problems with	non-indig.	32.06%	children under	non-indig.	0.49
where to get	Quechua	47.51%	three in the	Quechua	0.61
medical help	Aymara	55.94%	household	Aymara	0.55
	other ind.	38.04%		other ind.	0.68
Mother has	indigenous	69.51%	Household	indigenous	5.77
problems with	non-indig.	54.05%	size	non-indig.	5.41
distance to	Quechua	69.04%		Quechua	5.83
medical help	Aymara	70.70%		Aymara	5.59
	other ind.	62.23%		other ind.	6.93

Based on these examples it becomes obvious that one needs to go beyond descriptive bivariate analysis to detect whether it is ethnic origin, residing in urban or rural areas, living in a certain geographical location, wealth or other household related characteristics that make a difference for health outcomes.

There are investigations that study health using a multivariate regression framework controlling for other relevant factors like altitude or wealth. Mayer-Foulkes and Larrea (2005) conducted a study on Bolivia. They base their inquiry on health concentration indices and a decomposition of inequality controlling for education and health. To proxy for health they build four indices taking as input variables on maternal and child health from Demographic and Health Survey (DHS) data (1997). These are a health knowledge index, a health service use index, a health status index and a summary measure that combines the three indices. They find health inequalities related to ethnic origin. Indigenous people living primarily in rural areas and having a lower educational status suffer more from health problems. Larrea and Freire (2002) investigate social inequality in child malnutrition in Bolivia, Ecuador, Peru and Colombia using multivariate regressions. In the case of Bolivia using DHS data (1997), indigenous people are found to have twice as high prevalence rates of stunting as their non-indigenous counterparts, with indigenous people in the highlands suffering more than those

living in the lowlands. Morales et al. (2004) also focus on malnutrition in Bolivia. In a mul-
tivariate regression framework, they find that belonging to the native group of the Quechua
and living at a high altitude increase malnutrition.

Although a multivariate regression framework allows for more precision concerning the
influence of indigeneity, these studies have their limitations as well. Some of them focus
only on malnutrition indicators and delimit implications concerning health for this problem
(Larrea and Freire, 2002; Morales et al., 2004). Others like Mayer-Foulkes and Larrea (2005)
use composite measures of health that make it difficult to derive policy recommendations
when it comes to prevention strategies for particular diseases. Another shortcoming of these
studies except for that of Morales et al. (2004) is that they fail to differentiate between the
ethnic groups living in Bolivia.

The indigenous population of Bolivia is not homogeneous but consists of distinct com-
munities which comprise different cultures, customs, traditions, and beliefs (Layton and Pa-
trinos, 2006). Over 30 different groups live on the Bolivian territory. The largest indigenous
groups are the Aymara (about 17% DHS 2003) and the Quechua (about 27% DHS 2003) fol-
lowed by the Guaraní and other groups (about 1% or less each DHS 2003). The Aymara and
the Quechua populate the high plains and the valleys. About 90% of the Aymara and about
40% of the Quechua live in the high plains. 58% of the Quechua live in the valleys. The
Quechua group makes up the largest share of the poor population followed by the Aymara
although over 50% of the other indigenous groups are poor. These examples show that ne-
glecting heterogeneity within the indigenous population might mask differences and compli-
cate policy implications. Even then native origin remains a black box. For example, Morales
et al. (2004) cautiously attribute differences in malnutrition outcomes between Aymara and
Quechua to cultural patterns as genetic differences between the two groups should be min-
imal. Underlying social norms and culture might result in differences in behavior between
the ethnic groups which might affect health outcomes. Although the investigations control
for household characteristics, there is still a need to go beyond the indigenous, Quechua or
Aymara dummy.

Using several indicators on childhood diseases and overall morbidity (diarrhea, stunting
and under-five mortality) and data on vaccinations (DPT/Polio, measles, tuberculosis/BCG)
taken from the DHS 2003, this study seeks to contribute to this literature, focusing on five
major research questions:

(1) Is there health inequality between indigenous and non-indigenous children? Accord-
ing to the previous discussion, one would expect that for every health indicator indigeneity is
associated with a higher probability of suffering from a disease or a worse health status and
a lower probability of receiving vaccination.

(2) Does the indigenous dummy mask heterogeneity in health outcomes between children of different indigenous origin? As there are different native people in Bolivia that live in different areas of the country the hypothesis is that there are significant differences in health outcomes between native groups. As the Aymara and Quechua people are the largest native groups in Bolivia we investigate whether both or only one of these two groups is significantly different from the non-indigenous population. When exploring the differences between the groups we again take into account factors like urban-rural differences, poverty and regional location as well as household and maternal characteristics to check whether these differences remain.

(3) Is the indigenous dummy a proxy for urban-rural differences, poverty and regional location, so that the effect of indigeneity vanishes if one takes these factors into account? The effect of indigeneity should capture the effect of these factors. However, based on the literature, the hypothesis is that ethnic origin has an effect on health even if one controls for regional location, urban-rural differences and wealth.

(4) Do statistically significant differences between indigenous and non-indigenous children disappear if one controls for household characteristics and characteristics of the mother? Underlying social norms and culture might be reflected in household characteristics and characteristics of the mother related to education, health knowledge and access to health care, and one aim of the study is to explore whether this is the case.

(5) Is health inequality related to wealth so that within the indigenous and non-indigenous group diseases are concentrated among the poor? Concentration indices have become standard tools of health inequality analysis and are therefore not neglected in this study (e.g. O'Donnell et al., 2008; Kakwani et al., 1997; Wagstaff et al., 1991). We compute these indices to complement the former investigation by a focus on health inequality within groups to get a more refined picture of the situation in Bolivia. The expectation is that there is concentration of ill-health among the poor and of vaccinations among the rich. To contribute to this literature we explore whether this type of health inequality is found for the different indigenous and non-indigenous groups.

To answer these questions, for each health variable we start with a bivariate analysis of contingency tables and run multivariate regressions to investigate between-group (ethnic origin) inequality in health. Then we compute concentration indices to estimate health inequality related to wealth within these groups. For under-five mortality we take a slightly different approach which enables to consider the problem of censoring. Doing such a case study and combining all these methods is a necessary step in order to detect disparities and the driving forces of child health inequality in Bolivia. It might help to derive policy im-

plications, for example to design intervention strategies, to identify target groups and create equity-promoting health systems (Braveman and Tarimo, 2002; Marmot, 2005; WHO, 2008).

The main findings of this study are the following. First, the bivariate analysis of health inequality due to ethnic origin hides possible relationships with other variables that might be proxied by the indigenous dummy. Therefore, conducting a multivariate analysis is a necessary exercise to get a precise picture. Secondly, the indigenous dummy masks variation within the indigenous group. Consequently, using dummies for different ethnic groups like the Aymara and the Quechua gives valuable information. Thirdly, dummies for these different ethnic groups also capture effects of other variables in particular characteristics of the mother should be accounted for when analyzing health inequality in Bolivia. Finally, the results are dependent on the health indicator under examination. Findings differ if one uses indicators for childhood diseases and morbidity or vaccination variables.

The next section deals with the data and variables used in this study. Section 5.3 presents the methods of the health inequality analysis and in section 5.4 the results are described. The last section concludes.

5.2   Data

For the analysis of health inequality, we use the Demographic and Health Survey (DHS) from 2003, the latest one available for Bolivia. The data set contains detailed information on birth records, anthropometric measures, several measures of childhood diseases and data on vaccinations. It also provides a set of dummies on durable goods and housing conditions that capture assets which are used to proxy wealth.

To measure health we use different outcome and health-related behavior variables. The DHS provides information on the prevalence of diarrheal diseases *(diarrhea)* in the last two weeks, which is a typical measure used in the health economics literature to capture childhood diseases and overall morbidity (Mayer-Foulkes and Larrea, 2005; PAHO, 2007; WHO, 2008). As a further indicator for morbidity, the DHS also offers data on *stunting*, defined as low height for age. We use this as evidence for chronic malnourishment, which affects the development of a child as a whole. Suffering from malnourishment has an impact on physical and cognitive development, and causes a higher risk of diseases due to insufficient vitamin and nutrient intake. Moreover, children with slow body growth are exposed to a higher risk of overweight in later years which is associated with further non-communicable diseases like diabetes (PAHO, 2007).[5] We complement this information on childhood diseases and overall

---

[5]Moreover, low height for age as the indicator for malnourishment has the advantage that in measuring the growth of children self-reporting bias is low. Furthermore, it is now established that distributions of the heights of healthy children are comparable (Sahn and Younger, 2006). For the calculation

morbidity with demographic information that we use to calculate under-five mortality, which is the most reliable indicator of child health.[6]

The DHS also contains variables on communicable diseases, in particular data on their prevention through vaccinations. Communicable diseases are still one of the major causes of child mortality in less-developed countries (Lopez et al., 2006).[7] Most of these so-called childhood cluster diseases like diphtheria or tuberculosis are preventable through vaccinations (Lopez et al., 2006). Therefore, indicators on vaccinations are useful measures to approach health inequities at a prevention level. Low levels of prevention can be associated either with problems in access to health services or with a lack of demand for vaccinations.[8] From the DHS, we take the available data on vaccinations against diphtheria, pertussis, tetanus (DPT) and polio *(DPT/Polio)*, data on vaccinations against *measles* and on bacille Calmette-Guérin vaccinations against tuberculosis *(BCG)*. Although in recent years none of the above vaccination-preventable diseases has caused high death rates, the analysis of inequality in vaccinations in the case of Bolivia is relevant, as the coverage rate with vaccinations is very low (PAHO, 2007). This makes the outbreak and spreading of diseases possible.[9]

When coding the variables we take into account that effective immunization is reached only when vaccinations are given in a certain time span and/or enough doses are administered. Moreover, age dependence of vaccinations is considered by not using information on children who at the date of the survey have not reached the age at which the immunization should be done. To deal with factors like improvements in health technology, knowledge about health, other improvements in infrastructure and possible major economic shocks that may affect generations during their lifetime, we restrict the sample for the childhood disease indicators

---

see (WHO Multicentre Growth Reference Study Group, 2006). We use the special programs of the WHO (http://www.who.int/childgrowth/software/en/) using the WHO Reference 2007.

[6]The so called neglected diseases - Chagas, dengue fever, leprosy and leishmaniasis - are still prevalent in Bolivia. Especially in rural areas poor people face a higher risk of being infected by these diseases as housing conditions are bad (Hotez et al., 2008). These diseases are expensive to diagnose and to cure. Moreover, they demand a long recovery period or can cause disability. The Bolivian departments of Santa Cruz and Pando report new incidence of these cases each year and have a large risk group of 8% in the local population (PAHO, 2007). Data on neglected diseases is hard to obtain on a survey-based level, and is thus not included in this study.

[7]Indeed non-communicable chronic diseases are the main cause of death, both in developed and less-developed countries, and are related to malnutrition and bad health conditions, or can be caused by other factors (WHO, 2008). But due to data restrictions we use the available information on communicable diseases.

[8]In general, it is possible to get six vaccinations for 11 different diseases to prevent infections that primarily affect children. The eleven diseases are: yellow fever; diphtheria, pertussis, tetanus (DPT); polio(myelitis); hepatitis B; measles, rubella, mumps; tuberculosis (BCG); and influenza (PAHO, 2007).

[9]There were numerous cases of rubella (945) in 2000/01 and countable cases of pertussis (68), and diphtheria (8) in the period 2001-2005. Also tuberculosis is of great concern since a strategy to analyze risk groups and to prevent tuberculosis is lacking. There is only an active detection of tuberculosis (Stephens et al., 2006).

to children that have not reached the age of five and for the vaccination variables to children under the age of three. Table 5.4 gives the definitions and coding scheme for the health variables used in this study.[10]

Table 5.4: Definition and Coding of Health Variables

Diarrhea	had diarrhea recently missing if child is older than 5
Stunting	low height for age, 1 if z-score<-2 missing if child is older than 5
DPT/Polio	combined variable of vaccinations against diphtheria, pertussis and tetanus and vaccinations against polio 1 if child received all three doses of DPT and polio in the first 12 months. Missing if the child has not reached the age of 12 months at the date of the interview WHO (2003): immunization against polio with three doses at 6, 10 and 14 weeks WHO (2006): immunization against DPT in the first year starting at 6 months with a distance of 4 weeks between the three doses
BCG	bacille Calmette-Guérin vaccine against tuberculosis 1 if child received BCG vaccination in the first four weeks Missing if the child has not reached the age of 1 month at the date of the interview Colditz et al. (1995), WHO (2004) immunization when received as soon as possible after birth (newborns and infants)
Measles	Measles vaccines 1 if the first measles vaccination is between 12 and 15 months of age WHO (2009): immunization against measles not before 12 months

As indigeneity is the focus of this study we have to be concerned about an appropriate measure. The most often used identifiers are language and self-identification, although geographical location may be combined with the two (Layton and Patrinos, 2006; Stephens et al., 2006). Using language can lead to an underestimation of the indigenous population as there may be indigenous descendants who declare their native tongue to be Spanish or who do not speak any indigenous language. Moreover, complications may arise due to the existence of multilingual populations. Self-identification overcomes the disadvantages of the language identifier but it can lead either to an underestimation of the indigenous group if there is discrimination against and social exclusion of indigenous people, or to an overestimation if there are benefits connected with being indigenous (Layton and Patrinos, 2006).

The DHS 2003 includes a simple measure of languages spoken that is used to identify the indigenous population and to build an *indigenous* dummy.[11] Additionally, we use information on different indigenous groups, as - besides the possible errors in measurement - one has to be

---

[10]When coding the variables we faced a trade-off between accuracy and having enough observations. This trade-off explains the deviations from the schedule recommended by the WHO.

[11]The DHS 1994 and 1998 from Bolivia only include information about the language of the household interview. We do not accept this as an adequate measure for ethnicity.

aware that the two groups, indigenous and non-indigenous, are not homogeneous in the sense that they are built out of distinct communities which comprise different cultures, customs, traditions and beliefs (Layton and Patrinos, 2006). We construct a dummy for the *Quechua* and the *Aymara* populations as they constitute the largest indigenous groups in Bolivia. We do not consider the other indigenous groups as even of they are combined in one category sample size is too small to consider them as a single category in a regression analysis.

The DHS provides information on household assets that we combine into an asset-index to proxy wealth using polychoric principal component analysis, which is the appropriate method for categorical variables (Kolenikov and Angeles, 2009).[12] Using the asset index as a long-term measure of economic wealth based on stock indicators (Deaton, 1997; Mayer-Foulkes and Larrea, 2005), we classify the population into quintiles of wealth defining the first quintile as *poor* and all other quintiles as non-poor.

Besides using indigeneity and wealth, we make the general distinction between *urban* and rural areas in Bolivia. Urban areas are defined as towns with more than 2,000 inhabitants and differ from rural areas in the infrastructure and public services provided to the population. The urban population may have advantages in access to health facilities and services and may benefit from better sanitation and water services, education and social networks (Heaton and Forste, 2003). Alongside these considerations, we include the urban-rural division in the health analysis, as a major part of the rural population is indigenous whereas the non-indigenous population mainly lives in urban areas. Geographically, we analyze the data according to the three ecozones *high plains-valleys-lowlands* mentioned in the Introduction, as geographical characteristics are assumed to be related to health.

Additionally, we use indicators of household characteristics and characteristics of the mother as control variables in the regression analysis described below. These household characteristics are the number of children in a household under age three or five (*children und. three/children und. five*), household size (*hh size*), a dummy indicating *low water quality* based on the source of water, information on the sex of the child (*girl*) and the sex of the household head (*female hh head*). The characteristics of the mother should capture mother's knowledge about health and her ability to deal with health problems and access to health care (see, e.g. Liberato et al., 2006; Mayer-Foulkes and Larrea, 2005; Morales et al., 2004). We use mother's age at birth (*m's age at birth*), mother's education (*m's education*), a variable

---

[12]The assets used to measure housing quality and access to public services are source of drinking water, type of toilet facility, has electricity, main floor material, main wall material, main roof material, share toilet with other, household's type of cooking fuel, place for hand washing, water tap, has an exclusive room for kitchen, number of rooms excluding kitchen and bathrooms, number of bedrooms, electric water pump. Durable goods are measured with the following items: has radio, has television, has refrigerator, has bicycle, has motorcy-cle/scooter, has car/truck, has telephone.

capturing whether the mother knows a modern contraceptive method (*knowl. of contracept*), a variable indicating whether the mother lacks knowledge about where to go to get medical help for herself (*probl. with med. help*) and a variable indicating problems in getting medical help caused by the distance to a health facility (*probl. with distance to med. help*). In Appendix 5, Table 5.33 presents descriptive statistics for the variables used.

5.3   Methodology - Health Inequality Analysis

5.3.1   Analysis of Health Inequality Between Groups: Contingency Tables and Multivariate Regressions

We start with a bivariate analysis of contingency tables to investigate whether there are differences in health 'levels' between groups and check the association between two binary variables h and x, with h measuring health of a child and x measuring the group affiliation of the child.[13]   First, we consider the *conditional distribution* of h at various levels of x and compare the conditional probability that a child is sick given it is a member of group$_1$ $p_1 = P(h = 1|x = 1)$ with the conditional probability that the child is sick given that it is a member of group$_2$ $p_2 = P(h = 1|x = 0)$. Secondly, using *Pearson Chi-square test* we check whether the two variables x and h are statistically independent.[14]   Thirdly, another useful summary measure of association between binary variables is the *relative risk ratio (rr)*. It compares two groups with respect to the probability that an event is occurring in the groups. In this study the relative risk ratio gives the extent to which one of the two groups is more likely to suffer from diseases. The relative risk of getting sick for group$_1$ compared to group$_2$ is

$$rr = \frac{(p_1)}{(p_2)},$$

and for group$_2$ compared to group$_1$ respectively. If $rr = 1$, then suffering from a disease is independent of group affiliation. If $rr > 1$ then diseases are more likely in group$_1$ than in group$_2$ and vice versa (Agresti, 1990).

The next step of the health inequality analysis consists of estimating a multivariate regression model that allows investigating the effect of ethnic origin by holding the other group characteristics constant. We estimate a regression for the whole population including the indigenous dummy as the main regressor (Model 1) and introduce step by step the other dummy

---

[13]For a good introduction to the analysis of contingency tables see Agresti (1990).

[14]To evaluate the null-hypothesis that the health status of a child is independent from its group affiliation χ^2 which compares observed with expected frequencies, and considers the degrees of freedom of the test is calculated. If the probability that the χ^2-statistic belongs to a χ^2-distribution with the calculated degrees of freedom is smaller than 0.05, then the null hypothesis of independence is rejected.

variables poor, urban, valleys and lowlands (with high plains being the reference category) (Models 2, 3, 4 and 5). Then we add a set of control variables measuring characteristics of the household (number of children in a household, household size, sex of the household head and the child, and water quality in the household) (Model 6). Next, we change the set of control variables and include indicators capturing characteristics of the mother related to education, health knowledge and access to health care (Model 7). In the final specification both sets of control variables are considered (Model 8). This procedure of incorporating different sets of control variables is used to explore whether the indigenous dummy stays significant. If it turns insignificant the indigenous dummy might only be a proxy for the other variables.

For each health outcome h for child i with $i = 1, ..., n$, we estimate a simple logit model of the form

$$
\begin{aligned}
P(h = 1) \quad = \quad & \Lambda(\alpha + \beta_1 Indigenous_i \\
+ \quad & \beta_2 Poor_i + \beta_3 Urban_i + \beta_4 Valleys_i + \beta_5 Lowlands_i \\
+ \quad & \text{controls household}_i + \text{controls mother}_i + \varepsilon_i),
\end{aligned}
$$

where Λ is the c.d.f of the logistic distribution. This equation presents the final specification that is estimated.

To go beyond the indigenous dummy and to account for possible heterogeneity within the indigenous population we re-estimate the eight regressions for the whole sample replacing the indigenous dummy with dummies for the Quechua and the Aymara group having the non-indigenous population as the reference category. We drop all observations belonging to other indigenous groups as sample size for these groups is too small to allow for statistical inference. This way we can detect whether both native groups or only one of them is significantly different from the non-indigenous population which would also give insights into the differences between these native groups.

5.3.2 Analysis of Health Inequality Within Groups: Concentration Indices

The bivariate analysis of contingency tables and the multivariate regression analysis give insights into the extent of health inequality *between* groups. *Concentration indices* have become standard tools of health inequality analysis and are therefore not neglected in this study (e.g. O'Donnell et al., 2008; Kakwani et al., 1997; Wagstaff et al., 1991). We compute these indices to complement the former investigation by a focus on health inequality *within* groups to get a more refined picture of the situation in Bolivia. Concentration indices measure the extent of health inequality that is systematically associated to wealth. Comparing concentration indices over population subgroups helps to examine differences in the distribution of

health between different groups (Wagstaff et al., 1991; Lindelow, 2006; O'Donnell et al., 2008). The concentration index can be computed as

$$
\begin{aligned}
C &= \frac{2}{(N\mu)} \sum_{i=1}^{N} h_i r_i - 1 \\
&= \frac{2}{\mu} cov(h, r),
\end{aligned}
$$

where h_i indicates health of household i, μ is the mean level of health and n is the sample size. r_i is the fractional rank of household i according to the wealth indicator y_i. For computing the concentration indices, we recode the vaccination variables assigning the value 1 if there was no vaccination. This makes the results comparable to those of the health status variables that also have the value 1 in case of illness. If ill-health measured by health status and (missing) vaccinations is concentrated among the poor, concentration indices will be negative. If $C = 0$, then there is equality, $C < 0$ indicates pro-rich inequality so that health is concentrated among the rich, and $C > 0$ indicates pro-poor inequality describing the opposite case (O'Donnell et al., 2008).[15]

Estimating the coefficient β from the following regression gives the concentration index

$$
2\sigma^2 \left(\frac{h_i}{\mu} \right) = \alpha + \beta r_i + \varepsilon_i.
$$

Using the standard error of $\hat{\beta}$ makes inference possible (Kakwani et al., 1997).[16]

5.3.3   Estimating and Explaining Under-five Mortality

To estimate under-five-mortality we use methods of survival analysis that take into account the issue of right-censoring. Right-censoring means that the relevant event (death of a child) had not occurred until the observation time ends. Consequently, the total length of time till the event will occur is unknown. For under-five-mortality this means that a child has not yet reached the age of five at the end of the observation period and we do not know whether it will survive up to age five or not.

To calculate mortality rates, we use the *lifetable method* that is suited to the situation when grouped survival time data is observed, although the underlying survival time is continuous. After having defined the intervals used to group the data, one can calculate the survival

---

[15]When computing the concentration indices, we correct for the number of children in a household to not penalize or reward households with high numbers or low numbers of children.

[16]The standard error of $\hat{\beta}$ does not take into account sampling variability of the mean of the health variable h that enters the left hand side. Moreover, the fractional rank r has no sampling variability, but according to O'Donnell et al. (2008) taking sampling variability into account in estimating the standard error makes only little difference, so we rely on the standard error of the regression for estimating and testing concentration indices.

probabilities.[17] Let T be the non-negative continuous survival time. Define intervals of time I_j where $j = 1, ...J : I_j : [t_j, t_j + 1)$ and let n_j denote the number of subjects at risk of dying at start of the interval, d_j the subjects that will die in the interval, c_j denotes the subjects that are censored in the interval. To handle these censored observations, one assumes that they are uniformly distributed over the interval so that half of the censored observations are at risk of dying. Therefore, one defines an adjusted number at risk of dying and reduces the size of the subjects censored in the interval by one-half. The average size of the risk set in the interval is then $n_j - (c_j/2)$ (e.g. Jenkins, 2005b; Hosmer et al., 2008). The life table estimator of the survival function is the product of the conditional probabilities of survival through the interval and is obtained using the following formula:

$$S_j \quad = \quad \prod_{j=1}^{J} \frac{n_j - (c_j/2) - d_j}{n_j - (c_j/2)}$$

The corresponding mortality rate per thousand of children is

$$M_j \quad = \quad (1 - S_j) * 1000.$$

We compute the mortality rate for the indigenous and non-indigenous population, for the Quechua and Aymara people, for urban and rural areas, and for the three regions high plains, valleys and lowlands. Under five mortality is also computed for the quintiles of wealth to explore health inequality according to wealth.

To study the effect of indigeneity on under five mortality in a multivariate framework we estimate a *discrete proportional hazard model with frailty* to allow for unobserved individual effects and to reduce bias due to omitted variables and measurement errors in observed survival times or regressors (Allison, 1982; Jenkins, 2005b).[18] We use the discrete time specification as exact survival times of the children are not known but fall within an interval of time. Let T be a non-negative continuous random variable representing survival time that again is grouped into intervals, in this case months, I_j, with $j = 1, ...J : I_j : [t_j, t_j + 1)$. Moreover, a vector of explanatory variables X is observed. This vector includes the variables that entered the logit model in the previous section. The discrete time (interval) hazard function

---

[17]The definition of intervals follows the proposition of the DHS with age segments 0, 1, 2, 3 to 5, 6 to 11, 12 to 23, 24 to 35, 36 to 47, 48 to 59 months. See http://www.measuredhs.com/help/Datasets/Methodology_of_DHS_Mortality_Rates_Estimation.htm

[18]ρ (reported in the regression Tables 5.39 and 5.45) is defined as the ratio of the heterogeneity variance to one plus the heterogeneity variance. If the hypothesis that ρ is zero cannot be rejected, frailty is not important (Jenkins, 2005a).

$h_j(X)$ for month j , which is the conditional probability that a child i with $i = 1, ..., n$ dies in month j given that the child i has not died up to this month is then defined by

$$h_i j(X) \quad = \quad Pr(T_i = j | T_i \geq j, X).$$

Specifying a functional form on how hazard depends on X we get

$$h_i j(X) \quad = \quad 1 - exp[-exp(\gamma_j + \beta X_i + \varepsilon_i)],$$

where γ_j summarizes the pattern of duration dependence. In our case we assume that duration dependence is piecewise constant so that the hazard differs between every six months. To achieve this we create duration specific interval dummy variables one for each interval of six months less 1. ε_i is the error term to account for unobserved heterogeneity or frailty and is assumed to be normally distributed. Using the complementary log-log transformation we get

$$log[-log(1 - h_i j(, X))] \quad = \quad \gamma_j + \beta X_i + \varepsilon_i.$$

This model is estimated using Maximum Likelihood estimation that takes into account censored observations.[19] The Likelihood is defined as

$$L \quad = \quad \prod_{i=1}^{n} [Pr(Ti = j)]^{c_i} [Pr(T_i > j)]^{(1-c_i)},$$

where c_i is a censoring indicator that takes the value 1 if there is no censoring and 0 if there is censoring (Allison, 1982; Jenkins, 2005b).

5.4   Results

5.4.1   General Description and Bivariate Analysis of Health in Bolivia

The under-five mortality rate of Bolivia, at 74 children per 1,000 live births, places the country at the bottom of the South American Countries.[20] The indicators for childhood diseases and overall morbidity of the DHS 2003 used in this study reflect that in 2003 about 20% to 30% of children under age of five were affected. About 22% of children had had diarrhea recently and about 26% of all children under five years of age were stunted. The low immunization rates for the typical vaccine-preventable diseases in Bolivia in 2003 also indicate problems either with health-related behavior or with supply of health services. About 9.4% of children

---

[19]Estimation is done using the Stata command xtcloglog.

[20]Estimates of http://www.childinfo.org/mortality_ufmrcountrydata.php (date of access, April 2010) provided by UNICEF for the year 2000 indicate that in 2000 Bolivia had a rate of 86 followed by Guyana with 72. In comparison to this, other Andean countries had a markedly lower rate. Under-five mortality rates for 2000 were 26 in Colombia, 11 in Chile, 34 in Ecuador and 41 in Peru.

under three had received all three doses of DPT and polio vaccinations in the first 12 months. Only 6.3% received a vaccination against measles between 12 and 15 months. The lowest rate is observed for BCG vaccinations, with 5.5% receiving this immunization in the first month after birth (Table 5.34 in Appendix 5).

Taking a bivariate perspective, Table 5.35 on under-five mortality in Bolivia shows that with about 103 children per 1,000 live births the indigenous group is worse off than the non-indigenous one with around 52 children per 1,000 live births. The analysis of contingency tables suggests that indigeneity increases the likelihood of suffering from diseases and decreases the probability of receiving a vaccination. Taking the extreme examples, indigenous children face a 77% higher risk of being stunted, whereas non-indigenous children are 44% more likely to get a BCG vaccination (Tables 5.36 and 5.37 in Appendix 5).

5.4.2   Results from Regression Analysis and Concentration Indices

The results from the regressions and the concentration indices will be analyzed by using the questions posed in the introduction. To answer the questions, it has been necessary to estimate many regression models which are presented in Appendix 5 to keep the overview in the text.

(1) Is there health inequality between indigenous and non-indigenous children? The multivariate analysis shows that whether there is a relationship between ethnic origin and health depends on the health indicator under examination. Whereas in the case of the *childhood diseases and morbidity indicators* a statistically significant relationship between ethnic origin and the probability of suffering from a disease is found, there is none if *vaccination variables* are considered. Indigenous origin is positively related to diarrhea, under-five mortality and stunting. Consequently, the hypothesis that indigenous origin is associated with a higher probability of suffering from a disease or a worse health status can be confirmed (Tables 5.38, 5.39, and 5.40). For the *vaccination variables* DPT, polio, measles and BCG vaccinations, the multivariate analysis contradicts the findings of the bivariate analysis as a significant robust effect of indigenous origin cannot be found (Tables 5.41, 5.42 and 5.43). Thus, the bivariate results hide a large amount of information in the case of vaccinations and the hypothesis that indigenous origin is related to a lower probability of receiving a vaccination has to be questioned.

(2) Does the indigenous dummy mask heterogeneity in health outcomes between children of different native origins? In all regressions, even those for the vaccination variables DPT, polio, BCG and measles vaccinations for which the indigenous dummy was not significant,

there is at least one native group that is significantly different from the non-indigenous one. Again an interesting pattern emerges. For the childhood diseases and morbidity indicators being *Quechua* is associated with a significantly higher probability of suffering from diarrhea, under-five mortality or stunting, whereas there is no statistically significant difference between the non-indigenous and the Aymara group if one controls for the whole range of control variables (Tables 5.44, 5.45 and 5.46).

Regarding the vaccination variables, the pattern is not consistent. For DPT, polio and measles vaccinations, if one controls for all the control variables only the coefficient of the *Aymara* dummy is negative and significant at the 10 percent level. While there is no statistical difference between the Quechua and non-indigenous people, being Aymara is associated with a lower probability of receiving one of these vaccinations (Tables 5.47 and 5.49). With respect to BCG vaccinations it is the Quechua dummy which is significant in the final specification but unexpectedly the Quechua face a higher likelihood to receive a BCG vaccination than the non-indigenous people (Table 5.48).

Having found that are differences in health outcomes between different native groups, to answer questions 3 and 4 we use the regression results where the Quechua and Aymara dummies are used instead of the simple indigenous dummy.[21]

(3) Are the Quechua and/or Aymara dummies proxies for urban-rural differences, poverty and regional location, so that their statistically significant effect vanishes if one takes these factors into account? As the Quechua and Aymara dummies do not turn insignificant if one takes these factors into account, they seem not to proxy them. Only in the regression for measles, the Quechua dummy is insignificant as soon as one controls for regions. However, the region dummies do not show any significant effect (Tables 5.49.

(4) Do statistically significant differences between Quechua and/or Aymara children and the non-indigenous children disappear if one controls for household characteristics and characteristics of the mother? For under-five mortality and stunting the *Aymara* dummy becomes insignificant if characteristics of the mother are included. A better health knowledge of the mother (measured with less problems of where to get medical help and a higher knowledge of modern contraceptive methods), a higher mother's education as well as a higher mother's age at birth are significantly related to a lower likelihood to die under five or to be stunted (Tables 5.39 and 5.46).

---

[21]Remember that all observations belonging to other indigenous groups have been dropped as they constitute a too small number to allow for inferences.

For the vaccination variables the picture is less coherent but again characteristics of the mother are important. For BCG vaccinations a similar picture to the one for the childhood diseases and morbidity indicators becomes apparent. The Aymara dummy becomes insignificant if maternal characteristics are controlled for. Better access to medical help for the mother, a better health knowledge of the mother and a higher mother's education are all significantly related to a higher likelihood of receiving this vaccination. With respect to DPT, polio vaccinations, it is the Quechua dummy that turns insignificant when maternal characteristics are included. Again a better health knowledge (proxied by the question whether the mother has problems with where to find medical help) and in this case a lower age at marriage of the mother are associated with a higher probability of getting vaccinated (Tables 5.47 and 5.48). For measles vaccinations and diarrhea we do not find such a result.

Concluding, the results of mother's health knowledge, access to health, mother's education and mother's age at birth show that the effect of ethnic origin vanishes as soon as these characteristics are taken into account. Acknowledging possible problems of endogeneity due to reverse causality or omitted variable bias one might interpret these variables as intermittent variables or the pathways ethnic origin takes when producing health outcomes.

(5) Is there health inequality related to wealth so that within the indigenous and non-indigenous group diseases are concentrated among the poor? Concerning wealth-related health inequality, which is reported in Tables 5.5 and 5.6, the expected pattern for ill-health so that diarrhea, stunting and under-five mortality are concentrated among the poor is confirmed. In the case of *diarrhea* the concentration index is negative and significant for the non-indigenous group and the urban children. This might be due to the fact that there is more wealth and therefore more dispersion of wealth in these groups. Regarding *stunting* there are significant and negative concentration indices for all groups. Interestingly, there is again more wealth-related health inequality for the non-indigenous than the indigenous group. Consequently, the poor of the non-indigenous are affected more. For *under-five mortality* we calculated mortality per quintile and found that under-five mortality decreases steadily with quintile of wealth.

The expected pattern that *vaccinations* are concentrated among the rich cannot be confirmed. All concentration indices indicate that vaccinations are concentrated among the poor. In 1979 the Expanded Program on Immunization (EPI) was established in Bolivia with support from PAHO and other donors. This was followed by the launch of EPI II as a response to a dramatic drop in vaccination coverage in 1996. With EPI II, vaccination schemes have become part of public health insurance, new vaccines were introduced and institutional deficiencies with respect to financing, surveillance and control were tackled. Although not

Table 5.5: Concentration Indices

	Diarrhea	Stunting	DPT/Polio	Measles	BCG
Indigenous	-0.006	-0.101***	0.107***	0.069***	0.095***
Non-indig.	-0.067*	-0.151***	0.121***	0.092***	0.118***
Urban	-0.082***	-0.153***	0.106***	0.073***	0.115***
Rural	-0.018	-0.067***	0.090***	0.048*	0.082***
High Plains	-0.009	-0.082***	0.179***	0.160***	0.140***
Valleys	-0.021	-0.213***	0.190***	0.171***	0.165***
Lowlands	-0.054	-0.150***	0.075**	0.067**	0.122***

Table 5.6: Under-five Mortality per Quintile

	Under five mortality rate	95 % Confidence Interval	
Quintile 1	118.4	127.1	110.2
Quintile 2	101.6	109.6	94
Quintile 3	79.6	86.5	73.2
Quintile 4	57.1	62.5	52.1
Quintile 5	31.5	35.3	28

without difficulties, the EPI II strategy lead to an increase in vaccination coverage (World Bank, 2001). The launch of EPI II certainly reached the poor better than previous initiatives. However, we do not have an explanation as to what might have led to the result of concentration of vaccinations among the poor.[22] It could be that nation-wide vaccination campaigns reach the poorer population better than the richer one. This should be an issue for further research.

The regression results support the findings for the concentration indices at least for the childhood disease. In this case, the poverty dummy carries a positive and significant sign. Regarding the results for the vaccination variables, the poverty dummy is not significant except for the BCG vaccinations however, the coefficient becomes insignificant when characteristics of the mother are included in the model (Table 5.42).

5.5 Conclusions, Further Research and Policy Implications

This paper is about inequities in child health related to indigenous origin in Bolivia. It aims at giving a detailed picture shedding light on whether ethnic origin is decisive for childhood diseases and vaccinations. Most of the studies investigating health inequality in Bolivia recognize that ethnic origin is a decisive factor. However, these studies have their limitations.

---

[22]According to (MMWR Weekly, 2000) there was a nationwide, house-to-house vaccination campaign initiated in September 2000 to administer all vaccines used in the routine infant vaccination schedule (diphtheria and tetanus toxoids and pertussis vaccine (DTP), measles, mumps, and rubella vaccine, and oral poliovirus vaccine).

They are limited to a bivariate analysis which should be interpreted carefully as ethnic origin is associated with other factors which might be responsible for the observed inequality by ethnic origin. Possible candidates are poverty, urban-rural differences or geographical region, and characteristics of the household and the mother. Moreover, if multiple regressions are used, it is appealing to employ only the indigenous dummy, but this has the disadvantage of hiding differences between distinct indigenous groups. Finally, most of the studies focus on only one indicator of health or combine several health indicators into an index, although the reality might be more diverse depending on the health indicator used.

In this study, we have adjusted for these limitations by conducting a multivariate regression analysis using different health indicators on childhood diseases and vaccinations, and several control variables to separate the effect of ethnic origin from other influential factors. Moreover, we have accounted for possible heterogeneity between distinct indigenous groups. The results support that it is necessary to take these issues into account.

This study yields the following insights: First, whether or not there is a relationship between ethnic origin and health depends on the health indicator under examination. Indigenous origin is positively related to childhood diseases and morbidity measured with under-five mortality, diarrhea and stunting. But for the vaccination variables a robust effect is not found.

Secondly, the indigenous dummy masks considerable heterogeneity between different native groups. Even in those regressions for which the indigenous dummy is not significant, there is at least one native group that is significantly different from the non-indigenous one. The Quechua are those, which are more likely to suffer from a bad health status than the non-indigenous children if all control variables are considered. With respect to the probability of receiving a vaccination, the Aymara are those, which are worse off than the non-indigenous children (DPT, polio and measles). Notably, in the case of BCG vaccinations the Quechua are even better off than their non-indigenous counterparts.

Thirdly, when investigating health outcomes the Aymara or Quechua dummies seem not to be proxies for regional location, poverty, urban-rural differences and characteristics of the household. However, in most of the regressions (under-five mortality, stunting, DPT, polio and BCG vaccinations) one of the Aymara and Quechua dummies turns insignificant if characteristics of the mother are included. Relevant characteristics are mother's access to health services, mother's age at birth, health knowledge of the mother and mother's education with the last two factors showing significant results in all of the here considered regressions.

Finally, health inequality related to wealth is more pronounced for the non-indigenous group than for the indigenous one. Ill-health is concentrated among the poor if childhood diseases are investigated. However, regarding vaccinations the result is rather unexpected. Vaccinations are concentrated among the poor. We do not have an explanation for this result.

In Bolivia the Expanded Program on Immunization led to an improvement in vaccination coverage and supported nation-wide vaccination campaigns. However, whether these campaigns reached the poor better than the rich or whether the rich might be more skeptical with regard to vaccinations, is open to further research.

In conclusion, conducting a multivariate analysis using dummies for different ethnic groups is essential in order to get a precise picture. But as the dummies for different ethnic groups can also capture effects of other variables, searching for the factors that are behind to ethnic differences is the most important task for future research to alleviate inequities in health. Is it institutional mechanisms that lead to discrimination or different cultural habits, is it household or maternal characteristics, wealth or geographical information, or something else? This study has shown that mother's education, mother's health knowledge, mother's age at birth and access to health are all relevant and considering them makes the dummies for ethnic origin insignificant. Moreover, these variables might be the pathways ethnic origin takes in influencing health outcomes. A further investigation on how maternal characteristics and ethnic origin are related and how they interact in producing health outcomes might be of high value. Finally, the rather unexpected and incoherent results for the vaccination variables point out that investigating these vaccinations variables in a separate study could give more insights into the functioning of health services and/or health-related behavior.

However, this study has its limitations. Having only cross-sectional data does not allow us to talk about causal effects. Comparable panel data would allow better inference. Conceptually, we do not know what is behind the indigenous, Quechua, Aymara, 'other indigenous' and the non-indigenous dummies. Furthermore, measurement errors can bias the results. For example, people with different cultural backgrounds might have a different tolerance to diseases and therefore report them differently. Moreover, the results might depend on the coding of the time span when an effective vaccination should take place. Finally, the choice of the diseases is rather arbitrary and affected by the information available in the DHS. Other diseases like non-communicable or chronic diseases should be examined in further research.

A last comment should be made here. The study certainly gives evidence that in Bolivia indigenous people have a worse health status than non-indigenous people. As indigenous people make up about 50% of the population, this has important implications for Bolivia's contribution in achieving the MDGs. However, our results emphasize that a simple formula of choosing indigenous people as a target group of health interventions falls short of recognizing the realities, as the term masks heterogeneity within this group. One should go beyond ethnic origin, "Quechua" and "Aymara" and look for factors like maternal characteristics that might be the pathways through which ethnic origin produces health outcomes. Moreover,

an operationalization of institutional mechanisms that are related to ethnic origin might give even more insights. This should be kept in mind when policies are designed.

Appendix 1

Table 5.7: Democracies (Polity2 score>1) and Autocracies (Polity2 score<=0) Classified According to their Levels of Income and Life Expectancy, 1970

	Low life expectancy		Middle life expectancy		High life expectancy	
	Mean: 44.28; Min: 34.75; Max: 51.83; Obs: 48		Mean: 58.87; Min: 51.85; Max: 66.16; Obs 48		Mean: 70.33; Min: 66.28; Max: 74.43; Obs: 47	
	Autocracies	Democracies	Autocracies	Democracies	Autocracies	Democracies
Low income	Benin; Bhutan; Burkina Faso; Burundi; Cambodia; Central African Republic; Chad; Congo, Dem. Rep.; Equatorial Guinea; Ethiopia; Indonesia; Lao PDR; Lesotho; Madagascar; Malawi; Mali; Mauritania; Nepal; Niger; Nigeria; Pakistan; Rwanda; Senegal; Somalia; Sudan; Tanzania; Togo; Uganda; Zambia	The Gambia; Ghana; India; Sierra Leone	China; Congo, Rep.; Kenya; Korea, Dem. Rep.; Mongolia; Syrian Arab Republic	Botswana; Sri Lanka	-	Comoros; Guinea-Bissau; Mozambique
Middle income	Afghanistan; Bolivia; Cameroon; Egypt. Arab Rep.; Guinea; Haiti; Liberia; Morocco; Swaziland	-	Algeria; Brazil; Dominican Republic; Ecuador; El Salvador; Honduras; Iraq; Jordan; Mexico; Panama; Paraguay; Peru; Tunisia	Colombia; Fiji; Guatemala; Korea, Rep.; Malaysia; Mauritius; Philippines; Thailand; Turkey; Zimbabwe	Cuba; Poland; Romania	Jamaica
High income	Gabon; Oman; Saudi Arabia	-	Iran, Islamic Rep.; Kuwait; Nicaragua	Chile; South Africa and Trinidad Tobago; Venezuela, RB	Argentina; Greece; Hungary; Portugal; Singapore; Spain	Australia; Austria; Belgium; Canada; Costa Rica; Cyprus; Denmark; Finland; France; Ireland; Israel; Italy; Japan; Netherlands; New Zealand; Norway; Sweden; Switzerland; United Kingdom; United States; Uruguay

Table 5.8: Democracies (Polity2 score>1) and Autocracies (Polity2 score<=0) Classified According to their Levels of Income and Life Expectancy, 1980

| | Low life expectancy
Mean: 47.88; Min: 37.59; Max: 56.29 | | Middle life expectancy
Mean: 62.63; Min: 56.43; Max: 68.60 | | High life expectancy
Mean: 71.96; Min: 68.94; Max: 76.25 | |
	Autocracies	Democracies	Autocracies	Democracies	Autocracies	Democracies
Low income	Afghanistan; Bangladesh; Benin; Bhutan; Burkina Faso; Burundi; Central African Republic; Chad; Comoros; Congo, Dem. Rep.; Congo, Rep.; Equatorial Guinea; Ethiopia; Guinea-Bissau; Indonesia; Lao PDR; Lesotho; Liberia; Madagascar; Malawi; Mali; Mauritania; Mozambique; Nepal; Niger; Pakistan; Rwanda; Senegal; Sierra Leone; Somalia; Sudan; Tanzania; Togo; Zambia	The Gambia; Ghana; India; Maldives; Nigeria; Uganda	China; Kenya; Korea, Dem. Rep.; Mongolia; Syrian Arab Republic	Sri Lanka	-	-
Middle income	Bolivia; Cameroon; Djibouti; Egypt, Arab Rep.; El Salvador; Guinea; Haiti; Swaziland	Papua New Guinea	Algeria; Brazil; Guatemala; Iran, Islamic Rep.; Iraq; Jordan; Korea, Rep.; Mexico; Morocco; Nicaragua; Paraguay; Philippines; Tunisia; Turkey	Botswana; Colombia; Dominican Republic; Ecuador; Fiji; Honduras; Malaysia; Mauritius; Peru; Solomon Islands; South Africa; Thailand; Zimbabwe	Chile; Cuba; Panama; Poland; Romania	Costa Rica; Jamaica
High income	Gabon	-	Bahrain; Oman; Qatar; Saudi Arabia; United Arab Emirates	Venezuela, RB	Argentina; Hungary; Kuwait; Singapore; Uruguay	Australia; Austria; Belgium; Brunei; Canada; Cyprus; Denmark; Finland; France; Greece; Ireland; Israel; Italy; Japan; Netherlands; New Zealand; Norway; Portugal; Spain; Sweden; Switzerland; Trinidad and Tobago; United Kingdom; United States

Table 5.9: Democracies (Polity2 score>1) and Autocracies (Polity2 score<=0) Classified According to their Levels of Income and Life Expectancy, 1990

| | Low life expectancy Mean: 48.71; Min: 37.54; Max: 61.31 | | Middle life expectancy Mean: 67.62; Min: 61.36; Max: 72.18 | | High life expectancy Mean: 76.24; Min: 72.45; Max: 81.23 | |
	Autocracies	Democracies	Autocracies	Democracies	Autocracies	Democracies
Low income	Afghanistan; Angola; Bhutan; Burkina Faso; Burundi; Chad; Congo, Dem. Rep.; Congo, Rep. Eritrea; The Gambia; Haiti; Kenya; Lao PDR; Liberia; Mauritania; Rwanda; Sierra Leone; Somalia; Sudan; Togo; Uganda; Yemen. Rep.	Benin; Cambodia; Central African Republic; Cote d'Ivoire; Ethiopia; Ghana; Guinea-Bissau; Lesotho; Madagascar; Malawi; Mali; Mozambique; Nepal; Niger; Nigeria; Senegal; Tanzania; Zambia	Korea, Dem. Rep.; Solomon Islands; Tajikistan; Vietnam	Bangladesh; Honduras; Moldova; Mongolia	Syrian Arab Republic	Serbia and Montenegro
Middle income	Cameroon; Equatorial Guinea; Guinea; Iraq; Pakistan; Swaziland; Zimbabwe	Botswana; Djibouti; Guyana; Namibia; Papua New Guinea; South Africa	Algeria; Azerbaijan; Belarus; China; Egypt, Arab Rep.; Jordan; Kazakhstan; Kyrgyz Republic; Morocco; Pakistan; Turkmenistan; Uzbekistan	Armenia; Bolivia; Brazil; Bulgaria; Colombia; Dominican Republic; El Salvador; Fiji; Georgia; Guatemala; Guyana; India; Indonesia; Iran, Islamic Rep.; Jamaica; Latvia; Lithuania; Nicaragua; Paraguay; Philippines; Romania; Russian Federation; Thailand; Turkey; Ukraine	Cuba; Libya; Tunisia	Albania; Costa Rica; Croatia; Ecuador; Macedonia, FYR; Mexico; Panama; Poland; Slovak Republic; Sri Lanka; Venezuela, RB
High income	Gabon	-	Qatar; Saudi Arabia	Estonia; Hungary; Mauritius; Trinidad and Tobago	Bahrain; Kuwait; Oman; Singapore; United Arab Emirates	Argentina; Australia; Austria; Belgium; Canada; Chile; Cyprus; Czech Republic; Denmark; Finland; France; Germany; Greece; Ireland; Israel; Italy; Japan; Korea, Rep.; Malaysia; Netherlands; New Zealand; Norway; Portugal; Slovenia; Spain; Sweden; Switzerland; United Kingdom; United States; Uruguay

Table 5.10: Democracies (Polity2 score>1) and Autocracies (Polity2 score<=0) Classified According to their Levels of Income and Life Expectancy, 2000

	Low life expectancy Mean: 48.71; Min: 37.54; Max: 61.31		Middle life expectancy Mean: 67.62; Min: 61.36; Max: 72.18		High life expectancy Mean: 76.24; Min: 72.45; Max: 81.23	
	Autocracies	**Democracies**	**Autocracies**	**Democracies**	**Autocracies**	**Democracies**
Low income	Afghanistan; Angola; Bhutan; Burkina Faso; Burundi; Chad; Congo, Dem. Rep.; Congo, Rep. Eritrea; The Gambia; Haiti; Kenya; Lao PDR; Liberia; Mauritania; Rwanda; Sierra Leone; Somalia; Sudan; Togo; Uganda; Yemen. Rep.	Benin; Cambodia; Central African Republic; Cote d'Ivoire; Ethiopia; Ghana; Guinea-Bissau; Lesotho; Madagascar; Malawi; Mali; Mozambique; Nepal; Niger; Nigeria; Senegal; Tanzania; Zambia	Korea, Dem. Rep.; Solomon Islands; Tajikistan; Vietnam	Bangladesh; Honduras; Moldova; Mongolia	Syrian Arab Republic	Serbia and Montenegro
Middle income	Cameroon; Equatorial Guinea; Guinea; Iraq; Pakistan; Swaziland; Zimbabwe	Botswana; Djibouti; Guyana; Namibia; Papua New Guinea; South Africa	Algeria; Azerbaijan; Belarus; China; Egypt. Arab Rep.; Jordan; Kazakhstan; Kyrgyz Republic; Morocco; Pakistan; Turkmenistan; Uzbekistan	Armenia; Bolivia; Brazil; Bulgaria; Colombia; Dominican Republic; El Salvador; Fiji; Georgia; Guatemala; Guyana; India; Indonesia; Iran, Islamic Rep.; Jamaica; Latvia; Lithuania; Nicaragua; Paraguay; Philippines; Romania; Russian Federation; Thailand; Turkey; Ukraine	Cuba; Libya; Tunisia	Albania; Costa Rica; Croatia; Ecuador; Macedonia, FYR; Mexico; Panama; Poland; Slovak Republic; Sri Lanka; Venezuela, RB
High income	Gabon	-	Qatar; Saudi Arabia	Estonia; Hungary; Mauritius; Trinidad and Tobago	Bahrain; Kuwait; Oman; Singapore; United Arab Emirates	Argentina; Australia; Austria; Belgium; Canada; Chile; Cyprus; Czech Republic; Denmark; Finland; France; Germany; Greece; Ireland; Israel; Italy; Japan; Korea, Rep.; Malaysia; Netherlands; New Zealand; Norway; Portugal; Slovenia; Spain; Sweden; Switzerland; United Kingdom; United States; Uruguay

Table 5.11: Democracies (Polity2 score>1) and Autocracies (Polity2 score<=0) Classified According to their Levels of Income and Literacy, 1970

	Low literacy rate Mean: 21.30; Min: 5.75; Max: 35.36		Middle literacy rate Mean: 54.21; Min: 35.64; Max: 72.89		High literacy rate Mean: 89.07; Min: 73.21; Max: 99	
	Autocracies	Democracies	Autocracies	Democracies	Autocracies	Democracies
Low income	Benin; Burkina Faso; Burundi; Central African Republic; Chad; Congo, Dem. Rep.; Congo, Rep.; Ethiopia; Mali; Mauritania; Nepal; Niger; Nigeria; Pakistan; Rwanda; Senegal; Sudan; Togo	The Gambia; Ghana; India	Cambodia; China; Equatorial Guinea; Indonesia; Kenya; Lao PDR; Lesotho; Madagascar; Malawi; Syrian Arab Republic; Tanzania; Uganda; Zambia	Botswana	Mongolia	Sri Lanka
Middle income	Algeria; Cameroon; Egypt, Arab Rep.; Haiti; Iraq; Liberia; Morocco; Tunisia	-	Bolivia; Brazil; Dominican Republic; El Salvador; Honduras; Jordan; Peru; Swaziland	Guatemala; Jamaica; Malaysia; Mauritius; Zimbabwe	Cuba; Panama; Romania	Ecuador; Paraguay; Colombia; Fiji; Philippines; Thailand
High income	Iran, Islamic Rep.; Oman; Saudi Arabia	-	Kuwait; Nicaragua; Singapore	South Africa	Argentina	Chile; Costa Rica; Cyprus; Israel; Trinidad and Tobago; Uruguay; Venezuela, RB

Table 5.12: Democracies (Polity2 score>1) and Autocracies (Polity2 score<=0) Classified According to their Levels of Income and Literacy, 1980

	Low literacy rate Mean: 29.76; Min: 7.95; Max: 47.55		Middle literacy rate Mean: 63.11; Min: 48.18; Max: 76.27		High literacy rate Mean: 91.90; Min: 79.43; Max: 99	
	Autocracies	**Democracies**	**Autocracies**	**Democracies**	**Autocracies**	**Democracies**
Low income	Bangladesh; Benin; Burkina Faso; Burundi; Central African Republic; Chad; Congo, Dem. Rep.; Ethiopia; Guinea-Bissau; Liberia; Madagascar; Malawi; Mali; Mauritania; Mozambique; Nepal; Niger; Pakistan; Rwanda; Senegal; Sudan; Togo	The Gambia; Ghana; India; Nigeria; Uganda	China; Comoros; Congo, Rep.; Equatorial Guinea; Indonesia; Kenya; Lao, PDR; Lesotho; Syrian Arab Republic; Tanzania; Zambia	-	Mongolia	Sri Lanka
Middle income	Algeria; Cameroon; Djibouti; Egypt, Arab Rep.; Haiti; Iraq; Morocco; Tunisia	-	Bolivia; Brazil; El Salvador; Guatemala; Iran, Islamic Rep.; Jordan; Nicaragua; Swaziland	Botswana; Dominican Republic; Honduras; Jamaica; Malaysia; Mauritius; Papua New Guinea; South Africa; Zimbabwe	Chile; Cuba; Panama; Paraguay; Philippines; Romania	Colombia; Costa Rica; Ecuador; Fiji; Peru; Thailand
High income	Oman	-	Bahrain; Kuwait; Qatar; Saudi Arabia; United Arab Emirates	-	Argentina; Singapore; Uruguay	Cyprus; Israel; Trinidad and Tobago; Venezuela, RB

Table 5.13: Democracies (Polity2 score>1) and Autocracies (Polity2 score<=0) Classified According to their Levels of Income and Literacy, 1990

	Low literacy rate Mean: 39.57; Min: 11.40; Max: 57.88		Middle literacy rate Mean: 72.17; Min: 57.96; Max: 82.18		High literacy rate Mean: 94.530; Min: 85.47; Max: 99	
	Autocracies	Democracies	Autocracies	Democracies	Autocracies	Democracies
Low income	Bangladesh; Benin; Burkina Faso; Burundi; Central African Republic; Chad; Congo, Dem. Rep.; Ethiopia; Guinea-Bissau; Lao, PDR.; Liberia; Malawi; Mali; Mauritania; Mozambique; Niger; Nigeria; Rwanda; Senegal; Sudan; Togo; Uganda; Yemen, Rep.	Comoros; The Gambia; Haiti; India; Nepal	China; Equatorial Guinea; Ghana; Kenya; Lesotho; Madagascar; Syrian Arab Republic; Tanzania; Zambia	*Cambodia*	Vietnam	*Mongolia*
Middle income	Algeria; Cameroon; Djibouti; Egypt; Arab Rep.; Iraq; Morocco	Pakistan; Papua New Guinea	Congo, Rep.; Indonesia; Iran, Islamic Rep.; Jordan; Swaziland; Tunisia; Zimbabwe	Albania; Bolivia; Brazil; Botswana; Dominican Republic; El Salvador; Guatemala; Honduras; Jamaica; Malaysia; Namibia; Nicaragua; South Africa	Cuba; Uzbekistan	Argentina; Chile; Colombia; Costa Rica; Ecuador; Fiji; Panama; Paraguay; Peru; Philippines; Romania; Sri Lanka; Thailand; Uruguay; Venezuela, RB
High income	Oman	-	Bahrain; Qatar; Saudi Arabia; United Arab Emirates	Mauritius	Croatia; Estonia; Russian Federation; Singapore; Slovenia	Cyprus; Israel; Trinidad and Tobago

Table 5.14: Democracies (Polity2 score>1) and Autocracies (Polity2 score<=0) Classified According to their Levels of Income and Literacy, 2000

	Low literacy rate Mean: 49.74; Min: 15.96; Max: 67.03		Middle literacy rate Mean: 80.03; Min: 68.01; Max: 88.67		High literacy rate Mean: 96.27; Min: 89.81389; Max: 99	
	Autocracies	Democracies	Autocracies	Democracies	Autocracies	Democracies
Low income	Burkina Faso; Burundi; Chad; Comoros; Congo, Dem. Rep.; Eritrea; The Gambia; Haiti; Lao, PDR; Liberia; Mauritania; Mozambique; Rwanda; Sudan; Togo; Uganda; Yemen, Rep.	Bangladesh; Benin; Central African Republic; Ethiopia; Guinea-Bissau; Madagascar; Malawi; Mali; Mozambique; Nepal; Niger; Nigeria; Senegal	Congo, Rep.; Kenya; Syrian Arab Republic	Cambodia; Ghana; Honduras; Lesotho; Tanzania; Zambia	Tajikistan; Vietnam	Moldova; Mongolia
Middle income	Algeria; Egypt, Arab Rep.; Iraq; Morocco; Pakistan	Djibouti; India; Nicaragua; Papua New Guinea	Cameroon; China; Equatorial Guinea; Libya; Swaziland; Tunisia; Zimbabwe	Albania; Bolivia; Botswana; Brazil; Dominican Republic; El Salvador; Guatemala; Indonesia; Iran, Islamic Rep.; Jamaica; Namibia; South Africa	Belarus; Cuba; Jordan; Kazakhstan; Uzbekistan	Armenia; Bulgaria; Colombia; Costa Rica; Ecuador; Fiji; Guyana; Latvia; Lithuania; Panama; Paraguay; Philippines; Romania; Russian Federation; Sri Lanka; Thailand; Ukraine; Venezuela, RB
High income	-	-	Bahrain; Kuwait; Oman; Qatar; Saudi Arabia; United Arab Emirates	Malaysia; Mauritius	Gabon	Argentinian; Chile; Cyprus; Estonia; Israel; Slovenia; Trinidad and Tobago; Uruguay

Table 5.15: Summary Statistics (over 1970, 1975, 1980, 1985, 1990, 1995, 2000)

Variable			Mean	Std. Dev.	Min	Max	Observations
Life expectancy	overall		61.87384	11.31717	28.457	81.226	N = 1204
	between			10.72435	37.93457	77.24057	n = 172
	within			3.693278	46.86955	73.37855	T = 7
Life expectancy (Non-OECD)	overall		59.36761	10.75307	28.457	78.991	N=994
	between			10.05347	37.93457	75.07271	n=142
	within			3.894476	44.36332	70.87233	T=7
Literacy rate	overall		71.96323	26.83287	5.746837	99	N = 1064
	between			25.79901	10.1236	99	n = 152
	within			7.626983	45.08338	98.25513	T = 7
Literacy rate (Non-OECD)	overall		65.93476	26.28703	5.746837	99	N=861
	between			24.98993	10.1236	99	n=123
	within			8.41834	39.05492	92.22666	T=7
Mpol	overall		-0.8811	7.234761	-10	10	N = 1058
	between			6.156831	-10	10	n = 159
	within			3.85223	-14.7954	11.34748	T = 6.65409
Mpol (Non-OECD)	overall		-2.56148	6.348191	-10	10	N=867
	between			5.100214	-10	10	n=131
	within			3.862746	-13.0186	9.667095	T-bar=6.61832
Demexp	overall		0.390926	0.467843	0	1	N = 1058
	between			0.382838	0	1	n = 159
	within			0.272527	-0.40907	1.190926	T = 6.65409
Demexp (Non-OECD)	overall		0.295502	0.433169	0	1	N=867
	between			0.336134	0	1	n=131
	within			0.278805	-0.5045	1.095502	T-bar=6.61832
Gini	overall		0.442405	0.111644	0.18205	0.791961	N = 931
	between			0.105537	0.227864	0.791961	n = 133
	within			0.037392	0.295015	0.64751	T = 7
Gini (Non-OECD)	overall		0.461669	0.110174	0.18205	0.791961	N=770
	between			0.103224	0.257522	0.791961	n=110
	within			0.039575	0.314279	0.666773	T=7
Fractional.	overall		0.463355	0.255516	0	0.9302	N = 1064
	between			0.256241	0	0.9302	n = 152
	within			0	0.463355	0.463355	T = 7
Fractional. (Non-OECD)	overall		0.514433	0.237711	0	0.9302	N=875
	between			0.238531	0	0.9302	n=125
	within			0	0.514433	0.514433	T=7
log GDP	overall		8.300632	1.147447	5.139029	11.11525	N = 1063
	between			1.079833	6.099427	10.60301	n = 173
	within			0.275149	6.658481	9.536229	T = 6.14451
log GDP (Non-OECD)	overall		7.996619	1.036885	5.139029	11.11525	N=861
	between			0.975748	6.099427	10.60301	n=143
	within			0.281106	6.354468	9.232215	T-bar=6.02098

Table 5.16: Correlation Matrix (over 1970, 1975, 1980, 1985, 1990, 1995, 2000)

		Life expectancy	Literacy rate	Demexp	Mpol	log GDP	Fractional.	Gini
Life expectancy	ρ	1						
	p-value							
	obs							
Literacy rate	ρ	0.821	1					
	p-value	0.000						
	obs	1064						
Demexp	ρ	0.431	0.416	1				
	p-value	0.000	0.000					
	obs	1051	939					
Mpol	ρ	0.447	0.439	0.957	1			
	p-value	0.000	0.000	0.000				
	obs	1051	939	1058				
log GDP	ρ	0.807	0.711	0.444	0.458	1		
	p-value	0.000	0.000	0.000	0.000			
	obs	1049	951	916	916			
Fractional.	ρ	-0.529	-0.473	-0.294	-0.304	-0.449	1	
	p-value	0.000	0.000	0.000	0.000	0.000		
	obs	1064	959	1018	1018	922		
Gini	ρ	-0.475	-0.403	-0.016	-0.077	-0.382	0.393	1
	p-value	0.000	0.000	0.630	0.020	0.000	0.000	
	obs	966	889	901	901	846	896	
War	ρ	-0.247	-0.178	-0.075	-0.084	-0.261	0.160	0.102
	p-value	0.000	0.000	0.021	0.010	0.000	0.000	0.003
	obs	1017	942	947	947	975	931	840
Aids	ρ	-0.306	-0.189	-0.056	-0.097	-0.203	0.232	0.450
	p-value	0.000	0.000	0.068	0.002	0.000	0.000	0.000
	obs	1204	1064	1058	1058	1063	1064	973
Socialist	ρ	0.213	0.281	-0.332	-0.326	-0.003	-0.117	-0.477
	p-value	0.000	0.000	0.000	0.000	0.918	0.000	0.000
	obs	1204	1064	1058	1058	1063	1064	973

Appendix 2

Table 5.17: Kendall Tau b: Dimension Family Code

		earmarr	polyg	parauth	inher
earmarr	Kendall tau b	1			
	Number of obs.	112			
	p-Value				
polyg	Kendall tau b	0.2950	1		
	Number of obs.	112	112		
	p-Value	0.0001			
parauth	Kendall tau b	0.2884	0.4792	1	
	Number of obs.	112	112	112	
	p-Value	0.0001	0.0000		
inher	Kendall tau b	0.234	0.5964	0.5742	1
	Number of obs.	112	112	112	112
	p-Value	0.0020	0.0000	0.0000	

earmarr stands for the variable Early marriage, polyg for Polygamy, parauth is the variable Parental authority and inher is the variable inheritance. For a description of these variables, see section 2.2. The p-values correspond to the null hypothesis that the two variables are independent.

Table 5.18: Kendall Tau b: Dimension Civil Liberties

		obliveil
freemov	Kendall tau b	0.613
	Number of obs.	123
	p-Value	0.0000

freemov stands for the variable Freedom of movement. obliveil is the variable Obligation to wear a veil in public. For a description of these variables, see section 2.2. The p-value correspond to the null hypothesis that two variables are independent.

Table 5.19: Kendall Tau b: Dimension Physical Integrity with Missing Women

		femmut	vio	misswom
femmut	Kendall tau b	1		
	Number of obs.	114		
	p-Value			
vio	Kendall tau b	0.1584	1	
	Number of obs.	114	114	
	p-Value	0.0382		
misswom	Kendall tau b	-0.1041	0.1098	1
	Number of obs.	114	114	114
	p-Value	0.2160	0.1634	

femmut stands for the variable Female Genital Mutilation, vio for Violence against women and miss-wom is the variable Missing women. For a description of these variables, see section 2.2. The p-values correspond to the null hypothesis that the two variables are independent.

Table 5.20: Kendall Tau b: Dimension Physical Integrity without Missing Women

		vio
femmut	Kendall tau b	0.1584
	Number of obs.	114
	p-Value	0.0382

femmut stands for the variable Female Genital Mutilation and vio for Violence against women. For a description of these variables, see section 2.2. The p-value correspond to the null hypothesis that two variables are independent.

Table 5.21: Kendall tau b: Dimension Ownership Rights

		womland	womloans	womprop
womland	Kendall tau b	1		
	Number of obs.	122		
	p-Value			
womloans	Kendall tau b	0.5943	1	
	Number of obs.	122	122	
	p-Value	0.0000		
womprop	Kendall tau b	0.6438	0.5975	1
	Number of obs.	122	122	122
	p-Value	0.0000	0.0000	

womland stands for the variable Women's access to land. womloans is the variable Women's access to loans and womprop is the variable Women's access to property other than land. For a description of these variables, see section 2.2. The p-values correspond to the null hypothesis that the two variables are independent.

Figure 5.1: MJCA for the Dimension Family Code
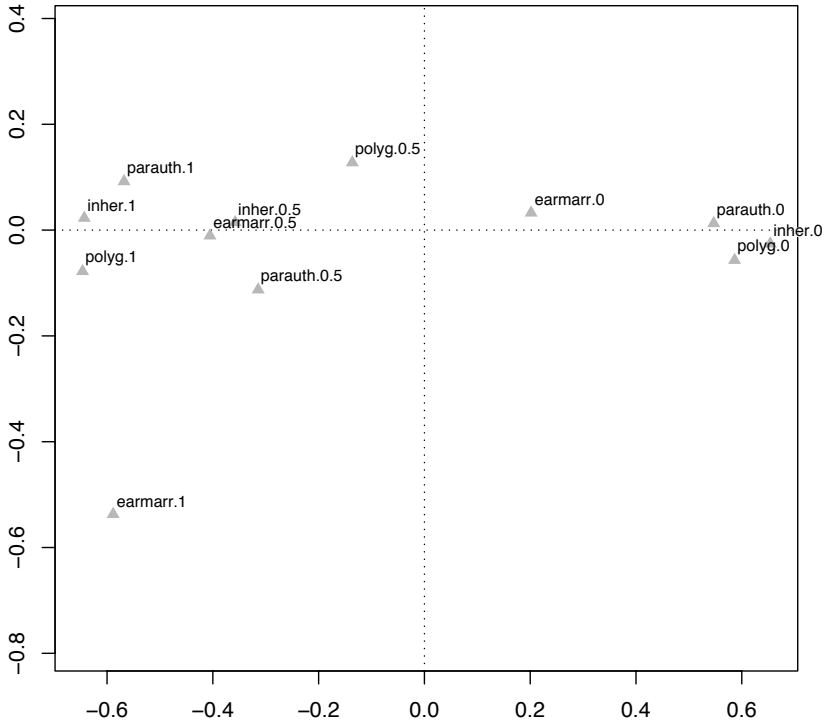
earmarr stands for the variables Early marriage, polyg for Polygamy, parauth is the variable Parental authority and inher is the variable inheritance. For a description of these variables, see section 2.2.

Figure 5.2: MJCA for the Dimension Civil Liberties
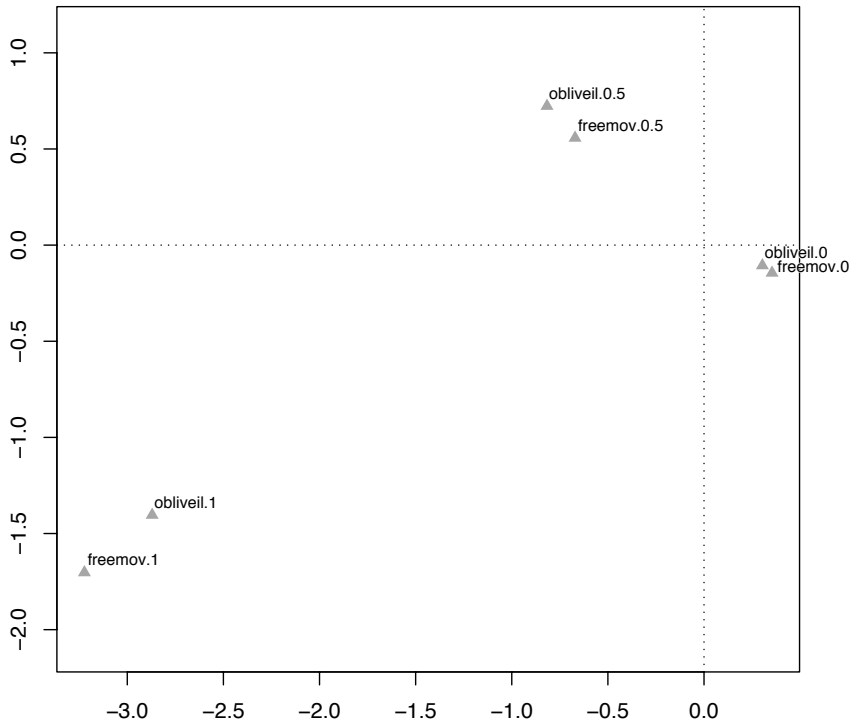
freemov stands for the variable Freedom of movement. obliveil is the variable Obligation to wear a veil in public. For a description of these variables, see section 2.2.

Figure 5.3: MJCA for the Dimension Physical Integrity with Missing Women
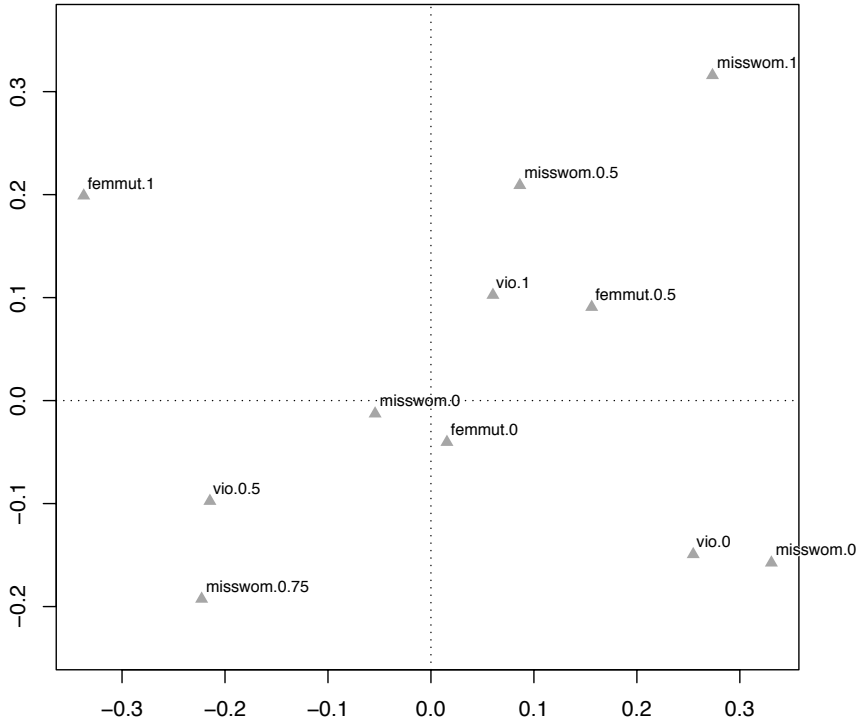
femmut stands for the variable Female Genital Mutilation, vio for Violence against women and misssk is the variable Missing women. For a description of these variables, see section 2.2.

Figure 5.4: MJCA for the Dimension Physical Integrity without Missing Women
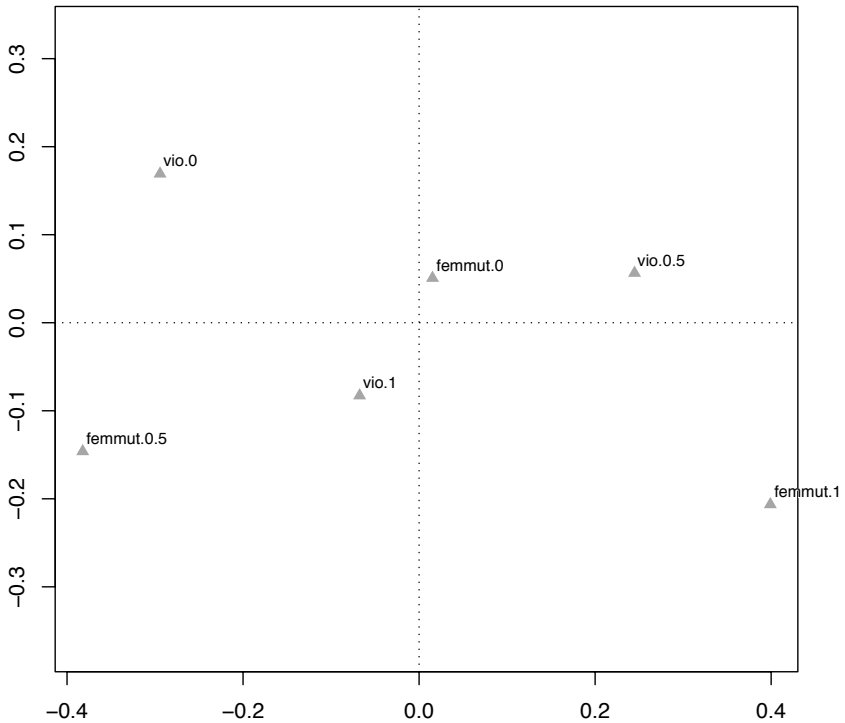
femmut stands for the variable Female Genital Mutilation and vio for Violence against women. For a description of these variables, see section 2.2.

Figure 5.5: MJCA for the Dimension Ownership Rights
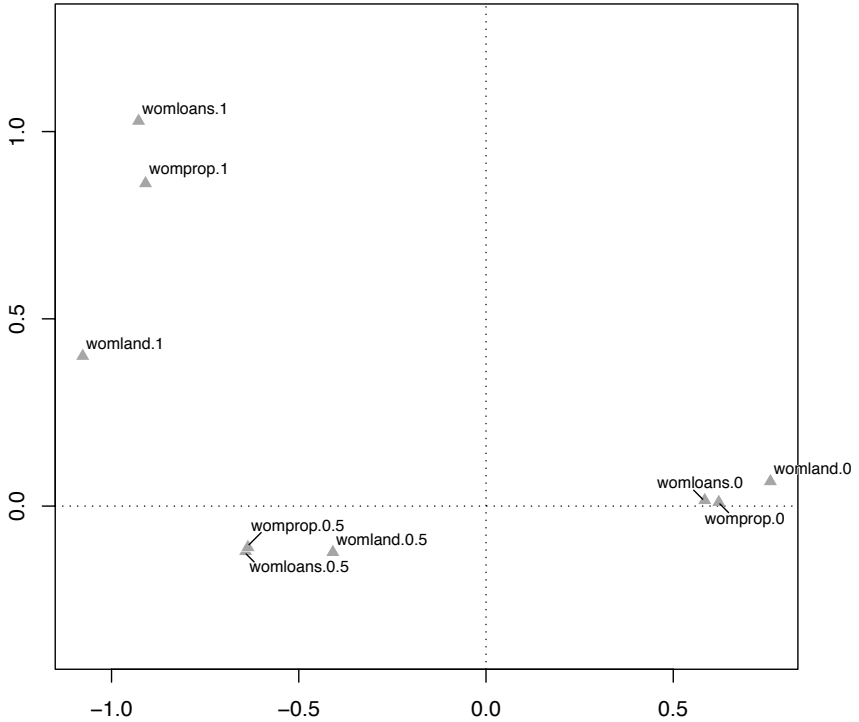
womland stands for the variable Women's access to land. womloan is the variable Women's access to loans and womprop is the variable Women's access to property other than land. For a description of these variables, see section 2.2.

Objectives, Properties and Proofs

In this section, we present the objectives and properties that we consider relevant for any composite index related to social institutions related to gender inequality. Moreover, we show that the proposed index fulfills all of them. We use the following notation. Let X^j, with $j = A, B$, be the vector containing the values of the subindices x_i^j, with $i = 1, ..., n$, for the country j[23]. $I(X)$ represents the composite index.

Objectives of the Index

The objectives of the index are the following:

1. The index $I(X)$ should represent the level of gender inequality, so that countries can be ranked.

2. The interpretation of $I(X)$ should be straightforward. As in the case of the subindices x_i, the value 0 should correspond to no inequality and the value 1 to complete inequality.

3. For any subindex x_i, we interpret the value 0, i.e. no inequality, as the goal to be achieved. The value zero can be thought of as a poverty line (see Ravallion, 1994; Deaton, 1997; Subramanian, 2007, and references therein). We define a deprivation function $\phi(x_i, 0)$, with $\phi(x_i, 0) > 0$ if $x_i > 0$, and $\phi(x_i, 0) = 0$ if $x_i = 0$. Higher values of x_i should lead to a penalization in $I(X)$ that should increase with the distance x_i to zero, i.e. $\frac{\partial I(X)}{\partial x_i} > 0$, and $\frac{\partial^2 I(X)}{\partial x_i^2} > 0$.

4. $I(X)$ should not allow for total compensation among variables, but permit partial compensation. This somehow relates to the transfer axioms that should be fulfilled by inequality as well as poverty measures. A decrease in x_i, i.e. less inequality, is rewarded more in $I(X)$ than an equivalent increase in another variable x_k (see Atkinson, 1970; Kakwani, 1984; Shorrocks and Foster, 1987; Subramanian, 2007; Alkire and Foster, 2008, and references therein).

5. $I(X)$ should be easy to compute and transparent.

---

[23]In what follows, the superscript j will only be used if it is necessary to distinguish countries.

Properties of the Index

Some of the properties that any index should fulfill are:

1. **Support and range of $I(X)$:**

   - $I(X)$ must be defined for $0 \leq x_i \leq 1$, $i = 1, \ldots, n$.

   - $0 \leq I(X) \leq 1$ must hold for any X.

   - If $x_i = 0 \; \forall i$, then $I(X) = 0$. If $x_i = 1 \; \forall i$, then $I(X) = 1$.

2. **Anonymity (symmetry)**: The value of $I(X^j)$ does not depend either on the names of the subindices nor on the name of the country (j).

3. **Unanimity (Pareto Optimality)**: If $x_i^A \leq x_i^B \; \forall i$, then $I(X^A) \leq I(X^B)$.

4. **Monotonicity**: If considering X^A and X^B country A is preferred to country B, and only x_i^A improves (i.e. decreases) for a given i, while $x_i^B \; \forall i$ remains unchanged, then country A should still be preferred over country B.

5. **Penalization of inequality in the case of equal means**: Let the mean of X^A be equal to the mean of X^B. If the dispersion of X^A is smaller than the dispersion of X^B, then $I(X^A) < I(X^B)$.

6. **Compensation property**: In a two-variable example, $\triangle x_1 \leq 1 - x_1$, and $\triangle x_2 \leq 1 - x_2$.

   (a) If x_1 increases by $|\triangle x_1|$ and x_2 decreases by $|\triangle x_2|$ and $|\triangle x_1| = |\triangle x_2|$, then $I(X)$ must increase.

   (b) For $I(X)$ to remain unchanged, we must have $|\triangle x_2| > |\triangle x_1|$.

Proofs

The composite index $I(X)$ is defined as

$$I(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - 0)^2.$$

The index proposed fulfills all the stated properties.

1. **Support and range of $I(X)$**

   - $I(X)$ is defined for $0 \le x_i \le 1$, $i = 1, \dots, n$.
   - For any X, we have that $0 \le I(X) \le 1$.
   - If $x_i = 0 \; \forall i$, then $I(X) = 0$. If $x_i = 1 \; \forall i$, then $I(X) = 1$.

2. **Anonymity (symmetry)**

   The value of $I(X^j)$ does not depend either on the names of the subindices nor on the name of the country (j).

3. **Unanimity (Pareto Optimality)**

   If we assume that $\forall i$

   $$x_i^A \le x_i^B,$$

   then we can show that

   $$
   \begin{aligned}
   (x_i^A)^2 &\le (x_i^B)^2 \\
   \frac{1}{n}\sum_{i=1}^{n}(x_i^A - 0)^2 &\le \frac{1}{n}\sum_{i=1}^{n}(x_i^B - 0)^2 \\
   I(X^A) &\le I(X^B).
   \end{aligned}
   $$

4. **Monotonicity**

   We assume that

   $$
   \begin{aligned}
   I(X^A) &\le I(X^B) \\
   \frac{1}{n}\sum_{i=1}^{n}(x_i^A - 0)^2 &\le \frac{1}{n}\sum_{i=1}^{n}(x_i^B - 0)^2.
   \end{aligned}
   $$

   Let us suppose, without loss of generality, that subindex x_1 improves (decreases) by $\delta > 0$ for country A. Then we have that

   $$\frac{1}{n}(x_1^A - \delta - 0)^2 + \frac{1}{n}\sum_{i=2}^{n}(x_i^A - 0)^2 \le \frac{1}{n}\sum_{i=1}^{n}(x_i^A - 0)^2,$$

and hence

$$\frac{1}{n}(x_1^A - \delta - 0)^2 + \frac{1}{n}\sum_{i=2}^{n}(x_i^A - 0)^2 \quad \leq \quad \frac{1}{n}\sum_{i=1}^{n}(x_i^B - 0)^2.$$

This means that

$$I(X^{A^*}) \quad \leq \quad I(X^B)$$

with X^{A^*} defined as the vector corresponding to country A with only one variable having improved (decreased) by δ.

5. **Penalization of inequality in the case of equal means**

If we assume equal means, so that

$$\mu = \frac{1}{n}\sum_{i=1}^{n}(x_i^A) = \frac{1}{n}\sum_{i=1}^{n}(x_i^B),$$

then we also have

$$\sum_{i=1}^{n}(x_i^A) = \sum_{i=1}^{n}(x_i^B).$$

If we assume that the variance of X^A is smaller than the variance of X^B so that

$$\frac{1}{n}\sum_{i=1}^{n}(x_i^A - \mu)^2 \quad < \quad \frac{1}{n}\sum_{i=1}^{n}(x_i^B - \mu)^2,$$

we can show that

$$\sum_{i=1}^{n}\left[(x_i^A)^2 - 2\mu x_i^A + \mu^2\right] \quad < \quad \sum_{i=1}^{n}\left[(x_i^B)^2 - 2\mu x_i^B + \mu^2\right],$$

$$\sum_{i=1}^{n}(x_i^A)^2 - 2\mu\sum_{i=1}^{n}x_i^A + n\mu^2 \quad < \quad \sum_{i=1}^{n}(x_i^B)^2 - 2\mu\sum_{i=1}^{n}x_i^B + n\mu^2.$$

As $\sum_{i=1}^{n}(x_i^A) = \sum_{i=1}^{n}(x_i^B)$, we have that

$$\sum_{i=1}^{n}(x_i^A)^2 \quad < \quad \sum_{i=1}^{n}(x_i^B)^2$$

$$\frac{1}{n}\sum_{i=1}^{n}(x_i^A - 0)^2 \quad < \quad \frac{1}{n}\sum_{i=1}^{n}(x_i^B - 0)^2$$

$$I(X^A) \quad < \quad I(X^B).$$

6. **Compensation property**

In a two-variable example, let $\triangle x_1 \leq 1 - x_1$, and $\triangle x_2 \leq 1 - x_2$.

(a) We can show that if $\triangle x_1 = \triangle x_2 = \delta > 0$, then

$$
\begin{aligned}
x_2 &< x_1 + \delta \\
0 &< x_1 - x_2 + \delta \\
0 &< 2\delta(x_1 - x_2 + \delta) \\
x_1^2 + x_2^2 &< x_1^2 + x_2^2 + 2\delta(x_1 - x_2 + \delta) \\
\frac{1}{2}\left(x_1^2 + x_2^2\right) &< \frac{1}{2}\left(x_1^2 + 2\delta x_1 + \delta^2 + x_2^2 - 2\delta x_2 + \delta^2\right) \\
\frac{1}{2}\left(x_1^2 + x_2^2\right) &< \frac{1}{2}\left[(x_1^2 + \delta)^2 + (x_2^2 - \delta)^2\right] \\
I(x_1, x_2) &< I(x_1 + \delta, x_2 - \delta),
\end{aligned}
$$

and hence we have shown that if x_1 increases by δ and x_2 decreases by δ, then $I(X)$ must increase.

(b) Let $x_1 = x_2 = x > 0$. We will show that if x_1 increases by $\triangle x_1$ and x_2 decreases by $\triangle x_1$ and the value of the index remains unchanged, the increase of x_1 must be smaller than the absolute value of the decrease in x_2.

$$
\begin{aligned}
I(x_1, x_2) &= I(x_1 + \triangle x_1, x_2 - \triangle x_2) \\
\frac{1}{2}\left(x_1^2 + x_2^2\right) &= \frac{1}{2}\left[(x_1 + \triangle x_1)^2 + (x_2 - \triangle x_2)^2\right] \\
x_1^2 + x_2^2 &= x_1^2 + 2x_1\triangle x_1 + (\triangle x_1)^2 + x_2^2 - 2x_2\triangle x_2 + (\triangle x_2)^2 \\
0 &= 2x_1\triangle x_1 + (\triangle x_1)^2 - 2x_2\triangle x_2 + (\triangle x_2)^2
\end{aligned}
$$

Using the fact that $x_1 = x_2 = x$, we can rewrite this as

$$
\begin{aligned}
0 &= 2x\triangle x_1 + (\triangle x_1)^2 - 2x\triangle x_2 + (\triangle x_2)^2 \\
0 &= 2x(\triangle x_1 - \triangle x_2) + (\triangle x_1)^2 + (\triangle x_2)^2.
\end{aligned}
$$

As $2x > 0$, $(\triangle x_1)^2 > 0$, and $(\triangle x_2)^2 > 0$, we must have that

$$
\begin{aligned}
\triangle x_1 - \triangle x_2 &< 0 \\
\triangle x_1 &< \triangle x_2.
\end{aligned}
$$

Table 5.22: Comparison of the SIGI and the Simple Average of the Subindices

Country	SIGI		Simple Aver.		Simple Aver. Rank
	Ranking	Value	Ranking	Value	minus SIGI rank
Paraguay	1	0.0024832	2	0.0312943	1
Croatia	2	0.00333	1	0.0273771	-1
Kazakhstan	3	0.0034778	3	0.0314302	0
Argentina	4	0.0037899	4	0.0354832	0
Costa Rica	5	0.0070934	5	0.0502099	0
Russian Fed.	6	0.0072524	11	0.0538114	5
Philippines	7	0.0078831	15	0.0603212	8
El Salvador	8	0.0082581	16	0.0647861	8
Ecuador	9	0.0091447	18	0.0700484	9
Ukraine	10	0.00969	6	0.051376	-4
Mauritius	11	0.009759	7	0.0521866	-4
Moldova	12	0.0098035	8	0.052673	-4
Bolivia	13	0.0098346	9	0.0529972	-4
Uruguay	14	0.0099167	10	0.0538078	-4
Venezuela, RB	15	0.0104259	13	0.0578608	-2
Thailand	16	0.010677	17	0.0652957	1
Peru	17	0.0121323	14	0.0586566	-3
Colombia	18	0.012727	24	0.0828911	6
Belarus	19	0.0133856	12	0.0563755	-7
Hong Kong, China	20	0.0146549	19	0.07076	-1
Singapore	21	0.0152573	20	0.0714613	-1
Cuba	22	0.0160304	22	0.0750193	0
Macedonia, FYR	23	0.0178696	23	0.0818509	0
Brazil	24	0.0188021	21	0.073534	-3
Tunisia	25	0.0190618	29	0.1012313	4
Chile	26	0.0195128	31	0.106534	5
Cambodia	27	0.0220188	27	0.0886198	0
Nicaragua	28	0.0225149	32	0.1117536	4
Trinidad & Tobago	29	0.0228815	34	0.1143368	5
Kyrgyz Rep.	30	0.0292419	36	0.12716	6
Viet Nam	31	0.0300619	25	0.0837546	-6
Armenia	32	0.0301177	26	0.0845632	-6
Georgia	33	0.0306926	28	0.0902375	-5
Guatemala	34	0.0319271	35	0.124404	1
Tajikistan	35	0.0326237	37	0.137724	2
Honduras	36	0.0331625	33	0.1122453	-3
Azerbaijan	37	0.0339496	30	0.1058964	-7
Lao PDR	38	0.0357687	39	0.1416411	1
Mongolia	39	0.0391165	43	0.1680587	4
Dominican Rep.	40	0.0398379	40	0.1440229	0
Myanmar	41	0.0462871	42	0.1553233	1
Jamaica	42	0.0484293	38	0.1399837	-4
Morocco	43	0.0534361	45	0.1973177	2
Fiji	44	0.0545044	41	0.1551223	-3
Sri Lanka	45	0.059141	47	0.2106919	2
Madagascar	46	0.0695815	44	0.1938462	-2
Namibia	47	0.0750237	49	0.241875	2
Botswana	48	0.0810172	46	0.2027736	-2
South Africa	49	0.0867689	53	0.2565411	4
Burundi	50	0.1069056	52	0.2488075	2
Albania	51	0.1071956	58	0.2715919	7
Senegal	52	0.1104056	50	0.2424129	-2
Tanzania	53	0.1124419	51	0.2445237	-2
Ghana	54	0.112694	54	0.2568415	0
Indonesia	55	0.1277609	57	0.2692867	2
Eritrea	56	0.1364469	48	0.2288967	-8
Kenya	57	0.1370416	56	0.2673039	-1
					Continued on next page

Table 5.22 – continued from previous page

| Country | SIGI | | Simple Aver. | | Simple Aver. Rank |
	Ranking	Value	Ranking	Value	minus SIGI rank
Cote d'Ivoire	58	0.1371181	59	0.2862332	1
Syrian Arab Rep.	59	0.1381059	74	0.3619356	15
Malawi	60	0.1432271	65	0.330963	5
Mauritania	61	0.1497032	68	0.3336183	7
Swaziland	62	0.1565499	70	0.3456205	8
Burkina Faso	63	0.1616069	60	0.3030649	-3
Bhutan	64	0.162508	63	0.3196661	-1
Nepal	65	0.1672252	84	0.3973769	19
Rwanda	66	0.1685859	61	0.3059172	-5
Niger	67	0.1755873	72	0.3537308	5
Equatorial Guinea	68	0.1759719	76	0.3676708	8
Gambia, The	69	0.1782978	62	0.3177497	-7
Central African Rep.	70	0.1843973	67	0.3323123	-3
Kuwait	71	0.1860213	79	0.3723096	8
Zimbabwe	72	0.1869958	78	0.3685864	6
Uganda	73	0.1871794	80	0.3735746	7
Benin	74	0.1889945	66	0.3319663	-8
Algeria	75	0.190244	87	0.4123239	12
Bahrain	76	0.1965476	89	0.4310629	13
Mozambique	77	0.1995442	82	0.3808849	5
Togo	78	0.202518	69	0.343517	-9
Congo, Dem. Rep.	79	0.2044817	64	0.3276955	-15
Papua New Guinea	80	0.2093579	83	0.3843125	3
Cameroon	81	0.2165121	85	0.4013174	4
Egypt, Arab Rep.	82	0.2176608	81	0.3779768	-1
China	83	0.2178559	55	0.2605644	-28
Gabon	84	0.2189224	86	0.4038617	2
Zambia	85	0.2193876	71	0.3526082	-14
Nigeria	86	0.2199123	92	0.4540078	6
Liberia	87	0.2265095	75	0.3629022	-12
Guinea	88	0.2280293	77	0.3678226	-11
Ethiopia	89	0.2332508	73	0.3559035	-16
Bangladesh	90	0.2446482	91	0.4491116	1
Libya	91	0.260187	94	0.5057952	3
Un. Arab Emir.	92	0.2657521	96	0.5082552	4
Iraq	93	0.2752427	97	0.522977	4
Pakistan	94	0.2832434	95	0.5062053	1
Iran, Islamic Rep.	95	0.3043608	98	0.5252544	3
India	96	0.318112	99	0.5295102	3
Chad	97	0.3225771	93	0.4733184	-4
Yemen	98	0.3270495	100	0.5567938	2
Mali	99	0.339493	88	0.422655	-11
Sierra Leone	100	0.3424468	90	0.4488637	-10
Afghanistan	101	0.5823044	101	0.746126	0
Sudan	102	0.6778067	102	0.800509	0

The data are sorted according to the value of the SIGI.

Table 5.23: Ranking according to the SIGI and the Five Subindices

Country	SIGI		Family code		Civil liberties		Physical integrity		Son preference		Ownership rights	
	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value
Paraguay	1	0.00248	19	0.06890	1	0	3	0.08757	1	0	1	0
Croatia	2	0.00333	3	0.00811	1	0	9	0.12878	1	0	1	0
Kazakhstan	3	0.00348	5	0.02837	1	0	9	0.12878	1	0	1	0
Argentina	4	0.00379	13	0.04864	1	0	9	0.12878	1	0	1	0
Costa Rica	5	0.00709	23	0.08106	1	0	15	0.16999	1	0	1	0
Russian Fed.	6	0.00725	35	0.14028	1	0	9	0.12878	1	0	1	0
Philippines	7	0.00788	8	0.04053	1	0	3	0.08757	1	0	53	0.17351
El Salvador	8	0.00826	17	0.06485	1	0	3	0.08757	1	0	43	0.17151
Ecuador	9	0.00914	24	0.08917	1	0	3	0.08757	1	0	53	0.17351
Ukraine	10	0.00969	8	0.04053	1	0	23	0.21635	1	0	1	0
Mauritius	11	0.00976	11	0.04458	1	0	23	0.21635	1	0	1	0
Moldova	12	0.00980	12	0.04701	1	0	23	0.21635	1	0	1	0
Bolivia	13	0.00983	13	0.04864	1	0	23	0.21635	1	0	1	0
Uruguay	14	0.00992	15	0.05269	1	0	23	0.21635	1	0	1	0
Venezuela, RB	15	0.01043	21	0.07295	1	0	15	0.16999	1	0	1	0
Thailand	16	0.01068	41	0.15649	1	0	15	0.16999	1	0	1	0
Peru	17	0.01213	15	0.05269	1	0	33	0.24059	1	0	1	0
Colombia	18	0.01273	21	0.07295	1	0	15	0.16999	1	0	1	0
Belarus	19	0.01339	4	0.02432	1	0	34	0.25756	1	0	43	0.17151
Hong Kong, China	20	0.01465	26	0.10380	1	0	1	0	89	0.25	1	0
Singapore	21	0.01526	25	0.09975	1	0	34	0.25756	1	0	1	0
Cuba	22	0.01603	28	0.11754	1	0	34	0.25756	1	0	1	0
Macedonia, FYR	23	0.01787	39	0.15169	1	0	34	0.25756	1	0	1	0
Brazil	24	0.01880	19	0.06890	1	0	48	0.29877	1	0	1	0
Tunisia	25	0.01906	32	0.12738	1	0	9	0.12878	89	0.25	56	0.17723
Chile	26	0.01951	34	0.13909	1	0	23	0.21635	1	0	1	0
Cambodia	27	0.02202	38	0.14433	1	0	48	0.29877	1	0	43	0.17151
Nicaragua	28	0.02251	33	0.12970	1	0	34	0.25756	1	0	1	0
Trinidad & Tobago	29	0.02288	39	0.15169	1	0	15	0.16999	89	0.25	56	0.17723
Kyrgyz Rep.	30	0.02924	42	0.15980	1	0	48	0.29877	1	0	1	0
Viet Nam	31	0.03006	6	0.03242	1	0	60	0.38634	1	0	1	0
Armenia	32	0.03012	7	0.03648	1	0	60	0.38634	1	0	1	0
Georgia	33	0.03069	17	0.06485	1	0	60	0.38634	1	0	1	0
Guatemala	34	0.03193	27	0.10538	1	0	54	0.34513	1	0	43	0.17151
Tajikistan	35	0.03262	47	0.25955	1	0	34	0.25756	1	0	43	0.17151
Honduras	36	0.03316	44	0.21610	1	0	54	0.34513	1	0	1	0
Azerbaijan	37	0.03395	37	0.14314	1	0	60	0.38634	1	0	1	0
Lao PDR	38	0.03577	51	0.32034	1	0	23	0.21635	1	0	43	0.17151
Mongolia	39	0.03912	30	0.12001	1	0	48	0.29877	89	0.25	43	0.17151
Dominican Rep.	40	0.03984	28	0.11754	1	0	34	0.25756	1	0	58	0.34502
Myanmar	41	0.04629	35	0.14028	1	0	60	0.38634	89	0.25	1	0

Table 5.23 – continued from previous page

Country	SIGI Ranking	SIGI Value	Family code Ranking	Family code Value	Civil liberties Ranking	Civil liberties Value	Physical integrity Ranking	Physical integrity Value	Son preference Ranking	Son preference Value	Ownership rights Ranking	Ownership rights Value
Jamaica	42	0.04843	1	0.00405	1	0	54	0.34513	1	0	76	0.35074
Morocco	43	0.05344	48	0.26279	1	0	9	0.12878	89	0.25	58	0.34502
Fiji	44	0.05450	8	0.04053	1	0	60	0.38634	1	0	66	0.34874
Sri Lanka	45	0.05914	46	0.23404	98	0.30069	15	0.16999	1	0	66	0.34874
Madagascar	46	0.06958	70	0.41138	1	0	60	0.38634	1	0	43	0.17151
Namibia	47	0.07502	58	0.35307	1	0	34	0.25756	89	0.25	66	0.34874
Botswana	48	0.08102	53	0.32163	1	0	15	0.16999	1	0	79	0.52225
South Africa	49	0.08677	73	0.42326	1	0	23	0.21635	1	0	58	0.34502
Burundi	50	0.10691	57	0.33545	84	0.29808	60	0.38634	1	0	79	0.52225
Albania	51	0.10720	31	0.12288	1	0	60	0.38634	101	0.5	66	0.34874
Senegal	52	0.11041	99	0.60250	1	0	45	0.26455	1	0	58	0.34502
Tanzania	53	0.11244	81	0.49886	1	0	22	0.20151	1	0	79	0.52225
Ghana	54	0.11269	61	0.36621	1	0	80	0.39575	1	0	79	0.52225
Indonesia	55	0.12776	59	0.35405	103	0.59876	79	0.39362	1	0	1	0
Eritrea	56	0.13645	76	0.45538	1	0	106	0.68910	1	0	1	0
Kenya	57	0.13704	63	0.37027	1	0	46	0.28152	1	0	111	0.68473
Cote d'Ivoire	58	0.13712	79	0.49012	1	0	85	0.43455	101	0.5	77	0.50650
Syrian Arab Rep.	59	0.13811	68	0.40269	98	0.30069	34	0.25756	1	0	66	0.34874
Malawi	60	0.14323	60	0.36087	84	0.29808	88	0.47362	1	0	79	0.52225
Mauritania	61	0.14970	71	0.42056	98	0.30069	103	0.60183	1	0	58	0.34502
Swaziland	62	0.15655	86	0.52144	84	0.29808	60	0.38634	1	0	79	0.52225
Burkina Faso	63	0.16161	88	0.53939	1	0	104	0.63092	1	0	58	0.34502
Bhutan	64	0.16251	43	0.20513	84	0.29808	54	0.34513	118	0.75	1	0
Nepal	65	0.16723	62	0.36779	84	0.29808	48	0.29877	101	0.5	79	0.52225
Rwanda	66	0.16859	56	0.32974	1	0	91	0.51512	1	0	111	0.68473
Niger	67	0.17559	104	0.64882	1	0	99	0.52482	89	0.25	58	0.34502
Equatorial Guinea	68	0.17597	82	0.50291	84	0.29808	91	0.51512	1	0	79	0.52225
Gambia, The	69	0.17830	103	0.64303	1	0	102	0.59698	1	0	66	0.34874
Central African Rep.	70	0.18440	92	0.55902	1	0	101	0.58029	1	0	79	0.52225
Kuwait	71	0.18602	83	0.50523	103	0.59876	34	0.25756	101	0.5	1	0
Zimbabwe	72	0.18718	80	0.49075	84	0.29808	59	0.36937	1	0	111	0.68473
Uganda	73	0.18899	102	0.63697	84	0.29808	81	0.41058	1	0	79	0.52225
Benin	74	0.18899	84	0.50633	1	0	87	0.46877	1	0	111	0.68473
Algeria	75	0.19024	69	0.40501	103	0.59876	60	0.38634	101	0.5	43	0.17151
Bahrain	76	0.19655	52	0.32147	103	0.59876	60	0.38634	101	0.5	66	0.34874
Mozambique	77	0.19954	109	0.69776	84	0.29808	60	0.38634	1	0	79	0.52225
Togo	78	0.20252	96	0.58833	1	0	86	0.44452	1	0	111	0.68473
Congo, Dem. Rep.	79	0.20448	66	0.39038	1	0	81	0.41058	118	0.75	119	0.83752
Papua New Guinea	80	0.20936	50	0.27697	1	0	60	0.38634	1	0	78	0.50825
Cameroon	81	0.21651	89	0.54344	84	0.29808	90	0.48332	1	0	109	0.68175
Egypt, Arab Rep.	82	0.21766	49	0.26647	98	0.30069	111	0.82273	101	0.5	1	0
China	83	0.21786	1	0.00405	1	0	48	0.29877	122	1	1	0

Continued on next page

Table 5.23 – continued from previous page

Country	SIGI Ranking	SIGI Value	Family code Ranking	Family code Value	Civil liberties Ranking	Civil liberties Value	Physical integrity Ranking	Physical integrity Value	Son preference Ranking	Son preference Value	Ownership rights Ranking	Ownership rights Value
Gabon	84	0.21892	107	0.68387	84	0.29808	91	0.51512	1	0	79	0.52225
Zambia	85	0.21939	108	0.69197	1	0	60	0.38634	1	0	111	0.68473
Nigeria	86	0.21991	71	0.42056	103	0.59876	89	0.47847	89	0.25	79	0.52225
Liberia	87	0.22651	87	0.53470	1	0	107	0.75756	1	0	79	0.52225
Guinea	88	0.22803	105	0.67140	1	0	105	0.64546	1	0	79	0.52225
Ethiopia	89	0.23325	55	0.32726	1	0	109	0.77424	1	0	108	0.67801
Bangladesh	90	0.24465	95	0.58334	103	0.59876	2	0.04121	101	0.5	79	0.52225
Libya	91	0.26019	67	0.39285	103	0.59876	91	0.51512	101	0.5	79	0.52225
Un. Arab Emir.	92	0.26575	93	0.56197	103	0.59876	100	0.53180	101	0.5	66	0.34874
Iraq	93	0.27524	77	0.47391	103	0.59876	98	0.51997	101	0.5	79	0.52225
Pakistan	94	0.28324	64	0.37821	103	0.59876	47	0.28180	118	0.75	79	0.52225
Iran, Islamic Rep.	95	0.30436	91	0.55792	119	0.78099	91	0.51512	89	0.25	79	0.52225
India	96	0.31811	100	0.60655	103	0.59876	15	0.16999	118	0.75	120	0.84049
Chad	97	0.32258	111	0.79330	98	0.30069	84	0.43212	1	0	79	0.52225
Yemen	98	0.32705	97	0.59439	119	0.78099	60	0.38634	101	0.5	58	0.34502
Mali	99	0.33949	112	0.79735	1	0	114	0.97091	1	0	121	0.84424
Sierra Leone	100	0.34245	98	0.60159	1	0	110	0.79849	122	1	109	0.68175
Afghanistan	101	0.58230	110	0.71598	121	0.81777	91	0.51512	101	0.5	122	1
Sudan	102	0.67781	106	0.67981	122	1	111	0.82273	89	0.25	1	0
Angola		NA	89	0.54344	1	0		NA	1	0	79	0.52225
Bosnia & Herzegov.		NA		NA	1	0	34	0.25756	101	0.5	1	0
Chinese Taipei		NA		NA	1	0	3	0.08757	101	0.5	1	0
Congo, Rep.		NA	101	0.62450		NA		NA	1	0	79	0.52225
Guinea-Bissau		NA		NA	1	0	107	0.75756	1	0	111	0.68473
Haiti		NA	65	0.37837	1	0	54	0.34513	1	0	1	NA
Israel		NA	45	0.22712	1	0		NA	1	0	1	0
Jordan		NA	85	0.51739	103	0.59876		NA	101	0.5	79	0.52225
Korea, Dem. Rep.		NA		NA	84	0.29808	91	0.51512	1	0	1	0
Lebanon		NA		NA	103	0.59876	60	0.38634	1	0	53	0.17351
Lesotho		NA	94	0.57149	84	0.29808		NA	1	0	79	0.52225
Malaysia		NA	53	0.32163	103	0.59876		NA	1	0	1	0
Occup. Palest. Terr.		NA	78	0.48607	103	0.59876		NA	1	0	66	0.34874
Oman		NA	74	0.45364	84	0.29808		NA	101	0.5	66	0.34874
Panama		NA		NA	1	0	8	0.11181	1	0	1	0
Puerto Rico		NA		NA	1	0	23	0.21635	1	0		NA
Saudi Arabia		NA	74	0.45364	122	1		NA	101	0.5	79	0.52225
Serbia & Monten.		NA		NA	1	0		NA		NA	43	0.17151
Somalia		NA		NA	103	0.59876	113	0.84213	101	NA	111	0.68473
Timor-Leste		NA		NA	1	0	83	0.42755	89	0.25	79	0.52225
Turkmenistan		NA		NA	1	0	60	0.38634	1	0	79	0.52225
Uzbekistan		NA		NA	1	0	60	0.38634	1	0	1	0

Table 5.24: Comparison of Ranks: the SIGI and other Gender-related Indices

Country	SIGI	GDI	GGI (capped)	GEM	GEM (revised)	GGG	WOSOC
Paraguay	1					32	19
Croatia	2	6	16	6	7	3	19
Kazakhstan	3	18	1			10	19
Argentina	4	2	21	2	3	11	3
Costa Rica	5	7	40	3	2	8	3
Russian Fed.	6	10	6	22	22	18	19
Philippines	7	22	30	10	8	1	19
El Salvador	8	29	35	13	14	20	19
Ecuador	9			14	11	17	19
Ukraine	10	19	7	23	23	25	19
Mauritius	11	12	46			44	3
Moldova	12						
Bolivia	13	35	24	19	15	41	3
Uruguay	14	5	17	15	17	39	19
Venezuela, RB	15	17	23	11	13	24	
Thailand	16	16	8	20	18	22	19
Peru	17	23	24	8	6	37	3
Colombia	18	15	11	16	16	7	3
Belarus	19	11	3			6	3
Hong Kong, China	20						
Singapore	21			1	11	38	19
Cuba	22		37			5	1
Macedonia, FYR	23	13	32	9	9	13	19
Brazil	24	14	20	20	19	36	3
Tunisia	25	26	72			55	64
Chile	26	3	44	16	20	45	3
Cambodia	27	45	10	28	26	52	3
Nicaragua	28	37	56			49	19
Trinidad & Tobago	29	9	33	4	5	19	1
Kyrgyz Rep.	30	34	11			33	19
Viet Nam	31	31	2			15	19
Armenia	32	20	4			34	19
Georgia	33			24	24	30	19
Guatemala	34	39	64			58	19
Tajikistan	35	40	19			40	19
Honduras	36	38	36	12	10	31	19
Azerbaijan	37	28	4			26	19
Lao PDR	38	47	45				3
Mongolia	39	36	27	25	25	27	3
Dominican Rep.	40	25	38			29	19
Myanmar	41		14				64
Jamaica	42	30	18			14	3
Morocco	43						19
Fiji	44						3
Sri Lanka	45	24	51	29	28	2	19
Madagascar	46	53	15			48	19
Namibia	47	43	33	5	4	9	19
Botswana	48	46	59	18	21	23	64
South Africa	49	41	42			4	19
Burundi	50	72	24				64
Albania	51						19
Senegal	52						64
Tanzania	53	66	27	7	1	12	19
Ghana	54	48	27			28	19
Indonesia	55	32	39			42	19
Eritrea	56						19

Continued on next page

Table 5.24 – continued from previous page

Country	SIGI	GDI	GGI (capped)	GEM	GEM (revised)	GGG	WOSOC
Kenya	57	57	42			43	64
Cote d'Ivoire	58	68	80				64
Syrian Arab Rep.	59	33	63			56	64
Malawi	60	70	41			46	19
Mauritania	61	60	48			60	64
Swaziland	62	59	82				64
Burkina Faso	63	76	50			66	64
Bhutan	64						3
Nepal	65	51	61			70	64
Rwanda	66	63	9				3
Niger	67	79	78				19
Equatorial Guinea	68	42	62				19
Gambia, The	69					50	19
Central African Rep.	70	75	67				19
Kuwait	71	1	48			51	64
Zimbabwe	72	58	57			47	19
Uganda	73	54	31			21	19
Benin	74	67	73			69	64
Algeria	75						64
Bahrain	76	4	76			64	64
Mozambique	77	71	47			16	64
Togo	78	61	70				64
Congo, Dem. Rep.	79	73	60				64
Papua New Guinea	80	50	22				19
Cameroon	81	55	54			65	64
Egypt, Arab Rep.	82			32	31	68	64
China	83	20	13			35	64
Gabon	84						64
Zambia	85	69	64			54	64
Nigeria	86	64	66			59	64
Liberia	87		68				19
Guinea	88	65	58				19
Ethiopia	89					62	64
Bangladesh	90	49	52	27	27	53	64
Libya	91		69				64
Un. Arab Emir.	92	8	74	30	32	57	64
Iraq	93		84				64
Pakistan	94	51	81	26	28	71	64
Iran, Islamic Rep.	95	27	54	31	30	67	64
India	96	44	77			63	19
Chad	97	74	75			72	64
Yemen	98	62	83	33	33	73	64
Mali	99	77	53			61	19
Sierra Leone	100	78	71				64
Afghanistan	101		85				19
Sudan	102	56	79				64
Number of obs.	102	79	85	33	33	73	99

Data for the Gender-related development Index (GDI) and the Gender Empowerment Measure (GEM) are from United Nations Development Programme (2006) and are based on the year 2004. The Gender Gap Index (GGI) capped and the revised Gender Empowerment Measure (GEM revised) are taken from Klasen and Schüler (2009) based on the year 2004. Data for the Global Gender Gap Index (GGG) are from Hausmann et al. (2007). The Women's Social Rights Index (WOSOC) data correspond to the year 2007 and are obtained from http://ciri.binghamton.edu/.

Appendix 3

Table 5.25: Description and Sources of Variables

Variables	Definition	Source
Response Variables		
Fertility	Total fertility rate (births per woman) (average of existing values over the last five years)	World Bank (2009)
Child mortality	Children under five mortality rate per 1,000 live births (year 2005)	World Bank (2008)
Female secondary school	School enrollment, secondary, female (% gross) (average of existing values over the last five years)	World Bank (2009)
Voice and accountability	Index that combines several data sources based on expert perceptions of "the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media" (Kaufmann et al., 2008); (average of existing values over the last five years)	Kaufmann et al. (2008)
Rule of law	Index that combines several data sources based on expert perceptions of "the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence" (Kaufmann et al., 2008); (average of existing values over the last five years)	Kaufmann et al. (2008)
Regressors		
SIGI	Social Institutions and Gender Index	Essay 2
Subindex Family code	Subindex Family code	Essay 2
Subindex Civil liberties	Subindex Civil liberties	Essay 2
Subindex Physical integrity	Subindex Physical integrity	Essay 2
Subindex Son preference	Subindex Son preference	Essay 2
Subindex Ownership rights	Subindex Ownership rights	Essay 2
Literacy female	Share of literate adult female population (15+) (%) year 2000	World Bank (2009)

Table 5.25 – continued from previous page

Variables	Definition	Source
	(average of the existing values over the last 10 years)	
Literacy population	Share of literate population (whole) (average of the existing values over the last 10 years)	Human Development Report (HDR) stats office
log GDP	Log of GDP per capita, PPP (constant 2005 international $) (average over the last 10 years)	World Bank (2008)
FH civil liberties	-1 * Index that measures the extent to which countries ensure civil liberties including freedom of expression, assembly, association, education, and religion as well as personal autonomy. It covers whether there is an established and generally equitable system of rule of law, free economic activity and equality of opportunity. (scale -1 (best) to -7 (worst)) (average of the existing values over the last 10 years)	Freedom House (2008)
Electoral democracy	Index that qualifies countries as electoral democracy when there exist competitive, universal and free and secret elections and a multiparty system that can access the media for political campaigning; (average of the existing values over the last 10 years)	Freedom House (2008)
Parliament	Proportion of seats held by women in national parliaments (%) (average of the existing values over the last 10 years)	World Bank (2009)
Aids	Adult (15-49) HIV prevalence percent by country, 1990-2007; Countries were coded 1 if Adult (15-49) HIV prevalence rate exceeds 5 per cent, otherwise 0.	UNAIDS (2008)
Ethnic	The ethnic fractionalization measure gives the probability that two individuals selected at random from a population are members of different groups. It is calculated with data on language and origin using the following formula $FRAC_j = 1 - \sum_{i=1}^{N} s_{ij}^2$, where s_{ij} is the proportion of group $i = 1, \ldots, N$ in country j going from complete homogeneity (an index of 0) to complete heterogeneity (an index of 1).	Alesina et al. (2003)
Openness	Share of imports of goods and services of total GDP	World Bank (2008)
Muslim	Countries get a 1 if at least 50 % of the population are muslim, 0 otherwise.	Central Intelligence Agency (2009)

Continued on next page

Table 5.25 – continued from previous page

Variables	Definition	Source
Christian	Countries get a 1 if at least 50 % of the population are christian, 0 otherwise.	Central Intelligence Agency (2009)
SA	Countries get a 1 if located in region South Asia, 0 otherwise.	
ECA	Countries get a 1 if located in region Europe and Central Asia, 0 otherwise.	
LAC	Countries get a 1 if located in region Latin America and the Caribbean, 0 otherwise.	
MENA	Countries get a 1 if located in region Middle East and North Africa 0 otherwise.	
EAP	Countries get a 1 if located in region East Asia and Pacific 0 otherwise.	

Table 5.26: Descriptive Statistics of the Variables Used

Variable	Observations	Mean	Std. Dev.	Min	Max
SIGI	102	0.126	0.122	0.002	0.678
Subindex Family code	112	0.326	0.223	0.004	0.797
Subindex Civil liberties	123	0.160	0.259	0	1
Subindex Physical integrity	114	0.358	0.191	0	0.971
Subindex Son preference	123	0.134	0.240	0	1
Subindex Ownership rights	122	0.298	0.266	0	1
Fertility	121	3.562	1.702	0.933	7.678
Child mortality	119	80.005	67.777	3.758	273.8
Female secondary school	108	59.210	30.484	6.037	113.275
Rule of law	123	-0.563	0.718	-2.142	1.658
Voice and accountability	123	-0.583	0.752	-2.102	1.088
SA	124	0.056	0.232	0	1
ECA	124	0.137	0.345	0	1
LAC	124	0.177	0.384	0	1
MENA	124	0.145	0.354	0	1
EAP	124	0.137	0.345	0	1
Muslim	124	0.331	0.472	0	1
Christian	124	0.435	0.498	0	1
log GDP	115	7.988	1.121	5.609	10.553
Literacy population	121	0.741	0.218	0.173	1
Literacy female	106	0.705	0.251	0.128	0.998
Electoral democracy	120	0.455	0.459	0	1
FH civil liberties	121	-4.366	1.434	-7	-1.4
Parliament	118	10.630	6.925	0	29.556
Aids	116	0.138	0.346	0	1
Openness	119	0.452	0.261	0.013	1.914
Ethnic	120	0.517	0.237	0.039	0.930

Table 5.27: Pearson Correlation Coefficient between the SIGI and the Subindices

		SIGI	Subindex Family code	Subindex Civil liberties	Subindex Physical integrity	Subindex Son preference	Subindex Ownership rights
SIGI	ρ	1					
	Obs.	102					
Subindex Family code	ρ	0.793	1				
	p-value	0.0000					
	Obs.	102	112				
Subindex Civil liberties	ρ	0.710	0.472	1			
	p-value	0.0000	0.0000				
	Obs.	102	112	123			
Subindex Physical integrity	ρ	0.661	0.594	0.282	1		
	p-value	0.0000	0.0000	0.0025			
	Obs.	102	103	113	114		
Subindex Son preference	ρ	0.535	0.179	0.530	0.020	1	
	p-value	0.0000	0.0594	0.0000	0.8312		
	Obs.	102	112	122	114	123	
Subindex Ownership rights	ρ	0.743	0.753	0.358	0.508	0.132	1
	p-value	0.0000	0.0000	0.0001	0.0000	0.1504	
	Obs.	102	111	121	112	121	122

Table 5.28: Correlation of the SIGI and the Subindices with the Control Variables

		SIGI	Subindex Family code	Subindex Civil liberties	Subindex Physical integrity	Subindex Son preference	Subindex Ownership rights
log GDP	ρ	-0.343	-0.390	0.196	-0.465	0.157	-0.481
	p-value	0.0005	0.0000	0.0362	0.0000	0.0948	0.0000
	Obs.	98	108	114	105	114	114
Muslim	ρ	0.504	0.421	0.570	0.401	0.361	0.226
	p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0122
	Obs.	102	112	123	114	123	122
Christian	ρ	-0.386	-0.332	-0.396	-0.271	-0.368	-0.052
	p-value	0.0001	0.0003	0.0000	0.0036	0.0000	0.5662
	Obs.	102	112	123	114	123	122
SA	ρ	0.298	0.134	0.326	-0.131	0.486	0.137
	p-value	0.0023	0.1589	0.0002	0.1652	0.0000	0.1319
	Obs.	102	112	123	114	123	122
ECA	ρ	-0.316	-0.379	-0.248	-0.167	-0.166	-0.329
	p-value	0.0012	0.0000	0.0057	0.0762	0.0659	0.0002
	Obs.	102	112	123	114	123	122
LAC	ρ	-0.424	-0.467	-0.289	-0.360	-0.240	-0.354
	p-value	0.0000	0.0000	0.0012	0.0001	0.0076	0.0001
	Obs.	102	112	123	114	123	122
MENA	ρ	0.231	0.164	0.533	0.083	0.417	0.017
	p-value	0.0196	0.0843	0.0000	0.3796	0.0000	0.8501
	Obs.	102	112	123	114	123	122
EAP	ρ	-0.194	-0.294	-0.111	-0.149	0.096	-0.284
	p-value	0.0505	0.0017	0.2205	0.1127	0.2934	0.0016
	Obs.	102	112	123	114	123	122
Electoral democracy	ρ	-0.388	-0.380	-0.344	-0.369	-0.217	-0.238
	p-value	0.0001	0.0000	0.0001	0.0001	0.0179	0.0091
	Obs.	101	110	119	111	119	119
FH civil liberties	ρ	-0.443	-0.298	-0.421	-0.415	-0.279	-0.251
	p-value	0.0000	0.0016	0.0000	0.0000	0.0021	0.0056
	Obs.	101	110	120	112	120	120
Parliament	ρ	-0.145	-0.150	-0.279	-0.182	-0.165	-0.105
	p-value	0.1514	0.1202	0.0023	0.0578	0.0750	0.2611
	Obs.	100	109	117	110	118	117
Literacy population	ρ	-0.657	-0.696	-0.189	-0.585	-0.252	-0.586
	p-value	0.0000	0.0000	0.0389	0.0000	0.0054	0.0000
	Obs.	102	112	120	112	121	119
Literacy female	ρ	-0.636	-0.679	-0.129	-0.581	-0.149	-0.617
	p-value	0.0000	0.0000	0.1891	0.0000	0.1286	0.0000
	Obs.	95	103	106	98	106	105
Openness	ρ	-0.195	-0.099	-0.071	-0.130	-0.125	-0.174
	p-value	0.0509	0.2995	0.4465	0.1784	0.1775	0.0605
	Obs.	101	111	118	109	118	117

Table 5.28 – continued from previous page

		SIGI	Subindex Family code	Subindex Civil liberties	Subindex Physical integrity	Subindex Son preference	Subindex Ownership rights
Ethnic	ρ	0.399	0.511	0.079	0.408	-0.105	0.463
	p-value	0.0000	0.0000	0.3918	0.0000	0.2548	0.0000
	Obs.	101	110	119	111	119	119
AIDS	ρ	0.121	0.356	0.019	0.016	-0.194	0.361
	p-value	0.2312	0.0002	0.8425	0.8684	0.0381	0.0001
	Obs.	99	108	115	107	115	115

Appendix 4

Table 5.29: Description and Sources of Variables

Variables	Definition	Source
Measures of corruption		
CPI	Corruption Perception Index (CPI); comprehensive measure of the level of corruption in a country that covers the different forms of grand and petty corruption in business, politics and administration. ranges from 0 (high corruption) to 10 (low corruption) (average of existing values over the last five years)	Transparency International (TI)
ICRG	Corruption in Government Index assesses corruption within the political system and focuses in particular on those types of corruption that lead to instability in the political system (average of existing values over the last five years)	International Country Risk Guide (ICRG)
Representation of women		
Parliament	Proportion of seats held by women in national parliaments (%) (average of the existing values over the last 10 years)	World Bank (2009)
Managers	Proportion of professional and technical, administrative and managerial positions held by women (%) (average of the existing values over the last 10 years)	World Bank (2009)
Labor force	Female labor force participation rate (average of the existing values over the last 10 years)	World Bank (2009)
Democracy		
Electoral democ.	Index that qualifies countries as electoral democracy when there exist competitive, universal and free and secret elections and a multiparty system that can access the media for political campaigning. (average of the existing values over the last 10 years)	Freedom House (2008)
Polity2	Measure of democracy taking account of competitiveness of participation, institutions and procedures openness and competitiveness of executive recruitment and constraints on the chief executive, ranges from -10 (highly autocratic) to 10 (highly democratic) , score 0 means country is democratic (average of the existing values over the last 10 years)	Marshall and Jaggers (2009)
Social inst. related to		

Continued on next page

Table 5.29 – continued from previous page

Variables	Definition	Source
gender ineq.		
Subindex civil lib.	Subindex Civil liberties that captures the freedom of social participation of women	Essay 2
Control variables		
log GDP	Log of GDP per capita, PPP (constant 2005 international $) (average over the last 10 years)	World Bank (2008)
SA	Countries get a 1 if located in region South Asia, 0 otherwise.	
ECA	Countries get a 1 if located in region Europe and Central Asia, 0 otherwise.	
LAC	Countries get a 1 if located in region Latin America and the Caribbean, 0 otherwise.	
MENA	Countries get a 1 if located in region Middle East and North Africa 0 otherwise.	
EAP	Countries get a 1 if located in region East Asia and Pacific 0 otherwise.	
Muslim	Countries get a 1 if at least 50 % of the population are Muslim, 0 otherwise.	Central Intelligence Agency (2009)
Christian	Countries get a 1 if at least 50 % of the population are Christian, 0 otherwise.	Central Intelligence Agency (2009)
Muslim percentage	percentage of Muslim population	Pew Forum on Religion and Public Life (2009)
Ethnic frac.	The ethnic fractionalization measure gives the probability that two individuals selected at random from a population are members of different groups. It is calculated with data on language and origin. The value 0 means complete homogeneity and 1 complete heterogeneity.	Alesina et al. (2003)
Literacy pop.	Literacy rate for the whole population (average of the existing values over the last 10 years)	Human Development Report (HDR) stats office
Openess	Imports of goods and services (% of GDP)	World Bank (2008)

Continued on next page

Table 5.29 – continued from previous page

Variables	Definition	Source
Not colony	Countries get a 1 if never colonized, 0 otherwise.	Correlates of War 2 Project (2003)
British colony	Countries get a 1 if former British colony, 0 otherwise.	Correlates of War 2 Project (2003)

Table 5.30: Descriptive Statistics of the Variables Used

Variable	N	mean	sd	min	max
Measures of corruption					
CPI	115	3.17	1.37	1.22	9.32
ICRG	97	2.17	0.74	0.25	4.32
Representation of women					
Parliament	119	10.76	7.03	0.00	29.56
Managers	120	7.98	5.26	0.00	23.70
Labor force	122	55.10	16.75	10.96	92.96
Democracy					
Electoral democ.	121	0.45	0.46	0.00	1.00
Polity2	98	1.09	6.08	-9.00	10.00
Social inst. related to gender ineq.					
Subindex Civil lib.	124	0.16	0.26	0.00	1.00
Control Variables					
log GDP	116	7.98	1.12	5.61	10.55
SA	125	0.06	0.23	0.00	1.00
ECA	125	0.14	0.34	0.00	1.00
LAC	125	0.18	0.38	0.00	1.00
MENA	125	0.14	0.35	0.00	1.00
EAP	125	0.14	0.35	0.00	1.00
Muslim	125	0.33	0.47	0.00	1.00
Christian	125	0.43	0.50	0.00	1.00
Muslim percentage	121	33.38	39.65	0 .00	99.7
Ethnic frac.	121	0.51	0.24	0.04	0.93
Literacy pop.	122	0.74	0.22	0.17	1.00
Openness	120	0.45	0.26	0.01	1.91
Not colony	121	0.21	0.41	0.00	1.00
British colony	121	0.30	0.46	0.00	1.00

Table 5.31: Pearson Correlation Coefficient between Subindex Civil liberties and Control Variables

log GDP	ρ	0.196	Muslim	ρ	0.570
	p-value	0.036		p-value	0.000
	Number of obs.	114		Number of obs.	123
SA	ρ	0.326	Christian	ρ	-0.396
	p-value	0.000		p-value	0.000
	Number of obs.	123		Number of obs.	123
ECA	ρ	-0.248	Ethnic	ρ	0.079
	p-value	0.006		p-value	0.392
	Number of obs.	123		Number of obs.	119
LAC	ρ	-0.289	Literacy population	ρ	-0.189
	p-value	0.001		p-value	0.039
	Number of obs.	123		Number of obs.	120
MENA	ρ	0.533	Openness	ρ	-0.071
	p-value	0.000		p-value	0.447
	Number of obs.	123		Number of obs.	118
EAP	ρ	-0.111	Not colony	ρ	-0.056
	p-value	0.221		p-value	0.549
	Number of obs.	123		Number of obs.	119
Muslim percent.	ρ	0.535	British colony	ρ	0.357
	p-value	0.000		p-value	0.000
	Number of obs.	120		Number of obs.	119

Table 5.32: Ranking According to the Subindex Civil Liberties

Country	Ranking	Subindex Civil liberties
Albania	1	0
Angola	1	0
Argentina	1	0
Armenia	1	0
Azerbaijan	1	0
Belarus	1	0
Benin	1	0
Bolivia	1	0
Bosnia and Herzegovina	1	0
Botswana	1	0
Brazil	1	0
Burkina Faso	1	0
Burundi	1	0
Cambodia	1	0
Central African Republic	1	0
Chile	1	0
China	1	0
Chinese Taipei	1	0
Colombia	1	0
Congo, Dem. Rep.	1	0
Congo, Rep.	1	0
Costa Rica	1	0
Cote d'Ivoire	1	0
Croatia	1	0
Cuba	1	0
Dominican Republic	1	0
Ecuador	1	0
El Salvador	1	0
Eritrea	1	0
Ethiopia	1	0
Fiji	1	0
Gambia, The	1	0
Georgia	1	0
Ghana	1	0
Guatemala	1	0
Guinea	1	0
Haiti	1	0
Honduras	1	0
Hong Kong, China	1	0
Israel	1	0
Jamaica	1	0
Kazakhstan	1	0
Kenya	1	0
Kyrgyz Republic	1	0
Lao PDR	1	0
Liberia	1	0
Macedonia, FYR	1	0
Madagascar	1	0
Mali	1	0
Mauritius	1	0
Moldova	1	0
Mongolia	1	0
Morocco	1	0
Myanmar	1	0
Namibia	1	0
Nicaragua	1	0
Niger	1	0
		Continued on next page

Table 5.32 – continued from previous page

Country	Ranking	Subindex Civil liberties
Panama	1	0
Papua New Guinea	1	0
Paraguay	1	0
Peru	1	0
Philippines	1	0
Russian Federation	1	0
Rwanda	1	0
Senegal	1	0
Serbia and Montenegro	1	0
Sierra Leone	1	0
Singapore	1	0
Tajikistan	1	0
Tanzania	1	0
Thailand	1	0
Timor-Leste	1	0
Togo	1	0
Trinidad and Tobago	1	0
Tunisia	1	0
Turkmenistan	1	0
Ukraine	1	0
Uruguay	1	0
Uzbekistan	1	0
Venezuela, RB	1	0
Viet Nam	1	0
Zambia	1	0
Bhutan	83	0.30059
Cameroon	83	0.30059
Equatorial Guinea	83	0.30059
Gabon	83	0.30059
Korea, Dem. Rep.	83	0.30059
Lesotho	83	0.30059
Malawi	83	0.30059
Mozambique	83	0.30059
Nepal	83	0.30059
Oman	83	0.30059
South Africa	83	0.30059
Swaziland	83	0.30059
Uganda	83	0.30059
Zimbabwe	83	0.30059
Chad	97	0.30211
Egypt, Arab Rep.	97	0.30211
Mauritania	97	0.30211
Somalia	97	0.30211
Sri Lanka	97	0.30211
Syrian Arab Republic	97	0.30211
Algeria	103	0.60269
Bahrain	103	0.60269
Bangladesh	103	0.60269
India	103	0.60269
Indonesia	103	0.60269
Iraq	103	0.60269
Jordan	103	0.60269
Kuwait	103	0.60269
Lebanon	103	0.60269
Libya	103	0.60269
Malaysia	103	0.60269
Nigeria	103	0.60269
Occupied Palestinian Territory	103	0.60269
Pakistan	103	0.60269
		Continued on next page

Table 5.32 – continued from previous page

Country	Ranking	Subindex Civil liberties
United Arab Emirates	103	0.60269
Iran, Islamic Rep.	118	0.78510
Yemen	118	0.78510
Afghanistan	120	0.81760
Saudi Arabia	121	1
Sudan	121	1

Appendix 5

Table 5.33: Descriptives of the Variables Used in the Regression Analysis

Variable	Observations	Mean	Std. Dev.	Min	Max
Diarrhea	22958	0.235	0.424	0	1
Stunting	21965	0.296	0.457	0	1
DPT/Polio	10125	0.099	0.299	0	1
Measles	10061	0.069	0.253	0	1
BCG	75998	0.012	0.110	0	1
Indigenous	82852	0.430	0.495	0	1
Quechua	82852	0.283	0.450	0	1
Aymara	82852	0.135	0.342	0	1
Other ind.	82852	0.013	0.111	0	1
Poor	47309	0.208	0.406	0	1
Urban	77193	0.613	0.487	0	1
High plains	77193	0.343	0.475	0	1
Valleys	77193	0.346	0.476	0	1
Lowlands	77193	0.311	0.463	0	1
Girl	77193	0.488	0.500	0	1
Female hh head	77193	0.179	0.383	0	1
Low water quality	67557	0.097	0.296	0	1
Children und 5	77193	0.915	0.887	0	6
Children und 3	77193	0.537	0.666	0	5
HH size	77193	5.621	2.370	1	20
M's education	74061	6.743	4.872	0	17
m's age at birth	69802	24.440	5.896	10	49
Knowl. of contracept.	112960	0.941	0.235	0	1
Probl. with med. help	112934	0.403	0.491	0	1
Probl. with distance to med. help	112909	0.596	0.491	0	1

Table 5.34: Prevalence Rates in Bolivia

U5 mortality rate	74
Diarrhea	21.96%
Stunting	25.99%
DPT/Polio	9.36%
Measles	6.28%
BCG	5.51%

Table 5.35: Under-five Mortality Rates per Thousand Live Births

	Under five	95 %	
	mortality rate	Confidence Interval	
Bolivia	74	71.1	76.700
Indigenous	103.2	98.2	108.400
Non-Indigenous	52.5	49.4	55.400
Urban	57.3	95.8	106.100
Rural	100.2	53.6	59.900
High plains	87.4	82.9	93.2
Valleys	76.4	72.3	81.9
Lowlands	55.7	50.1	58.7
Quechua	109.1	103	115.8
Aymara	95.1	85.6	103.8
Other	58.5	43.5	87.5
Poor	118.4	127.1	110.2
Rich	31.5	35.3	28

Table 5.36: Contingency Tables: Childhood Diseases

DIARRHEA	Indigenous	Non-indig.	Urban	Rural	High Plains	Valleys	Lowlands	Poor	Non-Poor
Conditional Probability	23.36%	20.94%	20.51%	25.05%	19.99%	23.76%	22.84%	27.57%	20.92%
χ^2- Test	dependent		dependent		no test			dependent	
Relat. Risk	1.115***	0.897***	0.819***	1.221***	0.858***	1.121***	1.058	1.318***	0.759***
STUNTING	**Indigenous**	**Non-indig.**	**Urban**	**Rural**	**High Plains**	**Valleys**	**Lowlands**	**Poor**	**Non-Poor**
Conditional Probability	34.68%	19.55%	20.01%	38.77%	30.91%	28.73%	16.80%	44.38%	22.43%
χ^2- Test	dependent		dependent		no test			dependent	
Relat. Risk	1.774***	0.564***	0.516***	1.938***	1.359***	1.158***	0.561***	1.979***	0.505***

Note: * p<0.10; ** p<0.05; *** p<0.01. Source: DHS 2003, own estimations

Table 5.37: Contingency Tables: Vaccinations

DPT/POLIO	Indigenous	Non-indig.	Urban	Rural	High Plains	Valleys	Lowlands	Poor	Non-Poor
Conditional Probability	8.39%	10.14%	9.61%	8.89%	8.62%	9.61%	10.09%	7.87%	9.69%
χ^2 – Test	dependent		independent		no test			dependent	
Relat. Risk	0.827**	1.209**	1.080	0.926	0.875	1.040	1.114	0.812**	1.231**
MEASLES	Indigenous	Non-indig.	Urban	Rural	High Plains	Valleys	Lowlands	Poor	Non-Poor
Conditional Probability	5.41%	6.97%	6.54%	5.80%	5.69%	5.83%	7.60%	5.07%	6.60%
Chi² Test	dependent		independent		no test			dependent	
Relative Risk	0.776***	1.289***	1.128	0.886	0.854	0.899	1.321***	0.768**	1.302**
BCG	Indigenous	Non-indig.	Urban	Rural	High Plains	Valleys	Lowlands	Poor	Non-Poor
Conditional Probability	4.42%	6.38%	6.06%	4.50%	4.63%	6.22%	5.97%	3.93%	5.92%
χ^2 – Test	dependent		dependent		no test			dependent	
Relat. Risk	0.693***	1.443***	1.346***	0.743***	0.760***	1.193**	1.124	0.663***	1.507***

Note: * p<0.10; ** p<0.05; *** p<0.01. Source: DHS 2003, own estimations

Table 5.38: Logit Regression Using Diarrhea as Dependent Variable

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
indigenous	0.190***	0.139***	0.133***	0.261***	0.191***	0.256***	0.161***	0.228***
	(0.031)	(0.033)	(0.033)	(0.035)	(0.037)	(0.040)	(0.044)	(0.047)
poor		0.322***			0.278***	0.256***	0.239***	0.223***
		(0.040)			(0.045)	(0.063)	(0.053)	(0.074)
urban			-0.194***		-0.050	-0.041	0.114**	0.159***
			(0.033)		(0.039)	(0.043)	(0.045)	(0.050)
valleys				0.225***	0.208***	0.170***	0.159***	0.129***
				(0.038)	(0.038)	(0.041)	(0.043)	(0.046)
lowlands				0.273***	0.232***	0.250***	0.245***	0.283***
				(0.043)	(0.044)	(0.047)	(0.049)	(0.053)
low water quality						0.116		0.109
						(0.071)		(0.082)
girl						-0.156***		-0.187***
						(0.034)		(0.038)
female hh head						0.082*		0.131**
						(0.047)		(0.053)
children under 5						-0.120***		-0.103***
						(0.029)		(0.032)
hh size						0.014*		0.015
						(0.009)		(0.010)
probl. distance med. help							0.117***	0.138***
							(0.039)	(0.042)
probl. med. help							0.177***	0.170***
							(0.039)	(0.043)
knowl. contracept.							0.189***	0.065
							(0.080)	(0.089)
m's educ							-0.019***	-0.022***
							(0.004)	(0.005)
m's age birth							-0.013***	-0.016***
							(0.003)	(0.004)
constant	-1.272***	-1.315***	-1.131***	-1.468***	-1.444***	-1.316***	-1.423***	-1.157***
	(0.021)	(0.022)	(0.032)	(0.036)	(0.049)	(0.073)	(0.133)	(0.152)
Number of obs.	22973	22614	22973	22973	22614	19647	19525	17057
Log-Likelihood	-12463.06	-12229.84	-12446.03	-12436.94	-12208.50	-10491.49	-10199.28	-8805.57
aic	24930.12	24465.68	24898.05	24881.89	24428.99	21004.99	20420.56	17643.14

Note: * $p<0.10$; ** $p<0.05$; *** $p<0.01$, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.39: Discrete Time Model for Under-five Mortality

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
indigenous	1.889***	1.651***	1.695***	1.780***	1.439***	1.472***	1.258***	1.248***
	(0.045)	(0.056)	(0.043)	(0.047)	(0.057)	(0.063)	(0.053)	(0.057)
poor		1.377***			1.287***	1.367***	1.206***	1.249***
		(0.049)			(0.053)	(0.076)	(0.051)	(0.072)
urban			0.699***		0.841***	0.828***	0.944	0.953
			(0.017)		(0.033)	(0.036)	(0.038)	(0.043)
valleys				0.928***	0.906***	0.893***	0.884***	0.893***
				(0.025)	(0.034)	(0.037)	(0.035)	(0.038)
lowlands				0.831***	0.764***	0.762***	0.720***	0.739***
				(0.028)	(0.037)	(0.041)	(0.036)	(0.041)
low water quality						0.982		0.950
						(0.059)		(0.059)
girl						0.897***		0.891***
						(0.032)		(0.033)
female hh head						0.814***		0.782***
						(0.044)		(0.044)
children und 5						1.032		1.039
						(0.033)		(0.034)
hh size						0.943***		0.925***
						(0.009)		(0.009)
probl. with distance to med. help							0.978	1.045
							(0.038)	(0.045)
probl. with med. help							1.035	1.071*
							(0.038)	(0.043)
knowl. of contracept.							0.957	0.958
							(0.056)	(0.063)
m's education							0.945***	0.937***
							(0.005)	(0.005)
m's age at birth							1.002***	1.001***
							(0.000)	(0.000)
constant	0.007***	0.007***	0.008***	0.007***	0.009***	0.014***	0.020***	0.032***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.006)
Number of obs.	689599	301259	689599	689599	301259	255549	296932	252098
Log-likelihood	-38564.97	-19634.03	-38460.46	-38549.66	-19608.77	-16259.12	-19068.07	-15777.99
aic	77147.94	39288.07	76940.93	77121.33	39243.54	32554.24	38172.13	31601.99
ρ	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.027

Note: * p<0.10; ** p<0.05; *** p<0.01, Standard errors in parenthesis. Source: DHS 2003, own estimations.
Time dummies measuring six months each included.

Table 5.40: Logit Regression Using Stunting as Dependent Variable

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
indigenous	0.703***	0.593***	0.548***	0.631***	0.415***	0.424***	0.185***	0.218***
	(0.030)	(0.031)	(0.032)	(0.032)	(0.035)	(0.038)	(0.040)	(0.043)
poor		0.616***			0.418***	0.387***	0.266***	0.345***
		(0.038)			(0.044)	(0.049)	(0.049)	(0.070)
urban			-0.541***		-0.405***	-0.339***	-0.225***	-0.190***
			(0.032)		(0.037)	(0.062)	(0.042)	(0.046)
valleys				-0.230***	-0.286***	-0.372***	-0.310***	-0.399***
				(0.035)	(0.036)	(0.039)	(0.040)	(0.044)
lowlands				-0.284***	-0.369***	-0.482***	-0.391***	-0.488***
				(0.041)	(0.042)	(0.047)	(0.047)	(0.052)
low water quality						0.200***		0.090
						(0.069)		(0.078)
girl						-0.404***		-0.342***
						(0.034)		(0.037)
female hh head						-0.066		-0.050
						(0.048)		(0.053)
children under 5						0.256***		0.181***
						(0.028)		(0.030)
hh size						0.040***		0.033***
						(0.008)		(0.009)
probl. distance med. help							0.076**	0.049
							(0.037)	(0.040)
probl. med. help							0.018	-0.053
							(0.038)	(0.041)
knowl. contracept.							-0.193***	-0.207***
							(0.068)	(0.078)
m's educ							-0.078***	-0.069***
							(0.004)	(0.005)
m's age birth							-0.009***	-0.010***
							(0.003)	(0.004)
constant	-1.206***	-1.284***	-0.821***	-1.010***	-0.720***	-1.131***	0.203*	-0.079
	(0.021)	(0.022)	(0.030)	(0.032)	(0.045)	(0.069)	(0.118)	(0.138)
Number of obs.	21986	21647	21986	21986	21647	18823	18766	16418
Log-Likelihood	-12995.06	-12654.54	-12849.57	-12963.61	-12552.87	-10545.65	-10563.28	-8987.31
aic	25994.12	25315.07	25705.13	25935.23	25117.75	21113.31	21148.56	18006.62

Note: * p<0.10; ** p<0.05; *** p<0.01, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.41: Logit Regression Using DPT/Polio Vaccinations as Dependent Variable

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
indigenous	-0.226***	-0.233***	-0.234***	-0.238***	-0.249***	-0.240***	-0.153*	-0.146
	(0.068)	(0.071)	(0.071)	(0.074)	(0.079)	(0.079)	(0.090)	(0.091)
poor		0.002			-0.022	-0.006	0.048	0.053
		(0.090)			(0.103)	(0.104)	(0.118)	(0.118)
urban			-0.027		-0.045	-0.045	-0.099	-0.097
			(0.071)		(0.082)	(0.083)	(0.091)	(0.091)
valleys				0.008	0.011	0.019	-0.005	0.006
				(0.080)	(0.081)	(0.081)	(0.086)	(0.087)
lowlands				-0.030	-0.018	0.013	-0.056	-0.028
				(0.088)	(0.090)	(0.091)	(0.099)	(0.100)
girl						0.091		0.074
						(0.066)		(0.072)
female hh head						-0.120		-0.160
						(0.095)		(0.107)
children under 3						-0.144*		-0.169*
						(0.083)		(0.093)
hh size						-0.016		-0.015
						(0.014)		(0.016)
probl. distance med. help							-0.034	-0.016
							(0.078)	(0.078)
probl. med. help							-0.187**	-0.168**
							(0.083)	(0.083)
knowl. contracept.							0.330*	0.335*
							(0.190)	(0.190)
m's educ							0.013	0.012
							(0.009)	(0.009)
m's age birth							-0.016***	-0.017**
							(0.006)	(0.006)
constant	-2.084***	-2.081***	-2.064***	-2.073***	-2.041***	-1.825***	-1.978***	-1.734***
	(0.041)	(0.043)	(0.067)	(0.069)	(0.099)	(0.147)	(0.281)	(0.317)
Number of obs.	10142	9959	10142	10142	9959	9959	8650	8650
Log-Likelihood	-3341.17	-3282.11	-3341.10	-3341.07	-3281.89	-3277.20	-2822.24	-2817.67
aic	6686.35	6570.22	6688.20	6690.13	6575.78	6574.39	5666.49	5665.35

Note: * $p<0.10$; ** $p<0.05$; *** $p<0.01$, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.42: Logit Regression Using BCG Vaccinations as Dependent Variable

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
indigenous	-0.347***	-0.382***	-0.249***	-0.360***	-0.336***	-0.244***	0.119	0.140
	(0.071)	(0.073)	(0.075)	(0.077)	(0.081)	(0.083)	(0.097)	(0.099)
poor		-0.633***			-0.420***	-0.452***	-0.094	-0.167
		(0.109)			(0.121)	(0.124)	(0.144)	(0.146)
urban			0.345***		0.385***	0.295***	0.130	0.125
			(0.078)		(0.086)	(0.088)	(0.100)	(0.100)
valleys				0.268***	0.331***	0.398***	0.250***	0.346***
				(0.083)	(0.085)	(0.086)	(0.094)	(0.096)
lowlands				0.064	0.049	0.234**	-0.052	0.175
				(0.093)	(0.095)	(0.096)	(0.107)	(0.109)
girl						-0.133*		-0.170**
						(0.070)		(0.077)
female hh head						-0.154		-0.220**
						(0.098)		(0.112)
children under 3						1.211***		1.311***
						(0.041)		(0.048)
hh size						-0.338***		-0.340***
						(0.022)		(0.027)
probl. distance med. help							-0.231***	-0.139
							(0.083)	(0.085)
probl. med. help							-0.368***	-0.199**
							(0.096)	(0.098)
knowl. contracept.							0.041	0.307
							(0.229)	(0.232)
m's educ							0.098***	0.085***
							(0.009)	(0.010)
m's age birth							0.002	0.014**
							(0.006)	(0.007)
constant	-4.303***	-3.694***	-4.569***	-4.417***	-4.129***	-3.500***	-4.827***	-4.850***
	(0.042)	(0.043)	(0.075)	(0.076)	(0.107)	(0.151)	(0.299)	(0.320)
Number of obs.	75976	46140	75976	75976	46140	46140	39496	39496
Log-Likelihood	-4825.79	-4289.94	-4815.37	-4819.90	-4272.00	-3944.64	-3474.26	-3199.80
aic	9655.58	8585.87	9636.74	9647.79	8556.00	7909.28	6970.52	6429.59

Note: * p<0.10; ** p<0.05; *** p<0.01, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.43: Logit Regression Using Measles Vaccinations as Dependent Variable

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
indigenous	**-0.230***	**-0.237***	**-0.258***	**-0.187**	**-0.204**	**-0.196**	**-0.113**	**-0.111**
	(0.081)	**(0.085)**	**(0.086)**	**(0.089)**	**(0.096)**	**(0.097)**	**(0.110)**	**(0.111)**
poor		0.049			-0.027	-0.016	0.132	0.135
		(0.106)			(0.120)	(0.121)	(0.139)	(0.140)
urban			-0.096		-0.097	-0.101	-0.104	-0.102
			(0.085)		(0.097)	(0.097)	(0.108)	(0.108)
valleys				0.132	0.117	0.120	0.102	0.105
				(0.096)	(0.097)	(0.097)	(0.104)	(0.104)
lowlands				0.162	0.166	0.192*	0.108	0.126
				(0.105)	(0.108)	(0.109)	(0.119)	(0.121)
girl						-0.043		-0.053
						(0.079)		(0.085)
female hh head						0.003		0.001
						(0.108)		(0.121)
children under 3						-0.128		-0.173
						(0.097)		(0.109)
hh size						-0.015		-0.010
						(0.017)		(0.020)
probl. distance med. help							0.005	0.028
							(0.093)	(0.093)
probl. med. help							-0.160	-0.153
							(0.097)	(0.099)
knowl. contracept.							0.246	0.251
							(0.221)	(0.221)
m's educ							0.012	0.010
							(0.011)	(0.011)
m's age birth							-0.007	-0.007
							(0.007)	(0.008)
constant	-2.479***	-2.482***	-2.408***	-2.595***	-2.516***	-2.272***	-2.667***	-2.380***
	(0.049)	(0.050)	(0.079)	(0.086)	(0.120)	(0.170)	(0.340)	(0.376)
Number of obs.	10027	9846	10027	10027	9846	9846	8545	8545
Log-Likelihood	-2570.33	-2528.73	-2569.67	-2568.89	-2526.78	-2524.76	-2169.52	-2167.44
aic	5144.66	5063.46	5145.33	5145.78	5065.55	5069.52	4361.04	4364.88

Note: * $p<0.10$; ** $p<0.05$; *** $p<0.01$, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.44: Logit Regression Using Diarrhea as Dependent Variable - Differentiated by Indigenous Group

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
quechua	**0.215***	**0.203***	**0.292***	**0.228***	**0.289***	**0.200***	**0.270***
	(0.036)	**(0.037)**	**(0.038)**	**(0.040)**	**(0.043)**	**(0.048)**	**(0.051)**
aymara	**-0.066**	**-0.069**	**0.084**	**0.038**	**0.129****	**0.007**	**0.080**
	(0.051)	**(0.051)**	**(0.057)**	**(0.058)**	**(0.063)**	**(0.064)**	**(0.070)**
poor	0.304***			0.265***	0.242***	0.227***	0.208***
	(0.041)			(0.046)	(0.065)	(0.054)	(0.075)
urban		-0.178***		-0.053	-0.044	0.106**	0.151***
		(0.034)		(0.039)	(0.043)	(0.045)	(0.050)
valleys			0.161***	0.149***	0.125***	0.098**	0.075
			(0.041)	(0.042)	(0.044)	(0.046)	(0.049)
lowlands			0.204***	0.178***	0.214***	0.188***	0.236***
			(0.045)	(0.046)	(0.049)	(0.051)	(0.054)
low water quality					0.123*		0.119
					(0.072)		(0.083)
girl					-0.151***		-0.184***
					(0.034)		(0.038)
female hh head					0.078*		0.121**
					(0.047)		(0.053)
children under 5					-0.134***		-0.117***
					(0.030)		(0.032)
hh size					0.013		0.016
					(0.009)		(0.010)
probl. distance med. help						0.109***	0.133***
						(0.039)	(0.042)
probl. med. help						0.180***	0.158***
						(0.040)	(0.043)
knowl. contracept.						0.189**	0.080
						(0.082)	(0.091)
m's educ						-0.019***	-0.022***
						(0.005)	(0.005)
m's age birth						-0.014***	-0.016***
						(0.003)	(0.004)
constant	-1.312***	-1.143***	-1.416***	-1.397***	-1.259***	-1.360***	-1.098***
	(0.022)	(0.032)	(0.037)	(0.049)	(0.074)	(0.135)	(0.154)
Number of obs.	22343	22680	22680	22343	19479	19312	16935
Log-Likelihood	-12042.36	-12238.01	-12239.58	-12031.61	-10378.85	-10059.40	-8722.02
aic	24092.72	24484.02	24489.16	24077.21	20781.70	20142.80	17478.05

Note: * $p<0.10$; ** $p<0.05$; *** $p<0.01$, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.45: Discrete Time Model for Under-five Mortality - Differentiated by Indigenous Group

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
quechua	2.025***	1.803***	1.798***	1.965***	1.607***	1.657***	2.042***	1.376***
	(0.053)	(0.066)	(0.049)	(0.056)	(0.068)	(0.075)	(0.216)	(0.066)
aymara	1.787***	1.488***	1.671***	1.614***	1.201***	1.250***	1.101	1.029
	(0.061)	(0.073)	(0.058)	(0.063)	(0.069)	(0.078)	(0.150)	(0.067)
poor		1.366***			1.265***	1.332***	1.662***	1.210***
		(0.050)			(0.052)	(0.075)	(0.182)	(0.070)
urban			0.710***		0.836***	0.817***	0.821**	0.941
			(0.018)		(0.033)	(0.036)	(0.080)	(0.042)
valleys				0.888***	0.816***	0.837***	0.554***	0.834***
				(0.027)	(0.034)	(0.037)	(0.059)	(0.038)
lowlands				0.841***	0.738***	0.785***	0.353***	0.747***
				(0.030)	(0.037)	(0.043)	(0.043)	(0.043)
low water quality						0.943		0.918
						(0.058)		(0.058)
girl						0.874***		0.865***
						(0.031)		(0.032)
female hh head						0.791***		0.761***
						(0.043)		(0.043)
children under 5						1.048		1.053
						(0.034)		(0.034)
hh size						0.930***		0.911***
						(0.009)		(0.009)
probl. distance med. help							0.912	0.989
							(0.083)	(0.042)
probl. med. help							1.255**	1.133***
							(0.112)	(0.046)
knowl. contracept.							0.756*	0.982
							(0.123)	(0.065)
m's educ							0.825***	0.934***
							(0.010)	(0.005)
m's age birth							1.007***	1.002***
							(0.001)	(0.000)
constant	0.007***	0.007***	0.008***	0.007***	0.010***	0.017***	0.092***	0.041***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.031)	(0.007)
Number of obs.	678951	296100	678951	678951	296100	253244	291879	249888
Log-likelihood	-38104.73	-19304.84	-38011.31	-38091.00	-19274.80	-16152.27	-17991.25	-15655.04
aic	76229.46	38631.68	76044.63	76206.00	38577.59	32342.54	36020.49	31358.08
ρ	0.000	0.000	0.000	0.000	0.000	0.000	0.943	0.000

Note: * p<0.10; ** p<0.05; *** p<0.01. Standard errors in parenthesis. Source: DHS 2003, own estimations.
Time dummies measuring six months each included.

Table 5.46: Logit Regression Using Stunting as Dependent Variable - Differentiated by Indigenous Group

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
quechua	0.643***	0.577***	0.733***	0.519***	0.531***	0.280***	0.329***
	(0.035)	(0.035)	(0.036)	(0.038)	(0.042)	(0.044)	(0.047)
aymara	0.537***	0.529***	0.427***	0.259***	0.224***	0.025	-0.000
	(0.046)	(0.046)	(0.050)	(0.052)	(0.058)	(0.058)	(0.064)
poor	0.630***			0.426***	0.400***	0.267***	0.342***
	(0.038)			(0.044)	(0.063)	(0.050)	(0.072)
urban		-0.545***		-0.399***	-0.326***	-0.221***	-0.176***
		(0.032)		(0.037)	(0.042)	(0.042)	(0.047)
valleys			-0.319***	-0.361***	-0.452***	-0.382***	-0.481***
			(0.038)	(0.039)	(0.042)	(0.043)	(0.046)
lowlands			-0.323***	-0.378***	-0.508***	-0.419***	-0.525***
			(0.043)	(0.044)	(0.049)	(0.043)	(0.054)
low water quality					0.175**		0.080
					(0.049)		(0.079)
girl					-0.398***		-0.337***
					(0.034)		(0.037)
female hh head					-0.073		-0.067
					(0.048)		(0.054)
children under 5					0.258***		0.182***
					(0.028)		(0.030)
hh size					0.040***		0.035***
					(0.008)		(0.009)
probl. distance med. help						0.081**	0.055
						(0.037)	(0.041)
probl. med. help						0.019	-0.043
						(0.038)	(0.042)
knowl. contracept.						-0.176**	-0.224***
						(0.069)	(0.079)
m's educ						-0.076***	-0.067***
						(0.004)	(0.005)
m's age birth						-0.010***	-0.011***
						(0.003)	(0.004)
constant	-1.286***	-0.819***	-0.968***	-0.700***	-1.111***	0.221*	-0.039
	(0.022)	(0.030)	(0.033)	(0.046)	(0.070)	(0.120)	(0.140)
Number of obs.	21392	21709	21709	21392	18673	18563	16308
Log-Likelihood	-12494.36	-12676.06	-12777.66	-12389.01	-10450.02	-10435.50	-8912.86
aic	24996.71	25360.11	25565.31	24792.03	20924.04	20895.00	17859.72

Note: * $p<0.10$; ** $p<0.05$; *** $p<0.01$, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.47: Logit Regression Using DPT/Polio Vaccinations as Dependent Variable - Differentiated by Indigenous Group

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
quechua	-0.220***	-0.222***	-0.229***	-0.245***	-0.234***	-0.146	-0.136
	(0.080)	(0.081)	(0.082)	(0.088)	(0.088)	(0.099)	(0.100)
aymara	-0.312***	-0.298***	-0.319***	-0.342***	-0.332***	-0.239*	-0.237*
	(0.114)	(0.113)	(0.123)	(0.126)	(0.125)	(0.138)	(0.138)
poor	-0.003			-0.024	-0.008	0.059	0.062
	(0.091)			(0.105)	(0.105)	(0.120)	(0.120)
urban		-0.028		-0.045	-0.046	-0.098	-0.096
		(0.072)		(0.083)	(0.083)	(0.091)	(0.092)
valleys			-0.013	-0.012	-0.004	-0.027	-0.017
			(0.085)	(0.086)	(0.086)	(0.091)	(0.091)
lowlands			-0.054	-0.050	-0.016	-0.083	-0.053
			(0.091)	(0.093)	(0.093)	(0.101)	(0.102)
girl					0.085		0.068
					(0.067)		(0.072)
female hh head					-0.128		-0.160
					(0.096)		(0.107)
children under 3					-0.141*		-0.173*
					(0.084)		(0.094)
hh size					-0.018		-0.017
					(0.015)		(0.016)
probl. distance med. help						-0.034	-0.014
						(0.078)	(0.078)
probl. med. help						-0.195**	-0.175**
						(0.084)	(0.084)
knowl. contracept.						0.293	0.296
						(0.191)	(0.191)
m's educ						0.013	0.013
						(0.009)	(0.010)
m's age birth						-0.016**	-0.016**
						(0.006)	(0.007)
constant	-2.080***	-2.064***	-2.056***	-2.019***	-1.792***	-1.931***	-1.671***
	(0.043)	(0.067)	(0.071)	(0.101)	(0.148)	(0.281)	(0.318)
Number of obs.	9853	10022	10022	9853	9853	8573	8573
Log-Likelihood	-3239.41	-3297.18	-3297.06	-3239.10	-3234.31	-2791.75	-2787.02
aic	6486.83	6602.35	6604.12	6492.21	6490.62	5607.51	5606.05

Note: * p<0.10; ** p<0.05; *** p<0.01, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.48: Logit Regression Using BCG Vaccinations as Dependent Variable - Differentiated by Indigenous Group

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
quechua	-0.231***	-0.078	-0.251***	-0.214**	-0.127	0.287***	0.304***
	(0.081)	(0.083)	(0.084)	(0.089)	(0.090)	(0.105)	(0.107)
aymara	-0.757***	-0.648***	-0.668***	-0.678***	-0.603***	-0.237	-0.242
	(0.132)	(0.132)	(0.144)	(0.143)	(0.144)	(0.162)	(0.163)
poor	-0.652***						
	(0.111)						
urban		0.374***		-0.426***	-0.459***	-0.071	-0.141
		(0.079)		(0.123)	(0.126)	(0.145)	(0.148)
valleys			0.178**	0.239***	0.303***	0.149	0.242**
			(0.090)	(0.089)	(0.090)	(0.098)	(0.099)
lowlands			0.010	-0.011	0.160*	-0.093	0.118
			(0.096)	(0.096)	(0.097)	(0.108)	(0.110)
girl					-0.130*		-0.172**
					(0.070)		(0.077)
female hh head					-0.161		-0.243**
					(0.098)		(0.112)
children under 3					1.211***		1.314***
					(0.042)		(0.049)
hh size					-0.338***		-0.340***
					(0.022)		(0.027)
probl. distance med. help						-0.238***	-0.146*
						(0.083)	(0.085)
probl. med. help						-0.355***	-0.181*
						(0.095)	(0.097)
knowl. contracept.						0.013	0.262
						(0.230)	(0.233)
m's educ						0.100***	0.088***
						(0.009)	(0.010)
m's age birth						0.002	0.013**
						(0.006)	(0.007)
constant	-3.693***	-4.591***	-4.363***	-4.083***	-3.452***	-4.769***	-4.791***
	(0.043)	(0.077)	(0.077)	(0.107)	(0.150)	(0.299)	(0.318)
Number of obs.	45448	75011	75011	45448	45448	38935	38935
Log-Likelihood	-4247.54	-4769.58	-4778.75	-4232.21	-3909.26	-3452.31	-3179.18
aic	8503.09	9547.17	9567.51	8478.42	7840.52	6928.62	6390.36

Note: * p<0.10; ** p<0.05; *** p<0.01, Standard errors in parenthesis. Source: DHS 2003, own estimations

Table 5.49: Logit Regression Using Measles Vaccinations as Dependent Variable - Differentiated by Indigenous Group

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
quechua	-0.157*	-0.176*	-0.113	-0.139	-0.130	-0.042	-0.040
	(0.093)	(0.095)	(0.098)	(0.105)	(0.106)	(0.119)	(0.120)
aymara	-0.458***	-0.466***	-0.365***	-0.390**	-0.377**	-0.324*	-0.318*
	(0.145)	(0.143)	(0.156)	(0.160)	(0.159)	(0.175)	(0.175)
poor	0.057			-0.013	0.000	0.147	0.148
	(0.107)			(0.122)	(0.123)	(0.141)	(0.141)
urban		-0.095		-0.094	-0.100	-0.107	-0.106
		(0.085)		(0.098)	(0.098)	(0.109)	(0.109)
valleys			0.073	0.058	0.063	0.039	0.044
			(0.103)	(0.103)	(0.103)	(0.110)	(0.110)
lowlands			0.144	0.141	0.171	0.082	0.103
			(0.109)	(0.111)	(0.112)	(0.121)	(0.123)
girl					-0.043		-0.053
					(0.079)		(0.086)
female hh head					0.011		0.001
					(0.109)		(0.121)
children under 3					-0.132		-0.182*
					(0.098)		(0.111)
hh size					-0.019		-0.012
					(0.017)		(0.020)
probl. distance med. help						0.001	0.027
						(0.093)	(0.094)
probl. med. help						-0.163*	-0.155
						(0.098)	(0.099)
knowl. contracept.						0.216	0.221
						(0.221)	(0.221)
m's educ						0.012	0.010
						(0.011)	(0.011)
m's age birth						-0.006	-0.007
						(0.007)	(0.008)
constant	-2.483***	-2.409***	-2.569***	-2.490***	-2.224***	-2.620***	-2.316***
	(0.050)	(0.080)	(0.088)	(0.121)	(0.171)	(0.339)	(0.375)
Number of obs.	9744	9911	9911	9744	9744	8470	8470
Log-Likelihood	-2503.85	-2543.93	-2543.68	-2502.50	-2500.09	-2149.53	-2147.18
aic	5015.69	5095.86	5097.36	5019.01	5022.19	4323.06	4326.36

Note: * p<0.10; ** p<0.05; *** p<0.01, Standard errors in parenthesis. Source: DHS 2003, own estimations

Bibliography

Abadian, S. (1996). Women's Autonomy and Its Impact on Fertility. *World Development*, 24(12):1793–1809.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The Colonial Origins of Comparative Advantage. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics*, 117:1231–1294.

Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2008). Income and Democracy. *American Economic Review*, 98(3):808–842.

Ades, A. and Tella, R. D. (1997). The New Economics of Corruption: A Survey and Some New Results. *Political Studies*, 45(3):496–515.

Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, New York.

Agresti, A. (1990). *Categorical Data Analysis*. John Wiley & Sons, New York.

Al-Marhubi, F. A. (2005). Openness and Governance: Evidence Across Countries. *Oxford Development Studies*, 33(3):453–471.

Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N., and Gangadharan, L. (2009). Gender, Culture, and Corruption: Insights from an Experimental Analysis. *Southern Economic Journal*, 75(3):663–680.

Alderman, H., Behrman, J. R., Ross, D. R., and Sabot, R. (1996). Decomposing the Gender Gap in Cognitive Skills in a Poor Rural Economy. *Journal of Human Resources*, 31(1):229–254.

Alesina, A., Baquir, R., and Easterly, W. (1999). Public Goods and Ethnic Divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2):155–194.

Alesina, A. and Ferrara, E. L. (2005). Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, 43(3):762–800.

Alesina, A. and Rodrik, D. (1994). Distributive Politics and Economic Growth. *The Quarterly Journal of Economics*, 109(2):465–490.

Alhassan-Alolo, N. (2007). Gender and Corruption: Testing the New Consensus. *Public Administration and Development*, 27:227–237.

Alkire, S. and Foster, J. E. (2008). Counting and Multidimensional Poverty Measurement. OPHI Working Paper 7, Oxford Poverty and Human Development Initiative, Queen Elizabeth House, University of Oxford.

Allison, P. D. (1982). Discrete-Time Methods for the Analysis of Event Histories. *Sociological Methodology*, 13:61–98.

Andreoni, J. and Vesterlund, L. (2001). Which Is The Fair Sex? Gender Differences In Altruism. *The Quarterly Journal of Economics*, 116(1):293–312.

Atkinson, A. B. (1970). On the Measurement of Inequality. *Journal of Economic Theory*, 2(3):244–263.

Bardhan, K. and Klasen, S. (1999). UNDP's Gender-Related Indices: A Critical Review. *World Development*, 27:985–1010.

Barro, R. J. (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, A. M. (2002). Why Does Education Lead to Lower Fertility? A Critical Review of Some of the Possibilities. *World Development*, 30(10):1779–1790.

Baum, M. A. and Lake, D. A. (2003). The Political Economy of Growth: Democracy and Human Capital. *American Journal of Political Science*, 47(2):333–347.

Becker, G. S. (1981). *A Treatise on the Family*. Harvard University Press, Cambridge, enlarged 1991 edition.

Besley, T. and Burgess, R. (2002). The Political Economy of Government Responsiveness: Theory and Evidence from India. *Quarterly Journal of Economics*, 117:1415–1451.

Besley, T. and Kudamatsu, M. (2006). Health and Democracy. *AEA Papers and Proceedings*, 96(2):313–315.

Bjørnskov, C., Dreher, A., and Fischer, J. A. (2009). On Gender Inequality and Life Satisfaction: Does Discrimination Matter? Mimeo.

Bloom, S. S., Wypij, D., and Das Gupta, M. (2001). Dimensions of Women's Autonomy and the Influence on Maternal Health Care Utilization in a North Indian City. *Demography*, 38(1):67–78.

Boix, C. (2001). Democracy, Development, and the Public Sector. *American Journal of Political Science*, 45(1):1–17.

Bourguignon, F. and Verdier, T. (2000). Oligarchy, Democracy, Inequality and Growth. *Journal of Development Economics*, 62:285–313.

Braunstein, E. (2007). The Efficiency of Gender Equity in Economic Growth: Neoclassical and Feminist Approaches. GEM-IWG Working Paper 07-4, The Gender and Macro International Working Group.

Braveman, P. and Tarimo, E. (2002). Social Inequalities in Health Within Countries: Not Only an Issue for Affluent Nations. *Social Science & Medicine*, 54:1621–1635.

Central Intelligence Agency (2009). The World Factbook. Electronic publication. https://www.cia.gov/library/publications/the-world-factbook/.

Chattopadhyay, R. and Duflo, E. (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica*, 72(5):1409–1443.

Colditz, G., Berkey, C. S., Mosteller, F., Brewer, T. F., Wilson, M. E., Burdick, E., and Fineberg, H. F. (1995). The Efficacy of Bacillus Calmette-Guerin Vaccination of Newborns and Infants in the Prevention of Tubercolosis: Meta-Analyses of the Published Literature. *Pediatrics*, 96(1):29–35.

Collier, P. (2001). Implications of Ethnic Diversity. *Economic Policy Analysis*, 16(32):129–166.

Correlates of War 2 Project (2003). Colonial / Dependency Contiguity Data V3.0. Electronic publication. http://correlatesofwar.org/.

Davidson, R. and MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. Oxford Univ. Press, New York.

De Soysa, I. and Jütting, J. (2007). Informal Institutions and Development: How They Matter and What Makes Them Change. In Jütting, J., Drechsler, D., Bartsch, S., and de Soysa, I., editors, *Informal Institutions. How Social Norms Help or Hinder Development*, pages 29–43. OECD Development Centre, Paris.

Deaton, A. (1997). *The Analysis of Household Surveys*. The John Hopkins University Press, Baltimore, Maryland.

Dijkstra, A. G. (2002). Revisiting UNDP's GDI and GEM: Towards an Alternative. *Social Indicators Research*, 57:301–338.

Dollar, D., Fisman, R., and Gatti, R. (2001). Are Women Really The "fairer" Sex? Corruption and Women in Government. *Journal of Economic Behavior & Organization*, 46(4):423–429.

Dollar, D. and Kraay, A. (2002). Growth is Good for the Poor. *Journal of Economic Growth*, 7:195–225.

Duflo, E. (2005). Gender Equality in Development. BREAD Policy Paper 011, Bureau for Research in Economic Analysis of Development.

Dyson, T. and Moore, M. (1983). On kinship structure, female autonomy and demographic behavior in India. *Population and Economic Review*, 9(1):35–60.

Eagly, A. H. and Crowley, M. (1986). Gender and Helping Behavior: A Meta-Analytic Review of the Social Psychological Literature. *Psychological Bulletin*, 100(3):283 – 308.

Eckel, C. C. and Grossman, P. J. (1998). Are Women Less Selfish Than Men? Evidence From Dictator Experiments. *The Economic Journal*, 108(448):726–735.

Economic Commission for Africa (2004). *The African Gender and Development Index*. ECA, Addis Ababa.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.

Emerson, P. M. and Souza, A. P. (2007). Child Labor, School Attendance, and Intrahousehold Gender Bias in Brazil. *The World Bank Economic Review*, 21(2):301–316.

Folbre, N. (1997). Gender Coalitions: Extrafamily Influences on Intrafamily Inequality. In Haddad, L., Hoddinott, J., and Alderman, H., editors, *Intrahousehold Resource Allocation*. The Johns Hopkins Press, Baltimore, MD.

Foster, G. and Williamson, J. (2000). A Review of Current Literature of the Impact of HIV/AIDS on Children in Sub-Saharan Africa. *AIDS*, 14(3):275–284.

Foster, J. E., Greer, J., and Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica*, 52:761–766.

Franco, Á., Álvarez-Dardet, C., and Ruiz, M. T. (2004). Effect of Democracy on Health: Ecological Study. *BMJ (British Medical Journal)*, 329:1421–1423.

Freedom House (2008). Freedom in the World 2008. Technical report, Freedom House. http://www.freedomhouse.org.

Gaspers, D. and van Staveren, I. (2003). Development as Freedom - and as What Else? *Feminist Economics*, 9(2-3):137–161.

Gatti, R. (2004). Explaining Corruption: Are Open Countries Less Corrupt? *Journal of International Development*, 16(6):851–861.

Glaeser, E. L., Ponzetto, G. A. M., and Shleifer, A. (2007). Why Does Democracy Need Education? *Journal of Economic Growth*, 12:77–99.

Glover, S. H., Bumpus, M. A., Logan, J. E., and Ciesla, J. R. (1997). Re-examining the Influence of Individual Values on Ethical Decision Making. *Journal of Business Ethics*, 16:1319–1329.

Goetz, A. M. (2007). Political Cleaners: Women as the New Anti-Corruption Force? *Development and Change*, 38(1):87–105.

Gradstein, M. and Milanovic, B. (2004). Does Liberté = Égalité? A Survey of the Empirical Links Between Democracy and Inequality with some Evidence on the Transition Economies. *Journal of Economic Surveys*, 18(4):515–537.

Greenacre, M. (2007). *Correspondence Analysis in Practice*. Interdisciplinary Statistics. Chapman and Hall, Boca Raton, second edition.

Grimm, M., Harttgen, K., Klasen, S., and Misselhorn, M. (2008). A Human Development Index by Income Groups. *World Development*, 36:2527–2546.

Grün, C. and Klasen, S. (2008). Growth, Inequality, and Well-Being: Comparisons Across Space and Time. *Oxford Economic Papers*, 60(2):212–236.

Haddad, L. and Hoddinott, J. (1994). Women's Income and Boy-Girl Anthropometric Status in Cote D'Ivoire. *World Development*, 22(4):543–553.

Hall, G., Layton, H. M., and Shapiro, J. (2006). Introduction: The Indigenous Peoples' Decade in Latin America. In Hall, G. and Patrinos, H. A., editors, *Indigenous Peoples, Poverty and Human Development in Latin America*, pages 1–24. Palgrave Macmillan, New York.

Hall, P. A. and Taylor, R. C. R. (1996). Political Science and the Three New Institutionalisms. Technical report, Köln: Max-Planck Institut für Gesellschaftsforschung.

Hamilton, L. C. (1992). How Robust is Robust Regression? *Stata Technical Bulletin*, 1(2).

Hatt, L. E. and Waters, H. R. (2006). Determinants of Child Morbidity in Latin America: A Pooled Analysis of Interactions Between Parental Education and Economic Status. *Social Science and Medicine*, 62:375–386.

Hausmann, R., Tyson, L. D., and Zahidi, S. (2007). *The Global Gender Gap Report 2007*. World Economic Forum, Geneva.

Heaton, T. B. and Forste, R. (2003). Rural/Urban Differences in Child Growth and Survival in Bolivia. *Rural Sociology*, 68(3):410–433.

Hill, M. A. and King, E. M. (1995). Women's Education and Economic Well-Being. *Feminist Economics*, 1(2):21 –46.

Hindin, M. J. (2000). Women's Autonomy, Women's Status and Fertility-Related Behavior in Zimbabwe. *Population Research and Policy Review*, 19:255–282.

Holzmann, H., Vollmer, S., and Weisbrod, J. (2008). Twin Peaks or Three Components? - Analyzing the World's Cross-Country Distribution of Income. Working Paper, University of Göttingen.

Hosmer, D. W., Lemeshow, S., and May, S. (2008). *Applied Survival Analysis: Regression Modeling of Time to Event Data*. Wiley, New York, 2nd edition.

Hotez, P. J., Bottazzi, M. E., Franco-Paredes, C., Ault, S. K., and Periago, M. R. (2008). The Neglected Tropical Diseases of Latin America and the Caribbean: A Review of Disease Burden and Distribution and a Roadmap for Control and Elimination. *PLoS Negl Trop Dis*, 2(9).

Inglehart, R., Norris, P., and Welzel, C. (2002). Gender Equality and Democracy. *Comparative Sociology*, 1(3-4):321–345.

Inglehart, R. and Welzel, C. (2005). *Modernization, Cultural Change and Democracy. The Human Development Sequence*. Cambridge University Press, New York.

Jenkins, S. P. (2005a). Lesson 7. Unobserved Heterogeneity ('frailty'). Essex Summer School Course Survival Analysis and EC968. Part II: Introduction to the Analysis of Spell Duration Data. http://www.iser.essex.ac.uk/files/teaching/stephenj/ec968/pdfs/ec968st7.pdf. Institute for Social and Economic Research. University of Essex.

Jenkins, S. P. (2005b). Survival Analysis. unpublished manuscript, Institute for Social and Economic Research, University of Essex. Downloadable from http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/pdfs/ec968lnotesv6.pdf.

Jolliffe, I. T. (1986). *Principal component analysis*. Springer, New York, NY.

Jütting, J., A. L. and Morrisson, C. (2009). Why Do so Many Women End Up in Bad Jobs? A Cross-Country Assessment. Technical report, OECD Development Centre. Mimeo.

Jütting, J. and Morrisson, C. (2005). Changing Social Institutions to Improve the Status of Women in Developing Countries. Technical Report Policy Brief 27, OECD Development Centre.

Jütting, J., Morrisson, C., Dayton-Johnson, J., and Drechsler, D. (2008). Measuring Gender (In)Equality: The OECD Gender, Institutions and Development Data Base. *Journal of Human Development*, 9(1):65–86.

Kakwani, N., Wagstaff, A., and Doorslaer, E. V. (1997). Socioeconomic Inequalities in Health: Measurement, Computation, and Statistical Inference. *Journal of Econometrics*, 77:87–103.

Kakwani, N. C. (1984). Issues in Measuring Poverty. In Basmann, R. L. and Rhodes, J. G. F., editors, *Advances in Econometrics*, volume 3. JAI Press, Greenwich, CT and London.

Kanbur, R. (2003). Education, Empowerment and Gender Inequalities. In Stern, N. and Pleskovic, B., editors, *The New Reform Agenda / Annual World Bank Conference on Development Economics*. Oxford University Press for The World Bank, Oxford.

Karlsson, M., Nilsson, T., Lyttkens, C. H., and Leeson, G. (2010). Income Inequality and health: Importance of a cross-country perspective. *Social Science & Medicine*, (70):875–885.

Kauermann, G. and Carroll, R. J. (2001). A Note on the Efficiency of Sandwich Covariance Matrix Estimation. *Journal of the American Statistical Association*, 96(456):1387–1396.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2007). Governance Matters VI: Gover Nance Indicators for 1996-2006. *World Bank Policy Research Working Paper No. 4280*, (4280).

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2008). Governance Matters VII: Aggregate and Individual Governance Indicators, 1996-2007. World Bank Policy Research Working Paper 4654, World Bank.

Kawachi, I., Kennedy, B. P., Lochner, K., and Smith, D. P. (1997). Social Capital, Income Inequality, and Mortality. *American Journal of Public Health*, 87(9):1491–1498.

Kazianga, H. and Klonner, S. (2009). The Intra-Household Economics of Polygyny: Fertility and Child Mortality in Rural Mali. MPRA Paper 12859, University Library of Munich, Germany.

Keefer, P. (2005). Democratization and Clientelism: Why are Young Democracies Badly Governed? *World Bank Policy Research Working Paper No. 3594*.

Keefer, P. and Khemani, S. (2005). Democracy, Public Expenditure, and the Poor: Understanding Political Incentives for Providing Public Services. *The World Bank Research Observer*, 20(1):1–27.

King, E. M. and Hill, M. A. (1993). *Women's Education in Developing Countries*. John Hopkins University Press, Baltimore.

Klasen, S. (2002). Low Schooling for Girls, Slower Growth for All? *World Bank Economic Review*, 16(3):345–373.

Klasen, S. (2004). In Search of the Holy Grail: How to Achieve Pro Poor Growth? In Tungodden, B., Stern, N., and Kolstad, I., editors, *Toward Pro Poor Policies - Aid, Institutions, and Globalization*, pages 63–94. Oxford University Press, New York.

Klasen, S. (2006). UNDP's Gender-Related Measures: Some Conceptual Problems and Possible Solutions. *Journal of Human Development*, 7(2):243–274.

Klasen, S. (2007). Gender-Related Indicators of Well-Being. In McGillivray, M., editor, *Human Well-Being: Concept and Measurement*, Studies in Development Economics and Policy, chapter 7, pages 167–192. Palgrave Macmillan, New York, NY.

Klasen, S. and Lamanna, F. (2009). The Impact of Gender Inequality in Education and Employment on Economic Growth in Developing Countries: Updates and Extensions. *Feminist Economics*, 15(3):91–132.

Klasen, S. and Schüler, D. (2009). Reforming the Gender-Related Development Index (GDI) and the Gender Empowerment Measure (GEM): Some Specific Proposals. Ibero-America Institute for Economic Research (IAI) Discussion Papers 186, University of Goettingen.

Klasen, S. and Wink, C. (2003). Missing Women: Revisiting the Debate. *Feminist Economics*, 9:263–300.

Kolenikov, S. and Angeles, G. (2004). The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomics Indices. CPC/MEASURE Working paper WP-04-85, Carolina Population Center.

Kolenikov, S. and Angeles, G. (2009). Socioeconomic Status Measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth*, 55(1):128–165.

Kraay, A. (2006). When is Growth Pro-Poor? Evidence from a Panel of Countries. *Journal of Development Economics*, 80:198–227.

Lahiri, S. and Self, S. (2007). Gender Bias in Education: The Role of Inter-Household Externality, Dowry and Other Social Institutions. *Review of Development Economics*, 11(4):591–606.

Lake, D. A. and Baum, M. A. (2001). The Invisible Hand of Democracy. Political Control and the Provision of Public Services. *Comparative Political Studies*, 34:587–620.

Lambsdorff, J. G. (2006). Measuring Corruption - The Validity and Precision of Subjective Indicators (CPI). In Sampford, C., Shacklock, A., Connors, C., and Galtung, F., editors, *Measuring Corruption*, Law, Ethics and Governance Series, chapter 5, pages 81–99. Ashgate, Aldershot.

Larrea, C. and Freire, W. (2002). Social Inequality and Child Malnutrition in Four Andean Countries. *Pan American Journal of Public Health*, 11(5/6):356–364.

Layton, H. M. and Patrinos, H. A. (2006). Estimating the Number of Indigenous Peoples in Latin America. In Hall, G. and Patrinos, H. A., editors, *Indigenous Peoples, Poverty and Human Development in Latin America*, pages 25–39. Palgrave Macmillan, New York.

Liberato, A. S., Pomeroy, C., and Fennell, D. (2006). Well-Being Outcomes in Bolivia: Accounting for the Effects of Ethnicity and Regional Location. *Social Indicators Research*, 76:233–262.

Lindelow, M. (2006). Sometimes More Equal Than Others: How Health Inequalities Depend on the Choice of Welfare Indicator. *Health Economics*, 15:263–279.

Lipset, S. M. (1959). Some Social Requisites of Democray: Economic Development and Political Legitimacy. *The American Political Science Review*, 53(1):69–105.

Long, J. S. and Ervin, L. H. (2000). Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician*, 54(3):217–224.

Lopez, A. D., Mathers, C. D., Ezzati, M., Jamison, D. T., and Murray, C. J. (2006). *Global Burden of Disease and Risk Factors*. The World Bank.

Lopez-Claros, A. and Zahidi, S. (2005). *Women's Empowerment: Measuring the Global Gender Gap*. World Economic Forum, Davos.

Lundberg, S. and Pollak, R. A. (1993). Separate Spheres Bargaining and the Marriage Market. *Journal of Political Economy*, 101(6):988–1010.

Lundberg, S. and Pollak, R. A. (2008). Family Decision Making. In Durlauf, S. N. and Blume, L. E., editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke, second edition.

Macdonald, M. (2006). Muslim Women and the Veil. *Feminist Media Studies*, 6(1):7–23.

Maitra, P. (2004). Parental Bargaining, Health Inputs and Child Mortality in India. *Journal of Health Economics*, 23(2):259–291.

Manser, M. and Brown, M. (1980). Marriage and Household Decision Making: A Bargaining Analysis. *International Economic Review*, 21:31–44.

Marmot, M. (2005). Social Determinants of Health Inequalities. *The Lancet*, 365:1099 – 1104.

Marshall, M. G. and Jaggers, K. (2005). Polity IV Project. Dataset User's Manual. Technical report, Center for Global Policy, George Mason University.

Marshall, M. G. and Jaggers, K. (2009). Polity IV Project: Political Regime Characteristics and Transitions, 1800-2008. Electronic publication. http://www.systemicpeace.org/polity/polity4.htm.

Mauro, P. (1995). Corruption and Growth. *The Quarterly Journal of Economics*, 110(3):681–712.

Mayer-Foulkes, D. and Larrea, C. (2005). Racial and Ethnic Health Inequalities: Bolivia, Brazil, Guatemala, and Peru. CIDE Working Paper 333, Mexico, DF.

McElroy, M. B. (1990). The Empirical Content of Nash-Bargained Household Behavior. *Journal of Human Resources*, 25:559–583.

McElroy, M. B. and Horney, M. J. (1981). Nash Bargained Household Decisions. *International Economic Review*, 22:333–349.

McGillivray, M. and White, H. (1993). Measuring Development? The UNDP's Human Development Index. *Journal of International Development*, 5(2):183–192.

McGuire, M. C. and Olson, M. (1996). The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force. *Journal of Economic Literature*, 34(1):72–96.

Meltzer, A. H. and Richard, S. F. (1981). A Rational Theory of the Size of Government. *Journal of Political Economy*, 89(5):914–927.

Miguel, E. and Gugerty, M. K. (2005). Ethnic Diversity, Social Sanctions, and Public Goods in Kenya. *Journal of Public Economics*, 89(11/12):2325–2368.

Milallos, M. T. R. (2007). Muslim Veil as Politics: Political Autonomy, Women and Syariah Islam in Aceh. *Contemporary Islam*, 1(3):289–301.

Minier, J. (1998). Democracy and Growth: Alternative Approaches. *Journal of Economic Growth*, 3(3):241–266.

MMWR Weekly (2000). Progress Toward Interrupting Indigenous Measles Transmission - Region of the Americas, January 1999-September 2000. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm4943a4.htm (date of access April 20, 2010).

Mohtadi, H. and Roe, T. L. (2003). Democracy, rent-seeking, public spending and growth. *Journal of Public Economics*, 87:445–466.

Morales, R., Aguilar, A. M., and Calzadilla, A. (2004). Geography and culture matter for malnutrition in Bolivia. *Economics and Human Biology*, 2:373–389.

Morrisson, C. and Jütting, J. P. (2005). Women's Discrimination in Developing Countries: A New Data Set for Better Policies. *World Development*, 33(7):1065–1081.

Mukherjee, C., White, H., and Wuyts, M. (1998). *Econometrics and Data Analysis for Developing Countries*. Routledge, London.

Munda, G. and Nardo, M. (2005a). Constructing Consistent Composite Indicators: The Issue of Weights. Technical Report EUR 21834 EN, European Commission.

Munda, G. and Nardo, M. (2005b). Non-Compensatory Composite Indicators for Ranking Countries: A Defensible Setting. Technical Report EUR 21833 EN, European Commission.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. (2005). Handbook on Constructing Composite Indicators: Methodology and User Guide. Technical Report 2005/3, OECD.

Navia, P. and Zweifel, T. (2003). Democracy, Dictatorship, and Infant Mortality Revisited. *Journal of Democracy*, 14(3):90–103.

Nenadić, O. (2007). *An Implementation of Correspondence Analysis in R and its Application in the Analysis of Web Usage*. PhD thesis, University of Goettingen, Goettingen.

North, D. C. (1990). *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, New York.

North, D. C. (2001). Institutions. *Journal of Economic Perspectives*, 5(1):97–112.

Nussbaum, M. C. (2003). Capabilities as Fundamental Entitlements: Sen and Social Justice. *Feminist Economics*, 9(2-3):33–59.

O'Donnell, O., Doorslaer, E. V., Wagstaff, A., and Lindelow, M. (2008). *Analyzing Health Equity Using Household Survey Data*. The World Bank, Washington, D.C.

Olson, M. (1993). Dictatorship, Democracy, and Development. *The American Political Science Review*, 87(3):567–576.

PAHO (2007). Bolivia. In *Health in the Americas 2007, Volume II - Countries*, volume II - Countries of *Health in the Americas 2007*, pages 114–129. PAHO, Washington, D.C. PAHO Scientific and Technical Publication No. 622.

Papaioannou, E. and Siourounis, G. (2008). Democratisation and Growth. *The Economic Journal*, 118:1520–1551.

Pardo, I. (2004). *Between Morality And The Law: Corruption, Anthropology And Comparative Society*. Ashgate, Aldershot.

Pasqua, S. (2005). Gender Bias in Parental Investments in Children's Education: A Theoretical Analysis. *Review of Economics of the Household*, 3(3):291–314.

Pérez-Cueto, F. J. A., Botti, A. B., and Verbeke, W. (2009). Prevalence of Overweight in Bolivia: Data on Women and Adolescents. *obesity reviews*, pages 1–5.

Persson, T. (2002). Do Political Institutions Shape Economic Policy? *Econometrica*, 70(3):883–905.

Persson, T., Roland, G., and Tabellini, G. (2000). Comparative Politics and Public Finance. *The Journal of Political Economy*, 108(6):1121–1161.

Persson, T. and Tabellini, G. (2000). *Political Economics. Explaining Economic Policy*. MIT Press, Cambridge, Massachusetts, London.

Persson, T. and Tabellini, G. (2006). Democracy and Development: The Devil in Details. *AEA Papers and Proceedings*, 96(2):319–320.

Persson, T. and Tabellini, G. (2007a). Democratic Capital: The Nexus of Political and Economic Change.

Persson, T. and Tabellini, G. (2007b). The Growth Effect of Democracy: Is It Heterogeneous and How Can It Be Estimated? *NBER Working Paper No. 13150*.

Peters, G. B. (2005). *Institutional Theory in Political Science*. Continuum, London, New York.

Pew Forum on Religion and Public Life (2009). Mapping the Global Muslim Population: A Report on the Size and the Distribution of the World's Muslim Population.

Pollak, R. A. (2003). Gary Becker's Contributions to Family and Household Economics. *Review of Economics of the Household*, 1:111–141.

Pollak, R. A. (2007). Bargaining Around the Hearth. NBER Working Papers 13142, National Bureau of Economic Research.

Pozo, W. J., Casazola, F. L., and Aguilar, E. Y. (2006). Bolivia. In Hall, G. and Patrinos, A., editors, *Indigenous Peoples, Poverty and Human Development in Latin America*, pages 40–66. Palgrave Macmillan, New York.

Pritchett, L. and Summers, L. H. (1996). Wealthier is Healthier. *The Journal of Human Resources*, 31(4):841–868.

Rasul, I. (2008). Household Bargaining over Fertility: Theory and Evidence from Malaysia. *Journal of Development Economics*, 86:215–241.

Ravallion, M. (1994). *Poverty Comparisons: A Guide to Concepts and Methods*. Harwood Academic, Chur.

Rivas, M. F. (2008). An Experiment on Corruption and Gender. ThE Papers 08/10, Department of Economic Theory and Economic History of the University of Granada.

Rizzo, H., Abdel-Latif, A.-H., and Meyer, K. (2007). The Relationship Between Gender Equality and Democracy: A Comparison of Arab Versus Non-Arab Muslim Societies. *Sociology*, 4(6):1151–1170.

Ross, M. (2006). Is Democracy Good for the Poor. *American Journal of Political Science*, 50(4):860–874.

Safaei, J. (2006). Is Democracy Good for Health? *International Journal of Health Services*, 36(4):767–786.

Sahn, D. E. and Younger, S. D. (2006). Changes in Inequality and Poverty in Latin America: Looking Beyond Income to Health and Education. *Journal of Applied Economics*, 0:215–234.

Saleem, S. and Bobak, M. (2005). Women's Autonomy, Education and Contraception Use in Pakistan: A National Study. *Reproductive Health*, 2(8).

Schüler, D. (2006). The Uses and Misuses of the Gender-Related Development Index and the Gender Empowerment Measure: A Review of the Literature. *Journal of Human Development*, 7(2):161–182.

Schultz, T. P. (1990). Testing the Neoclassical Model of Family Labor Supply and Fertility. *Journal of Human Resources*, 25(4):599–634.

Schultz, T. P. (2002). Why Governments Should Invest More to Educate Girls. *World Development*, 30(2):207–225.

Schultz, T. P. (2004). School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program. *Journal of Development Economics*, 74:199– 250.

Seebens, H. (2008). *The Economics of Gender and the Household in Developing Countries*. Peter Lang, Frankfurt.

Sen, A. K. (1988). The Concept of Development. In Chenery, H. and Srinivasan, T., editors, *Handbook of Development Economics*, pages 9–26. North Holland, Amsterdam.

Sen, A. K. (1991). Capability and Well-Being. In Nussbaum, M. and Sen, A. K., editors, *The Quality of Life*. Clarendon Press, Oxford.

Sen, A. K. (1992). Missing Women. *British Medical Journal*, 304:586–7.

Sen, A. K. (1999a). Democracy as a Universal Value. *Journal of Democracy*, 10(3):3–17.

Sen, A. K. (1999b). *Development as Freedom*. Oxford University Press, Oxford.

Sen, A. K. (2000a). A Decade of Human Development. *Journal of Human Development*, 1(1):17–23.

Sen, A. K. (2000b). Social Exclusion: Concept, Application, and Scrutiny. *Social Development Papers*, (1). http://www.adb.org/documents/books/ social_exclusion/Social_exclusion.pdf, (date of access: May 2010).

Sen, A. K. (2003). Development as Capability Expansion. In Fukuda-Parr, S. and Kumar, A. S., editors, *Readings in Human Development*, pages 41–58. Oxford University Press, New Delhi,New York.

Shorrocks, A. F. and Foster, J. E. (1987). Transfer Sensitive Inequality Measures. *Review of Economic Studies*, 54(1):485–497.

Shroff, M., Griffiths, P., Adair, L., Suchindran, C., and Bentley, M. (2009). Maternal Autonomy is Inversely Related to Child Stunting in Andhra Pradesh, India. *Maternal and Child Nutrition*, 5:64–74.

Smith, L. C., Ramakrishnan, U., Ndiaye, A., Haddad, L., and Martorell, R. (2002). *The Importance of Women's Status for Child Nutrition in Developing Countries*. Research Report 131. International Food Policy Research Institute, Washington, DC.

Social Watch (2005). *Roars and Whispers Gender and Poverty: Promises versus Action*. Social Watch, Montevideo.

Song, L., Appleton, S., and Knight, J. (2006). Why Do Girls in Rural China Have Lower School Enrollment? *World Development*, 34(9):1639–1653.

Stasavage, D. (2005). Democracy and Education Spending in Africa. *American Journal of Political Science*, 49(2):343–358.

Stephens, C., Porter, J., Nettleton, C., and Willis, R. (2006). Disappearing, Displaced, and Undervalued: A Call to Action for Indigenous Health Worldwide. *The Lancet*, 367:2019–2028.

Subramanian, S. (2007). Indicators of Inequality and Poverty. In McGillivray, M., editor, *Human Well-Being: Concept and Measurement*, Studies in Development Economics and Policy, chapter 6, pages 135–166. Palgrave Macmillan, New York, NY.

Sung, H.-E. (2003). Fairer Sex or Fairer System? Gender and Corruption Revisited. *Social Forces*, 82(2):703–723.

Svensson, J. (2005). Eight Questions About Corruption. *Journal of Economic Perspectives*, 19(3):19–42.

Swamy, A., Knack, S., Lee, Y., and Azfar, O. (2001). Gender and Corruption. *Journal of Development Economics*, 64(1):25–55.

Tavares, J. and Wacziarg, R. (2001). How Democracy Affects Growth. *European Economic Review*, 45:1341–1378.

Telles, E. E. (2007). Race, Ethnicity and the United Nation's Millennium Development Goals in Latin America. *Latin American and Caribbean Ethnic Studies*, 2(2):185–200.

The PRS Group (2008). Bolivia. Country report, The PRS Group, Inc.

Thomas, D. (1997). Incomes, Expenditures and Health Outcomes: Evidence on Intrahousehold Resource Allocation. In Haddad, L., Hoddinott, J., and Alderman, H., editors, *Intrahousehold Resource Allocation in Developing Countries: Models, Methods and Policy*. The Johns Hopkins Press, Baltimore, MD.

Treisman, D. (2007). What Have We Learned About the Causes of Corruption from Ten Years of Cross-National Empirical Research? *Annu. Rev. Polit. Sci.*, 10:211–244.

Tripp, A. M. (2001). Women's Movements and Challenges to Neopatrimonial Rule: Preliminary Observations from Africa. *Development and Change*, 32:33–54.

Tsai, M.-C. (2006). Does Political Democracy Enhance Human Development in Developing Countries. *American Journal of Economics and Sociology*, 65(2):233–268.

UDAPE and OPS (2004). Caracterización de la Exclusión En Salud En Bolivia. Technical report.

UNAIDS (2008). Report on the Global AIDS Epidemic. Technical report, Geneva.

United Nations (2009). The Millennium Development Goals Report. Technical report, United Nations, New York.

United Nations (2010). Keeping the promise: a forward-looking review to promote an agreed action agenda to achieve the Millennium Development Goals by 2015. Report of the Secretary-General A/64/665, United Nations.

United Nations Development Programme (1995). *Human Development Report*. Oxford University Press, New York.

United Nations Development Programme (2006). *Human Development Report*. Oxford University Press, New York.

Wagstaff, A., Paci, P., and Doorslaer, E. V. (1991). On the Measurement of Inequalities in Health. *Social Science and Medicine*, 33(5):545–557.

Waylen, G. (1993). Women's movements and democratization in Latin America. *Third World Quarterly*, 14(3):573 – 587.

WHO (2003). Introduction of Inactivated Poliovirus Vaccine Into Oral Poliovirus Vaccine-Using Countries. WHO Position Paper. *Weekly epidemiological record*, 78(28):241–252.

WHO (2004). BCG Vaccine. WHO Position Paper. *Weekly epidemiological record*, 79(4):25–40.

WHO (2006). Diphtheria Vaccine. WHO Position Paper. *Weekly epidemiological record*, 81(3):21–32.

WHO (2008). World Health Report. Report, World Health Organization.

WHO (2009). Measles Vaccines. WHO Position Paper. *Weekly epidemiological record*, 84(35):349–360.

WHO Multicentre Growth Reference Study Group (2006). WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development. Technical report, World Health Organization, Geneva.

Wilcox, R. (2001). Comment. *The American Statistician*, 55(4):374–375.

Wilkinson, R. G. (1992). Income Distribution and Life Expectancy. *BMJ (British Medical Journal)*, 304:165–168.

Williamson, C. R. (2009). Informal Institutions Rule: Institutional Arrangements and Economic Performance. *Public Choice*, 139:371–387.

Williamson, O. E. (2000). The New Institutional Economics: Taking Stock, Looking Ahead. *Journal of Economic Literature*, 38(3):595–613.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge.

World Bank (2001). *Engendering Development: Through Gender Equality in Rigths, Resources and Voices*. World Bank/Oxford University Press, New York, NY.

World Bank (2005). *World Development Report 2006*. *Equity and Development*. World Bank.

World Bank (2008). World development indicators. Technical report, World Bank.

World Bank (2009). GenderStats. Electronic publication. http://genderstats.worldbank.org.

Göttinger Studien zur Entwicklungsökonomik
Göttingen Studies in Development Economics

Herausgegeben von / Edited by Hermann Sautter
und / and Stephan Klasen

Die Bände 1-8 sind über die Vervuert Verlagsgesellschaft (Frankfurt/M.) zu beziehen.

Bd. / Vol. 9 Hermann Sautter / Rolf Schinke (eds.): Social Justice in a Market Economy. 2001.

Bd. / Vol. 10 Philipp Albert Theodor Kircher: Poverty Reduction Strategies. A comparative study applied to empirical research. 2002.

Bd. / Vol. 11 Matthias Blum: Weltmarktintegration, Wachstum und Innovationsverhalten in Schwellenländern. Eine theoretische Diskussion mit einer Fallstudie über „Argentinien 1990-1999". 2003.

Bd. / Vol. 12 Jan Müller-Scheeßel: Die Privatisierung und Regulierung des Wassersektors. Das Beispiel Buenos Aires / Argentinien. 2003.

Bd. / Vol. 13 Ludger J. Löning: Economic Growth, Biodiversity Conservation, and the Formation of Human Capital in a Developing Country. 2004.

Bd. / Vol. 14 Silke Woltermann: Transitions in Segmented Labor Markets. The Case of Brazil. 2004.

Bd. / Vol. 15 Jörg Stosberg: Political Risk and the Institutional Environment for Foreign Direct Investment in Latin America. An Empirical Analysis with a Case Study on Mexico. 2005.

Bd. / Vol. 16 Derk Bienen: Die politische Ökonomie von Arbeitsmarktreformen in Argentinien. 2005.

Bd. / Vol. 17 Dierk Herzer: Exportdiversifizierung und Wirtschaftswachstum. Das Fallbeispiel Chile. 2006.

Bd. / Vol. 18 Jann Lay: Poverty and Distributional Impact of Economic Policies and External Shocks. Three Case Studies from Latin America Combining Macro and Micro Approaches. 2007.

Bd. / Vol. 19 Kenneth Harttgen: Empirical Analysis of Determinants, Distribution and Dynamics of Poverty. 2007.

Bd. / Vol. 20 Stephan Klasen / Felicitas Nowak-Lehmann: Poverty, Inequality and Migration in Latin America. 2008.

Bd. / Vol. 21 Isabel Günther: Empirical Analysis of Poverty Dynamics. With Case Studies from Sub-Saharan Africa. 2007.

Bd. / Vol. 22 Peter Dung: Malaysia und Indonesien: Wirtschaftliche Entwicklungsstrategien in zwei Vielvölkerstaaten. 2008.

Bd. / Vol. 23 Thomas Otter: Poverty, Income Growth and Inequality in Paraguay During the 1990s. Spatial Aspects, Growth Determinants and Inequality Decomposition. 2008.

Bd. / Vol. 24 Mark Misselhorn: Measurement of Poverty, Undernutrition and Child Mortality. 2008.

Bd. / Vol. 25 Julian Weisbrod: Growth, Poverty and Inequality Dynamics. Four Empirical Essays at the Macro and Micro Level. 2008.

Bd. / Vol. 26 Johannes Gräb: Econometric Analysis in Poverty Research. With Case Studies from Developing Countries. 2009.

Bd. / Vol. 27 Sebastian Vollmer: A Contribution to the Empirics of Economic and Human Development. 2009.

Bd. / Vol. 28 Wokia-azi Ndangle Kumase: Aspects of Poverty and Inequality in Cameroon. 2010.

Bd. / Vol. 29   Adriana Rocío Cardozo Silva: Economic Growth and Poverty Reduction in Colombia. 2010.

Bd. / Vol. 30   Ronald Kröker: Ansätze zur Implementierung von RSE (CSR) in einem lateinamerikanischen Entwicklungsland. Das Beispiel Paraguay – Eine wirtschafts- und unternehmensethische Untersuchung. 2010.

Bd. / Vol. 31   Maria Ziegler: Institutions, Inequality and Development. 2011.

www.peterlang.de