The resource curse and social development

Exploring its relation in Latin America

Marianne Knutsdotter Andenæs

Master thesis at the Department of Political Science

Faculty of Social Sciences

UNIVERSITY OF OSLO

December 2017

Word count: 29.355

The resource curse and social development: Exploring its relation in Latin America

Marianne Knutsdotter Andenæs

December 18, 2017

Year: 2017

The resource curse and social development: Exploring its relation in Latin America

Marianne Knutsdotter Andenæs

http://www.duo.uio.no/

Print: Copy Cat Nydalen, Oslo

Abstract

For many decades, scholars have debated the relationship between riches in natural resources and development. The aim of this thesis is to complement that debate, by investigating the relationship between riches from natural resource and *social development* in Latin America, with a specific focus on how quality of governance influences this relationship. In addition, I investigate if the production of different types of natural resources can effect social development differently.

Using fixed effects models with country level data on Latin American countries from 1991 to 2012, I find no evidence of a positive, or negative, relationship between riches in natural resources and social development. The null result holds even when economic development (GDP per capita) is used as the dependent variable instead of social development (HDI). Another surprising finding is that the quality of governance does not have an effect on social development, and in turn does not condition the effect of natural resources on social development. I get the same "null findings" when I use production data on different types of natural resources. However, my results indicate that Governance has a positive effect on economic development (GDP per capita), but not on social development in Latin America. My results remain stable when running extensive models, using lagged variables and when excluding influential values from the regression. In short, my results give no support for a resource curse in Latin America.

Preface

Countless times I have doubted my own ability to finish this thesis, so producing a final product is a milestone in so many ways. Without the wonderful help and moral support from my family, friends and thesis advisor, the result would have been something quite different.

First of all, I would like to thank my thesis advisor Benedicte Bull for her priceless guidance. Her extensive knowledge about Latin America and social development in particular, has been invaluable for the writing of this thesis. The rest of the staff at Centre for Development and the Environment (SUM) also deserves great attention for their encouragement, and in particular Kristoffer Ring, for technical support and insight.

I am fortunate to have a great family and a broad circle of friends that always supports me and encourage me to reach my goals. To my parents and sister: thank you for the unconditional backing. Every phone call, Taco Friday and snapchat gave me a positive boost!

Amongst my friends at the University of Oslo, I would first like to thank Karl Bjurstrøm and Mads Motrøen, for their good company and advice in several methodological discussions. I am forever grateful! Writing a thesis has its ups and downs, and two of my closest friends have helped me turn bad days in to good days. Heidi Bang and Helene Mykland, the countless lunches, hours at the gym and your incredible company, made life as a master student much more fun!

I would also like to direct gratitude to The National Union of students in Norway (NSO), for their encouragement to "get back on track" with the thesis after ending a year in their service.

Through all the different "stages" of this project, there has been a solid rock putting up with me. I am talking about moral support, technical support, ideas, insight, you name it. My dear husband, Ådne, has always contributed. Without you, I´m quite sure this thesis would not have been completed. So, if it is even possible, I dedicate this thesis to you.

All mistakes and shortcomings are of course my own!

Marianne Knutsdotter Andenæs

Contents

List of figures

List of Tables

1. Introduction

Are riches in natural resources a "curse" for a country´s development? This relationship has been a hot topic amongst among social scientist, especially since the 1990s. The aim of this thesis is to complement that academic debate, by investigating the relationship between riches from natural resource and social development in Latin America, with a specific focus on how quality of governance influences this relationship. In addition, I investigate if the production of different types of natural resources can effect social development differently.

In a traditional perspective, natural resources are seen as a positive contribution to development, especially in western countries, like Australia, the U.S. and Canada (North 1963). The riches from natural resources provided new incomes to states and their public expenses and contributed to successful development (North 1963). Frederick van der Ploeg writes that “the positive experiences of the United States with its mineral abundance from the mid-nineteenth to the mid-twentieth century explain much of subsequent economic growth” (van der Ploeg 2011: 369). However, in more recent times, some of the best performing developing countries, like Korea and Taiwan, are resource poor whereas the “majority of the resource-rich countries, such as Argentina, Mexico, Peru, Saudi Arabia and Venezuela, have lower than average annual rates of gross domestic product (GDP) growth in the 1980-1993 period” (Mikesell 1997: 191). Extensive exploitation in light of resource booms and shifting prices in the commodity markets have led some researches to argue that natural resources can be a “curse” to a country’s economic growth (Auty 1990; Gelb et al. 1988; Sachs and Warner 1995, 1999, 2001). Richard M. Auty clearly states that in recent decades “resource-poor countries have out-performed resource-abundant ones” (Auty 2001: 839).

The studies by Jeffrey D. Sachs and Andrew M. Warner (1995, 1999 and 2001) have been noteworthy in the research debate concerning natural resources as a curse. They study the relationship between primary-product exports and GDP growth between 1970 and 1989 in 97 countries and claim to have found a negative effect (Sachs and Warner 1995: 1). In later publications they hold their position, arguing that “the curse of the natural resources is a demonstrable empirical fact, even after controlling for trends in commodity prices” (Sachs and Warner 2001: 828). Their work has been challenged by many scholars. Some raise questions about lacking measurement validity (Wright and Czelusta 2004), others claim that

Sachs and Warner (1995) monotonic understanding of the causal relationship between natural resources and economic growth ignore the role of institutions (Mehlum, Moene, and Torvik 2006). Lederman and Maloney (2007: xv) on the other hand, argue that it is how natural resources are managed by the government that is important. A somewhat similar argument is found in the works of van der Ploeg (2011), who states that several "resource rich economies may have performed badly, not because they relied too much on resources, but because they failed in developing their mineral potential through appropriate policies" (van der Ploeg 2011: 369–70). In light of these contributions, I find it relevant to investigate how governance affects the relationship between resource abundance and development. In order to get a broader understanding of the possible effect resource abundance and governance have on the well-being of the people, this thesis will emphasize on social development instead of economic development.

1.1 Why Latin American countries?

Why is it interesting to study the relationship between resource abundance and social development in Latin America specifically? First of all, the region has vast natural resource reservoirs. South American countries alone, hold some of the world's most extensive mineral reserves: 65% of global reserves of lithium; 42% for silver; 38% for copper; 33% for tin; 21% for iron; 18% for bauxite and 14% for nickel" (ECLAC/UNASUR 2013: 7). The region has excessive oil reserves, where Venezuela particularly stands out with its extra-heavy crude reserves in the Orinco Belt (ECLAC/UNASUR 2013: 7). A second aspect is that countries in Latin America are often used as an example of the "resource curse" in the literature[1]. Last, but not least, one of the distinguishing features of the region has been its social inequality. The region is commonly stated to be the most unequal in the world and characterized by its contrast (Bull 2015; Gasparini and Lustig 2011). To summarize, Latin America is an interesting region to study the relationship between natural resource abundance and social development, since the region has large natural resource reservoirs, has been central in the resource curse debate and is a region with inequality issues. In addition, Latin America is the region where the Dependency theory had its strongest foothold in policy- and academic circles in the post war period, resulting in the priority of increasing the manufacturing sector in many Latin American countries (Baer 1972).

[1] See for instance Auty (2001), Maloney (2002) or Sachs and Warner (1995, 1999 and 2001).

Latin America is defined as countries south of Rio Grande with an Iberian language (Spanish and Portuguese) as the official language (Bull 2015). This excludes the Caribbean countries, the English- and French speaking countries, like Guyana and French Guiana, as well as the previous Dutch colony, Suriname. This is not a precise definition, never the less, similarities in history and culture, makes the term "Latin America" recognizable for the majority of the population living south of Rio Grande (Bull 2015: 20).

1.2 Research questions

The question of natural resources as a gift or a curse to a country's economy has been heavily debated the last couple of decades. My thesis will be a complement to that debate, investigating how riches from natural resources affect social development in Latin American countries, with special focus on governance as an intermediate variable. My main focus will be on the following research questions:

1. *Do riches from natural resources affect social development in Latin American countries?*

2. *Does the quality of governance affect the relationship between natural resources and social development in Latin American countries?*

3. *Can production of different types of natural resources affect social development differently, in Latin American countries?*

These questions will investigated using fixed effects regressions with country level data from 17 Latin American countries in the period from 1991-2012. The details regarding the research design is presented in chapter 3.

1.3 A complement to the resource literature

This thesis presents both similarities and contrasts to previous research. The most obvious difference is my choice of dependent variable. Several contributions to the resource literature focus on natural resource's effect on economic development, whereas I use social development as my indicator. The choice of independent variables and controls, are on the other hand recognizable in the resource literature. For instance, my operationalization of "Quality of governance", bares a lot of resemblance to the institutional quality indicator found in Mehlum et al. (2006), The third research question, regarding how different types of natural resources effect social development, is an additional effort to complement existing literature, that commonly focus on all natural resources[2], or on one specific type[3]. Since I seek to investigate different aspects of the relationship between natural resources and social development, the choice of indicators is devoted considerable time due to disagreement in the literature. For instance, Stijns (2015) claims that the results presented by Sachs and Warner are not as robust as they claim themselves, precisely because of an export based measure. He claims that there are two other ways to measure natural resource abundance, namely by using production data or reserves data (Stijns 2005: 110). Others claim that resource rents or revenues are better suited to reflect the value or riches from natural resources (Johnson 1999; Ross 2001). This academic debate will be addressed to further extent in the literature review (section 2.2) and in the choice of measurement (section 3.4).

An important question to address is if “riches from natural resources” alone is a satisfactory explanation for a country’s development (or lack of development). Mehlum et al. (2006) claim that the main reason for differences in growth is not the abundance of natural resources, but the quality of institutions. Critics of the works by Sachs and Warner (1995) argue that intermediate indicators such as the policies by the government to manage natural resources (Lederman and Maloney 2007, Ploeg 2011) or issues of ideology and history (McNeish and Borchgrevink 2015) have been ignored. In other words, it is necessary to include a broad set of variables in the analysis in order to determine the actual relationships.

1.4 Structure of the thesis and important findings

In the following chapter theory related to the key concepts of the thesis is discussed. The chapter leads to deduction of three hypotheses that will be investigated in the analysis. In Chapter three, the research design and method is presented. Chapter four contains the results of the empirical analysis and discussion. In chapter five I add my concluding remarks.

In my results, I find no evidence of a positive, or negative, relationship between riches in natural resources and social development. The null result holds even when economic development (GDP per capita) is used as the dependent variable instead of social development (HDI). Another surprising finding is that the quality of governance does not have an effect on social development, and in turn does not condition the effect of natural resources on social development.

[2] See Sachs and Warner (1995, 1999 and 2001) for instance.
[3] See Mikesell (1997)

I get the same “null findings” when I use production data on different types of natural resources. However, my results indicate that Governance has a positive effect on economic development (GDP per capita), but not on social development in Latin America. My results remain stable when running extensive models, using lagged variables and when excluding influential values from the regression. In short, my results give no support for a resource curse in Latin America.

2. Theory and key concepts

In this chapter theory related to the main concept related to the subject will be presented and discussed. The objective is to define and delimit these concepts, as well as building a theoretic foundation for deduction of hypothesis, that are presented in the end of this chapter.

2.1 Social development

2.1.1 Development, from economic growth to basic human needs

In the introductory chapter we could read that in a traditional perspective, natural resources have been regarded as a positive contribution to development in Western countries (See North 1963). In this perspective, development is seen as progress in industrialization and technology, and a leading manufacturing sector would in this sense be regarded as key for successful development (Wright and Czelusta 2004: 8). A similar view of development can be found amongst scholars in the economic resource literature who argue that resource abundant countries tend to perform badly (Auty 1990; Gelb et al. 1988; Sachs and Warner 1995). ‘Perform badly’ refers to slow economic growth and shrinking of the manufacturing sector. A different perspective on both economic development, as well as development in general, is found in another branch of economics with Amartya Sen (1983). He argues that "economic growth is one aspect of the process of economic development. And it happens to be the aspect on which traditional development economics - rightly or wrongly - has concentrated" (Sen 1983: 748). In his view, growth only matters if benefits are realized through the process of economic growth. Hence, if we were only to rely on growth to achieve economic development in a broad sense, we would have to wait a long time (Sen 1983: 753–754). Sen’s own contribution and answer is that economic development should be more concerned with “what people can or cannot do” (Sen 1983: 754). This idea of development with individuals’ capabilities and basic liberties as the main focus is referred to as the ‘capability approach’, which Sen (1983) developed together with philosopher Martha Nussbaum. Overcoming problems such as poverty, elementary needs, hunger, violations of political freedom and basic liberties and neglect of interest of women, is seen as the principal exercise for development (Sen 2001: 3). Based on this understanding, development involves a variety of different dimensions.

The idea of development as multidimensional, and including the granting of basic needs such as education, health and decreased poverty has strongly influenced the analysis and studies of development (Grynspan and López-Calva 2011), and in particular shaped the work of the United Nations Development Programme (UNDP). In 1990 the UNDP published their first Human development Report, and together with the report, the Human Development Index (HDI) was launched (UNDP 2010: 4). The aim of the HDI was "international comparison of achievements on the basis that human development is the process of enlarging people's choices and their level of achieved wellbeing" (Grynspan and López-Calva 2011: 6). Each year since 1990 UNDP has published the Human Development Report and gradually updated and modified the HDI, which focuses on three basic dimensions of human development: health, education and income" (UNDP 2015: 3). An important aspect worth mentioning is that the HDI cannot be seen as a measure of poverty, but a measure of outcomes or achievements (Grynspan and López-Calva 2011).

Gryspan and López-Calva (2011: 1) argue that a multidimensional approach is the best way to study 'the many faces' of poverty. Both in academia and policy circles this is an established viewpoint, but which dimensions that are relevant and how to measure them, is still an active debate (Grynspan and López-Calva 2011: 1). One of the common methods is to focus on human basic needs. The unsatisfied basic need approach (UBN) is based on the absolute concept of poverty, where the poor are defined as households or people who cannot access basic goods and services (Grynspan and López-Calva 2011; Stewart 1985). UBN "suggest that an improvement in wellbeing is related not only to income growth, but to the satisfaction of fundamental needs like clothing, shelter, and access to education, health, and other public services" (Grynspan and López-Calva 2011: 4). This method of studying poverty is common in Latin America (Grynspan and López-Calva 2011: 4), and you can find a similar method behind ECLAC's multidimensional poverty index, which includes a broader set of indicators, such as deprivations in employment, educational achievement, and fulfilling of basic human needs (for instance access to water, sanitation facilities and electricity) (ECLAC 2014b: 11)[4]. In ECLAC's own words this measure provides "more up-to-date standards that reflect the reality in the region" (ECLAC 2014b: 11).

[4] A detailed description of ECLAC's Multidimensional poverty index can be found in Social Panorama of Latin America (2014: 18).

2.1.2 Why speak of social development

The previous section indicates that there are several ways to understand development, several more than the ones mentioned. This thesis aims to study more than the relationship between natural resource abundance and economic growth; it aims to detect the more complex human and social aspects about development. Inspired by both (Sen 1983, 2001) and the multidimensional approach on poverty, I find it more accurate to talk of *social development*, addressing both human development, including dimensions such as education, health and poverty and access to basic services, as well as social aspects like employment. The understanding of social development as a broad concept can be summarized in achieving a "basic level of economic and social well-being for all members of society" (Cecchini, Filgueira, and Robles 2014).

2.1.3 A glance at development in Latin America

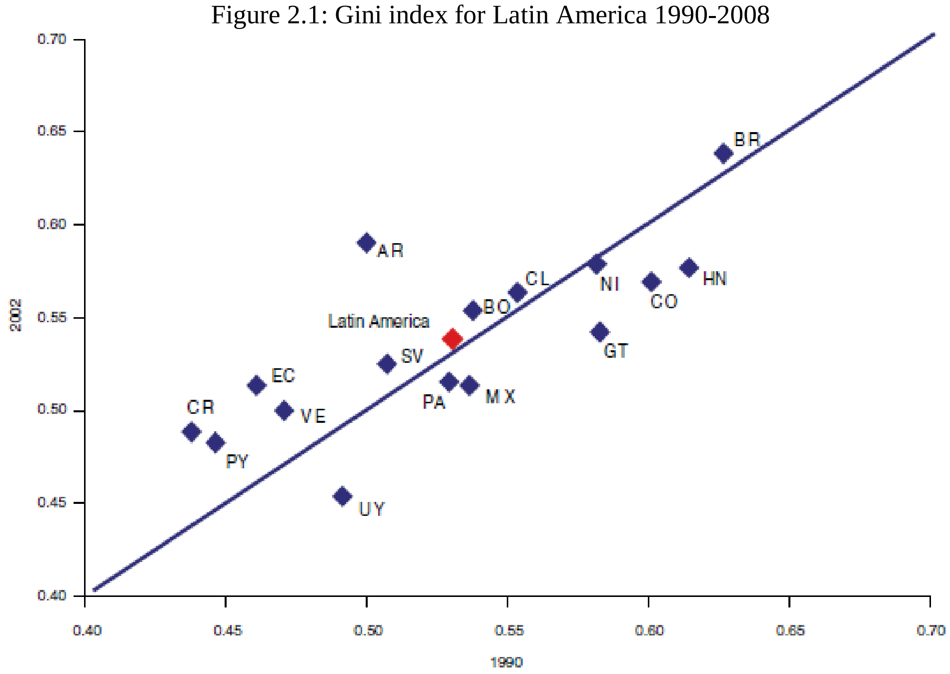
> *"Inequality is a distinctive, pervasive characteristic of the region. In fact, it is often stated that Latin America is the world's most unequal region"*
> *(Gasparini and Lustig 2011: 1)*

This section is not an attempt to describe all features of development in Latin America, but to give the reader a brief glance about the general development trends[5], starting with economic development. Latin America consists mainly of middle-income countries (Bull 2015: 58). However, Latin American "economies have historically been defined by sharp structural heterogeneity, which is largely responsible for the high levels of social inequality found in the region's countries"(ECLAC 2017: 11). After decades with slow economic growth, Latin America experienced a historical economic growth period between the years 2004-2007, partly explained by sky rocking oil, cobber and gas prices (Bull 2015: 58). At the same time as the economic growth has increased, poverty has decreased in the region (Bull 2015). Nevertheless, poverty remains a great issue for Latin American countries (Grynspan and López-Calva 2011: 1). "In 2012, as many as 164 million Latin American people were living in poverty, 66 million of them in extreme poverty" (Cecchini, Filgueira, and Robles 2014: 7). Figure 2.1 indicates a very unequal distribution of the riches in the region (Gini Coefficient 0-1), where Brazil stands out as the country with most inequality in income and Uruguay as the

[5] For more detailed information on development in Latin America, the yearly "Social Panorama of Latin America" publication by ECLAC is a good place to start.

most equal in the period between 1990 and 2002 (ECLAC 2010: 272). After 2000, by contrast, “inequality began to decrease in most of the countries of the region and this trend is still holding” (ECLAC 2014a: 55).

Figure 2.1: Gini index for Latin America 1990-2008



Source: ECLAC (2010:272)

Poverty and inequality is often linked to other aspects of development, such as education and health. Reports on educational level in the same period indicate an increase in the share that completed secondary education. “However, persistently high levels of inequality and poverty, as well as the high proportion of people living in rural areas, posed additional difficulties to the expansion of quality education throughout most of the region”, the regional UNESCO reports (UNESCO 2014). In addition, the educational system is in many of the countries characterized by divisions between private and public sectors. Elite schools often offer better quality in education and in result better prospects for the future. In short, the division in the education system can hinder social mobility, as access to education may be conditioned upon social status (Bull 2015: 147).

2.1.4 Summary

This section has presented different ways of thinking about development. Based on a multidimensional understanding I define of social development as achieving a basic level of well-being, including dimensions such as health, education, employment, access to basic services and poverty. The section has also given a brief overview of social development in Latin America, with specific focus on the inequality in the region. The details on how social development can be operationalized are presented in chapter 3.3.

2.2 Natural resources and development

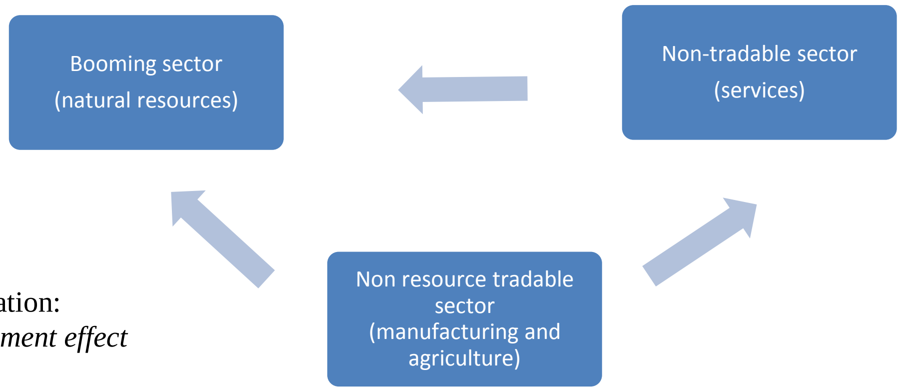
Whether or not natural resources are a curse or a gift for a country's development has been a hot topic amongst academics for decades. One of the furthermost discussed contributions to this debate is the works by Sachs and Warner (1995, 1999, 2001), and I will therefore dedicate some time and space to explain their research and the logic behind Dutch Disease models. From there we will move on to alternative ways researchers have studied the same phenomenon. A continuation of this discussion will include a presentation of natural resources in Latin America, including a part on how the Dependency theory influenced political circles and economic policies during the post-war period in Latin America

2.2.1 The Dutch Disease model and natural resources as a curse

Sachs and Warner study the effect of natural resource abundance on GDP in more than 90 developing countries between 1970 and 1989 (Sachs and Warner 1995). As we already mentioned, they argue that their results are evidence of the resource curse- a negative relationship between natural resource abundance and economic growth. In their study they use an economic approach, acquiring the principal from the *Dutch disease model*. The model was developed by Max Corden and Peter Neary to explore the effects of a "boom" in the natural resource sector on other sectors (Corden and Neary 1982: 825). The term *Dutch Disease* refers to the unfortunate effects on the Dutch manufacturing sector caused by the large natural gas field discoveries in the Netherlands in the nineteen sixties. In the Dutch Disease model, "the economy consist of three sectors: the natural resource sector, the nonresource tradables sector (usually understood as agriculture and manufacturing), and the non-tradables sector (including non-tradable services and construction) (Brahmbhatt, Canuto, and Vostroknutova 2010: 1). The first, natural resources, and the second, manufacturing, have prices set by the world market whereas the last sector, non-tradables are set by the domestic marked (Corden and Neary 1982, 826). According to the model, there are two main effects of a resource boom that could lead to Dutch disease and real exchange rate appreciation: *the spending effect* and *the resource movement effect* (Corden and Neary 1982: 827). The illustration below simplifies the framework of the Dutch disease model.

Figure 2.2 Dutch Disease model

Deindustrialization: *spending effect*



Deindustrialization:
resource movement effect

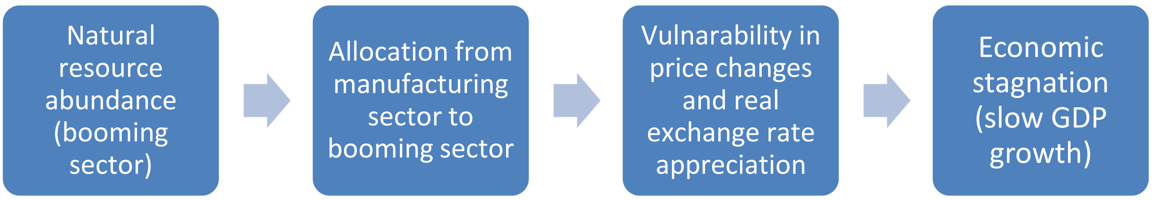
Illustration made by author based on Corden and Neary (1982).

The spending effect takes place "when increased domestic income from the booming natural resource sector leads to higher aggregate demand and spending by the public and private sector" (Brahmbhatt, Canuto, and Vostroknutova 2010: 2). The increased demand for nontradable goods (services) essentially result in higher prices on services and therefore increased spending, consequentially leading to real appreciation (Corden and Neary 1982: 827–828). *The resource movement effect*, occurs "when a boom in the natural resources attracts capital and labor from other parts of the economy" (Brahmbhatt, Canuto, and Vostroknutova 2010: 2). This gives rise to different adjustments in the economy, one important, being the adjustment mechanism of the real exchange rate (Corden and Neary 1982: 827). Both effects lead to higher prices in the nontradable goods (services) relative to the tradable, which essentially implies real exchange rate appreciation. Sachs and Warner (1995) arguments focus mainly on *the resource movement effect* of the model. They write that "the greater the natural resource endowment, the higher is the demand for non-tradeable goods, and consequently, the smaller will be the allocation of labor and capital to the manufacturing sector" (Sachs and Warner 1995: 6). In continuation, they reason that the reduced size of the manufacturing sector as a result of a boom in the natural resource sector, makes a country more vulnerable to changes of prices in the commodity market (Sachs and Warner 1995: 6). It might be an obvious statement, but still, I would like to underline that the model Sachs and Warner use in their analysis has an overall assumption of natural resources as something negative for a country's economy. The American economist Paul Krugman claims that the model presents a paradox to conventional trade models and is a continuation of a "widespread concern that the contraction of a country's manufacturing sector which follows natural resource discoveries is a bad thing" (Krugman 1987: 49). This surprises Krugman, because in the field of economics, and trade models in particular, the common idea is that "countries should simply specialize in whatever is their comparative advantage. If an oil discovery shifts this comparative advantage, so be it" (Krugman 1987: 49). As pointed out by Mehlum et al. (2006: 3), natural resources can have a positive outcome for a countries economy, like Norway and Canada have experienced. There are also examples of "success stories" within Latin American countries, like Monterrey in Mexico, that became an industrial center based on mining and Medellin (Colombia) and Sao Paulo (Brazil) based on coffee (Maloney 2002: 112).

2.2.2 How natural resources are studied can influence the results

In the analysis Sachs and Warner use resource-based export defined as agriculture, minerals and fuels as their operationalization of natural resources (Sachs and Warner 1995: 2). In later publications, they acquire the same method and hold on to their previous conclusion: "Almost without exception, the resource-abundant countries have stagnated in economic growth since the early 1970s, inspiring the term 'curse of natural resources'. Empirical studies have shown that this curse is a reasonably solid fact" (Sachs and Warner 2001: 837). They "explore different pathways for this negative relationship by studying the cross country effects of resource endowments on trade policy, bureaucratic efficiency and other determinants of growth" (Sachs and Warner 1995: 1). Their argument is still quite simple: countries with high resource-based export (as share of GDP) in most cases have a lower growth rate. With the Dutch disease model in mind, the causal relationship in Sachs and Warner (1995) can be illustrated as following:

Figure 2.3: The causal relationship in Sachs and Warner (1995)



Source: Illustration made by author using the arguments from Sachs and Warner (1995)

There are several aspects about the studies by Sachs and Warner that can be put to question. One relevant aspect is how they measured “natural resource abundance”. In their works, they do not distinguish between different types of natural resources and use resource-based exports as their indicator without explaining why they made this choice of operationalization. They have made a choice of linking natural resource abundance to resource-based export as if it was self-explanatory. As the researchers seem certain of their findings, they also dedicate a quite generalizing use of the term “curse of natural resources”. However, is a resource-based export the best way to measure the effect natural resource abundance can have on a country’s economy?

Wright and Czelusta (2004) argue that “the resource-curse literature pays little attention to the economic character of mineral resources or the concept of “resource abundance” (Wright and Czelusta 2004: 7). They raise questions of measurement validity in these studies arguing that it says too little about the actual concept itself when an equal sign is being put between resource abundance and export of mineral products. “When the revenues from this activity are described, terms such as “windfalls” and booms are generally not far behind” they argue (Wright and Czelusta 2004: 7). A similar argument is found in the works of Brunnschweiler and Bulte (2008), who state that export based measures are more closely linked to resource dependency, rather than resource wealth. Another important argument from Wright and Czelusta is that “comparative advantage in resource products is not equivalent to “resource abundance” (Wright and Czelusta 2004: 7). This argument is somewhat in the same alley as the one from Krugman, only that Krugman is critical to the fact the comparative advantage is often ignored in resource economy. Wright and Czelusta on the other hand, underline that a “comparative advantage in natural resources may simply reflect an absence of other internationally competitive sectors in the economy- in a word, underdevelopment” (Wright and Czelusta 2004: 7). I want to study, in more detail, which other factors that can contribute to give resource abundant countries an increased social development.

(Wright and Czelusta 2004)Wright and Czelusta (2004) claims that more specified measures, that take into account the size of the population, like reserves per capita or the level of natural resource export per worker, give more accurate (and different) results than the ones we find in (Sachs and Warner 1995). Wright and Czelusta focus on nonrenewable resources, minerals in particular and their results lead to the quite self-explanatory headline of their paper “The myth of the resource curse"(Wright and Czelusta 2004: 7). Even though they do not include or touch the subject of agriculture or renewable resources in particular, they indicate that their argument may apply to these sectors as well. They specifically refer to the works by Maloney (2002) and Stijns (2003) that use similar measures as themselves, and do not find a negative relationship between natural resources and economic growth. When using resource stock-based measures, Brunnschweiler and Bulte (2008) find similar results. Raymond Mikesell (1997) studies mineral exporting countries explicitly and argues that the mineral exporting countries are "the most visible victims of the resource curse", however the explanations for the resource curse covers a range of economic, structural and political factors, and it is rare that you find the similar combination of explanations in more than one country (Mikesell 1997: 191, 192). In short, this indicates that focusing on natural resources share of exports alone, might not tell us the whole story.

An additional "note" is the findings of Alexeev and Conrad (2009). They argue that the reasons for either positive or negative findings in the resource literature can be caused by misinterpretation of the data. They find that most of the major oil exploitation began before the 1950s, so even if the empirical evidence in the literature is correct, it might not consider that large oil endowments results in high growth rates in early stages of extraction, and slower rates when the oil deposits mature (Alexeev and Conrad 2009). Since most of the empirical studies are from more recent periods, one should not expect increasing growth rates, according to this argument.

2.2.3 Natural resources in Latin America

> "Throughout the history of the Latin America and Caribbean (LAC) region, natural resource wealth has been critical for its economies. Production of precious metals, sugar, rubber, grains, coffee, copper, and oil have at various periods of history made countries in Latin America—and their colonial powers—some of the most prosperous in the world. In some ways, these commodities may have changed the course of history in the world at large" (Sinnott, Nash, and Torre 2010: ix)

Countries in Latin America are often used as an example in resource literature (Sinnott, Nash, and Torre 2010: 28). Countries like Venezuela, Mexico, Argentina and Peru are the most mentioned in this context, used as "victims of the curse" (Auty 2001; Maloney 2002; Sachs and Warner 1995). Latin America is a resource rich region, with great diversity in resources. In the mining sector, "thirteen countries in Latin America are among the 15 largest producers of the world” (ECLAC/UNASUR 2013: 13). Chile, for instance, has been the world´s largest producer of cobber since 1982 (ECLAC/UNASUR 2013: 13). The cobber and the exports tied to it stand for more of the half of the country’s exports (Franko 2007: 247). The oil reserves in the region are also immense. In 2011, Latin America as a region had the second-largest oil reserves in the world (after the Middle East), counting around 20 percent of the share (ECLAC/UNASUR 2013: 35). The growth of Latin America´s share of the world’s oil reserves is mainly due to the certification in the Orinco Belt in Venezuela (The Magna project) and to successful exploration, principally in Brazil and Colombia (ECLAC/UNASUR 2013: 35). In 2015, Brazil, Mexico and Venezuela are the largest oil producers, with Venezuela considered as the ”petro-state” of the region, where oil and gas make up 90 per cent of the exports (McNeish and Borchgrevink 2015: 9).

Of other important mineral producing countries, Brazil is one of the top three producers of iron in the world, Peru one of the principal producers of silver, cobber and gold, Bolivia the fourth largest producer of tin ore and the sixth largest of silver in the world, Colombia the seventh producer of refined nickel in the world, Cuba the eight largest producer of nickel ore in the world, Mexico the largest producer of silver and Jamaica the seventh largest producer of bauxite (ECLAC/UNASUR 2013: 21). In terms of global output, the region has increased its share of gold from 10.3 to 19.2 percent, cobber ore from 24.9 to 45.4 percent and refined cobber from 15.7 to 21.9 in the period between 1990-2010 (ECLAC/UNASUR 2013: 21). These quantities are undoubtedly a proof of the importance of Latin American minerals in the world. The share each country has in global mineral exports is shown in Table 2.1.

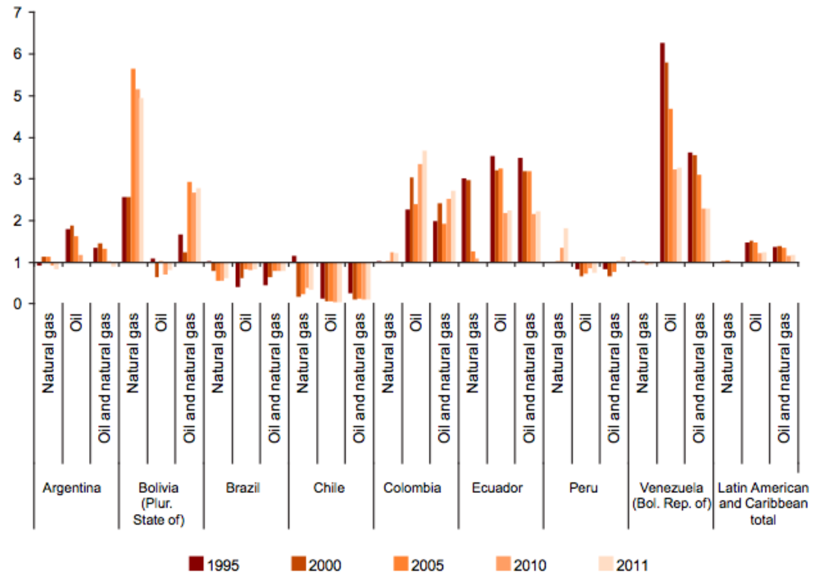
In light of the works of (Sachs and Warner 1995) and their focus on exports of natural resources, it is interesting to take a closer look on export and import of fuels in Latin America. On the average, Latin America, as a region, is a net exporter of oil and almost self-sufficient in natural gas (ECLAC/UNASUR 2013: 46). However, of the oil producing countries it was ”only Venezuela (with net oil exports representing more than 30% of its GDP), Ecuador (less than 10%), Colombia (less than 10%), Trinidad and Tobago (less than 10%) and Mexico (close to balance)” that were net oil exporting countries in Latin America in 2014 (“Uneven Recovery in Latin America: A Regional Overview” 2015). The ratio of production and consumption is also relevant. A country can have vast natural resources reserves and still be net importers of a type of natural resources. Figure 2.4 illustrates the ratio between production and consumption:

Table 2.1: Share of the total mining in Latin America and the Caribbean (1990-2012)

Mineral or metal	Percentages of world total						Production in 2011 (thousands of tons, except gold and silver, which are in tons)	Top three producers in the region in 2012 [b]
	1990	1995	2000	2005	2010	2012 [a]		
Bauxite	22.9	26.7	26.0	27.5	21.5	20.1	51 392.9	Brazil, Jamaica and Suriname
Primary aluminium	9.2	10.4	8.9	7.5	5.5	6.1	2 186.9	Brazil, Argentina and Venezuela (Bolivarian Republic of)
Copper ore	24.9	32.2	43.0	46.5	45.2	44.4	7 293.5	Chile, Peru and Mexico
Refined copper	15.7	23.2	25.1	23.7	21.5	19.6	4 134.5	Chile, Mexico and Peru
Gold	10.3	12.5	14.4	18.1	19.2	21.4	541.1	Peru, Mexico and Colombia
Silver	34.2	38.3	38.5	41.8	48.7	49.3	10 512.0	Peru, Mexico and Bolivia (Plurinational State of)
Tin ore	28.3	27.8	26.0	21.2	19.5	20.4	57.6	Peru, Bolivia (Plurinational State of) and Brazil
Refined tin	23.1	15.8	14.9	18.0	16.6	13.9	52.1	Peru, Bolivia (Plurinational State of) and Brazil
Iron	22.6	24.9	26.1	26.0	23.1	n.a.	341.2	Brazil, Venezuela (Bolivarian Republic of) and Mexico
Molybdenum ore	15.8	18.2	35.2	37.3	31.8	23.5	70.9	Chile, Peru and Mexico
Nickel ore	11.5	11.7	14.1	15.1	12.9	13.4	196.5	Brazil, Cuba and Colombia
Refined nickel	9.7	10.1	10.7	13.4	11.6	9.9	139.3	Brazil, Colombia and Cuba
Lead ore	13.3	15.5	14.7	14.6	14.5	11.0	595.1	Mexico, Peru and Bolivia (Plurinational State of)
Refined lead	7.8	7.6	8.4	7.2	7.4	4.1	484.3	Mexico, Brazil and Venezuela (Bolivarian Republic of)
Zinc ore	16.8	20.6	19.0	21.0	21.7	19.8	2 607.0	Peru, Mexico and Bolivia (Plurinational State of)
Refined zinc	7.5	8.5	7.3	7.9	7.0	7.7	969.4	Mexico, Peru and Brazil

Source: (ECLAC/UNASUR 2013: 22)

Figure 2.4 Ratio between production and consumption of oil and gas (1995, 2000, 2005, 2010 and 2011)



Source: (ECLAC/UNASUR 2013: 47)

The ratio between production and consumption of oil and natural gas (Figure 2.4) tell us that countries like Brazil, Chile and Peru are net importers. The latter, Peru, imports crude oil from Ecuador, but is increasing its consumption from local production every year (ECLAC/UNASUR 2013: 49). Brazil produces vast amounts of both oil and natural gas, so it might seem like a paradox that the country has a negative ratio between production and consumption. If we look aside of the fact that it is the country with the highest population and consumption of oil, one aspect of the explanation of the ratio could be that Brazilian refineries do not have the power to refine all the heavy oil from the local production. This implies that the country exports crude oil and imports oil products like gasoline ("The Brazilian Oil & Gas Industry" 2013: 3). To overcome the consumption deficit in natural gas, Brazil has to import, mainly from Bolivia ("The Brazilian Oil & Gas Industry" 2013: 3). The operationalization of natural resources as part of total exports, as Sachs and Warner (1995) have used, does not take into account actual resource abundance itself, but more accurately say something about self-sufficiency. But the case of Brazil for instance, illustrates that there can be other, more technology-based issues that forces the country to import gasoline. The main point is, that natural resources as part of a country's export does not necessarily give an accurate picture of the riches generated from natural resources.

According to World Bank publications, the share of natural resources in exports has declined in Latin America in general terms; however, it has not declined as much is other regions, making Latin America and the Caribbean further vulnerable to terms-of-trade shocks than it would be with a more differentiated export basket (Sinnott, Nash, and Torre 2010: 7). If we take a closer look, we do find heterogeneity among the countries in the region. The dependency on natural resources remains high in for instance Chile, Peru and Venezuela, with a commodity share of export rising over 75 percent, while is falling on the other hand, in countries like Mexico and Brazil (Sinnott, Nash, and Torre 2010: 9). Another characteristic of the commodity export of the region is a shift in the destination market, with a decline in exports to the U.S. and an increase in China´s share in the period between 1990 and 2008 (Sinnott, Nash, and Torre 2010: 9).

2.2.4 Import substitution industrialization (ISI)

Before we continue to the next section, one specific feature about economic policies in Latin America should be mentioned, namely the influence of the Dependency theory and the following import substitution industrialization (ISI) policies in the post second world war time[6]. The dependency theory can be seen as a theory trying to explain underdevelopment in poor countries, especially Latin American countries (Ardanaz, Scartascini, and Tommasi 2011) in a time when industrialized (western) countries experienced economic growth. Explained in a very simple way, Prebicsh, and the dependency theorist, argued that the uneven distribution of labor in the world, where poor countries (periphery) export raw materials to the manufacturing countries (center) and import back manufactured products to a higher cost, was a leading cause for underdevelopment in Latin America (Baer 1972: 95). Thus the name dependency theory, since the periphery countries become dependent countries, because they rely on exporting commodities to get foreign exchange (Sunkel 1969). The suggested solution for poor countries was to embark on programs of import substitution, with the aim of not needing to buy (import) manufactured products from rich countries (Baer 1972).

Some scholars argue that the inward-looking policies in the 1950s and 1960s in Latin America led the countries to invest in sectors they had no comparative advantage in, leaving them worse off than initially. Blomström and Meller (1991) have been quite harsh in their critique of ISI policies, stating that Latin American countries tried to adjust their input-output table all at the same time, some countries even trying to compete with United States production numbers. “When Latin America decided to force industrialization by import substitution, it was not an industrialization based on the countries' endowments that was supported”, they write (Blomström and Meller 1991: 9). A continuation of this critique is apparent in the works of Maloney (2002:135), who argues that these sectors were neither in “line with comparative advantage and walled off from competition and the sources of innovation, but they needed to be subsidized, or at least they diverted attention from sectors that had the potential for innovation”. Said differently, Maloney’s argument indicates that Latin American countries have missed their chance of becoming pioneers in refining and mining due to focus on “all other sectors at once”, instead of concentrating on their comparative advantage. However, it should be noted that ISI led to a decrease in resource

[6] ISI policied derivies from the dependency theory, developed at the United Nations Economic Commission for Latin America (ECLAC), with Raul Prebisch as director at that time.

dependency and an increase in regional trade, the latter still being an important priority in the region (Bull 2015: 52). The fact that Latin American countries, to a further extent than other regions, introduced import substitution makes it an even more interesting region in light of my research questions.

2.2.5 Summary

This section has presented some of the major contributions to the resource literature, as well as given insight to natural resource production in Latin America. The results presented by different scholars, indicate that the measurement one chooses can give divergent results. The promotion of alternatives measures to the export-based indicators is especially relevant in light of my own considerations in choosing an appropriate measure for "resource riches" to be included in my analysis (section 3.4)

2.3 Governance

Is the relationship between natural resource abundance and economic growth as direct as some scholars argue? According to Mehlum, Moene and Torvik (2006) the role of institutions has been ignored as an intermediate factor in this relationship[7]. Other scholars argue that other important aspect influence this relationship, as the policies by the government to manage natural resources (Lederman and Maloney 2007; van der Ploeg 2011) or issues of ideology and history (McNeish and Borchgrevink 2015). These contributions indicate that the negative relationship might not be as direct as assumed by Sachs and Warner (1995). Mikesell (1997: 198) reasons that too much emphasis has been placed on resource abundance countries lacking ability to promote a productive manufacturing sector. As we have seen earlier in this thesis, for economist like Sachs and Warner (1995), a growing manufacturing sector is the key to technological success. Mikesell (1997: 199) on the other hand, argues that the relationship between natural resource abundance and economic growth is more complicated than that. He suggest that the reason for lacking economic growth in resource rich countries can be explained by "incorrect government policies or exogenous conditions that have little to do with mineral exports and could just as well have occurred in resource-poor countries" (Mikesell 1997: 198). I aim to explore how governance conditions the relationship between natural resource riches and social development, and the following section therefore

[77] The role of institutions is also relevant in the studies of Wright and Czelusta, and of course, (Acemoglu, Robinson, and Johnson 2001)

emphasizes on how we can understand governance as a concept and its relation to development.

2.3.1 Governance as a concept

"The concept of governance is notoriously slippery; it is frequently used among both social scientist and practitioners without a definition which all agree on" (Pierre and Peters 2000: 7). Said in other words, governance is a complex concept with a variety of ways to be defined. According to the American political scientist R.A.W. Rhodes there are at least six different ways to use the term governance: as the minimal state, as corporate governance, as the new public management, as 'good governance', as socio-cybernetic systems and as self-organizing networks (Rhodes 1996). I will not address all of these in detail but try to connect other scholars' understanding of governance to the categorization of Rhodes.

Political scientist, Mark Bevir, who has written several books on governance in cooperation with Rhodes, argues that "governance can be used to refer to all patterns of rule" (Bevir 2008: 3). Still, this has not been the case in more recent understanding of governance. Bevir seems concerned when he writes that in recent times "the interest in governance derives from its specific use in relation to changes in the state in the late twentieth century" (Bevir 2008: 4). With this he refers to the "liberal ideal" of a small state with little market interference. According to Leftwich (1993) neoliberalism is more than just an economic theory. It addresses theories of politics and the state and introduced terms like 'bad government' and necessary government to distinguish the type of rule they favored (Leftwich 1993: 608). The neoliberal understanding of government falls logically into Rhodes first category "governance as the minimal state", where the idea of governance relates to reducing the size of the government, increase privatization and in consequence, a shrinkage in public services (Bevir 2008: 5).

A more common understanding among social scientist today is that governance is a "complex and fragmented pattern of rule composed of multiplying networks" (Bevir 2008: 7). The literature on governance therefore highlights different aspects, like "the role of markets, networks, and none state actors" (Bevir 2008: 14), hence three of Rhodes (1996) uses of governance; corporate, socie-cybernentic systems and self-organizing networks. Rhodes himself argues that "governance is broader than government" (Rhodes 1996: 660). On the other "side", we find the contrast and reaction to this pluralistic understanding of governance. The most obvious is Francis Fukuyama. He considers it a paradox that few people pay attention to the most important institution that "accumulates and uses power, the state" (Fukuyama 2013: 347).

Fukuyama define governance as "a government's ability to make and enforce rules, and to deliver services, regardless of whether that government is democratic or not" (Fukuyama 2013: 350). His understanding of governance is closely connected to government power and reasons that "governance is about execution" (Fukuyama 2013: 351). The reason for his main focus on government is due to several issues that governments have in common, like education, health and public safety, that can be used in an instrumental measure of the quality of governance (Fukuyama 2013: 351). The notion of Fukuyama is strait forward; good governance relies on a well-functioning state that executes these roles in a proper manner. A slightly similar argument one finds in Pierre and Peters (2000). They view the state as the leading political actor and provider of basic services, hence they believe that the state should be included in analytical models of governance (Pierre and Peters 2000: 25). This is also supported by Bull (2016) who states that an analysis of governance should start by defining the state "not an ideal type, as that of Weber, but one that is able to capture real existing states, in all their variations and weaknesses" (Bull 2016: 4).

Another feature that is common whilst studying governance, and especially good governance, is democracy and impartiality. Rothstein and Teorell (2008: 166), who are both connected to the Quality of Governance Institute, argue that democracy is a necessary criterion to include whilst studying the quality of governance, because it "concerns the access to government power". Their understanding of governance is two-sided. First, one should have political equality and impartial treatment on the "input side of the system". The notion is inspired by (Dahl 1991) understanding of political equality, where democracy can be understood as the way to ensure political equality. Second, this has to be "complemented with impartiality on the output side of the political system, that is, the exercise of public authority" (Rothstein and Teorell 2008: 170). This includes impartiality in the implementation of laws and policies and how government officials handle cases (reported in a corruption scale). The notion of how the government is selected and how the government exercises its power is also an important feature in the World Bank Institute's conceptualization of governance as "the traditions and institutions by which authority in a country is exercised" (Kaufmann, Kraay, and Mastruzzi 2011: 222). The Worldwide Governance indicators developed by the World Bank Institute include the selection of government, the capacity of government to implement policies and the state and citizen's respect "for the institutions that govern economic and social interactions among them". (Fukuyama 2013: 348)Fukuyama (2013: 348) is, not surprisingly, critical to this blending of democracy with governance. Leftwich (1993: 613), also being critical, states that many of the success stories of economic growth, such as Brazil, South Korea and Taiwan in the third world did in fact not happen in stable democracies. He argues that "effective adjustment in practice required both a strong and relatively autonomous state, whether democratic or not" (Leftwich 1993, 607). Putting the democratic aspect aside, I would still argue that enforcing the rules and policies by the government, as described by Rothstein and Teorell(2008), represents some of the core elements of governance.

2.3.2 The role of the state and its capacity

Fukuyama's state centric definition if governance is quite similar to the notion of state capacity that one can find in contributions in development studies. (Johnson 1982) finds that in order to have successful development, you need a strong state. His studies on "the Japanese Miracle" have been noteworthy in the debate of the role of the state in industrialization. Johnson highlights that political stability (authoritarian regime) and predictability for the system can strengthen the state and its ability to use markets forces to promote development. He does not promote an authoritarian state, but a developmental state (Johnson 1999). Strong states are in other words able to use its authority in order to promote development. Said more differently, "The governance capacity of states depends on their autonomy and authority (…) Autonomy depends on the control of resources" (Bull 2016: 4). Control of resources implies not only the state's control natural resources, but control over capital and mean, as tax collector for instance (Tilly 1990). In political economy, they don't use the same term state capacity or quality of governance to great extent, but when they refer to the quality of institutions, it is often comparable or referring to the same (Acemoglu, Robinson, and Johnson 2001; Mehlum, Moene, and Torvik 2006).

In a revision of several ways to measure state capacity, Hanson and Sigman (2013) promotes a focus on three main dimensions of state capacity: extractive capacity (tax collection), monopoly of the use of force and administrative capacity, based on the Weberian notion of a good bureaucracy (Hanson and Sigman 2013: 3). A similar view in found in Bull (2015). When talking about how to measure state capacity, she argues that three criterions are often included: the state's ability to protect its citizens, the state's ability to collect taxes and the state's ability to provide services (Bull 2015: 128). The Quality of Governance institute provides data on several indicators that can reflect these dimensions, which will be discussed to a further extent in section 3.3, where I present my choice of measure.

2.3.3 Summary

This section has presented an overview of different understandings of governance. Based on the existing theory, I find it difficult to speak of governance without including the role of the state. Following Fukuyama's definition, the key aspects of governance is the state's role in enforcing rules and providing services for its population. Therefore I will focus on the states capacity to fulfill these roles when I discuss alternative ways to measure the quality of governance.

2.4 Hypotheses

Based on the conceptual discussions in the previous sections, several hypotheses have been deduced.

The first hypothesis seeks to investigate the direct relationship between riches from natural resources and social development. Can Sachs and Warner's (1995, 1999, 2001) strong claim of a resource curse weakened when looking at social development instead of economic growth?

H_1: Riches from natural resource have a positive effect on social development

The argument of governance quality (a strong state) as key to development has been raised by scholars like Leftwich (1993) and Johnson (1982). In the second hypothesis, I seek to combine elements from the development research with a "hot topic" amongst resource economist (Mehlum, Moene, and Torvik 2006; Wright and Czelusta 2004). Therefore, I explore if a positive relationship between natural resources and social development is conditioned by high quality of governance.

H_2: Riches from natural resource have a positive effect on social development, when the score on governance is high

A third aspect that this thesis explores is whether different types of natural resources can lead to diverse outcomes on social development. As discussed in previous sections, it has been common in the resource literature to either focus on one category of natural resources (see Mikesell 1997) or regarding all natural resources as "the same" (Sachs and Warner 1995). This is the rationale behind the third hypothesis, were I assume the following:

H3: Production of different types of natural resources effects social development differently

3. Research design

This chapter addresses questions related to methodology and justification of the choices related to operationalization of variables that will be part of the analysis. The last part of the chapter deals with assessment of missing data, presentation of the data set and the statistical model. Due to missing data, different imputation methods are discussed, concluding on using multiple imputation.

3.1 The quantitative approach – using panel data

Qualitative and quantitative analysis can "appear quite differently; indeed they sometimes seem to be at war" (King, Keohane, and Verba 1994: 3). King, Keohane, and Verba (1994: 4, 7) argues that all good scientific research, both quantitative and qualitative, derive from the same logic of inference, based on empirical information. What separates the two traditions is specific techniques and styles (King, Keohane, and Verba 1994: 3). Research in social sciences attempts to explain complex social situations. In our case, we are trying to describe the causal relationship between natural resource abundance, quality of governance and social development. As we have seen in earlier chapters, the relationship between natural resource abundance and (economic) development has been studied in different ways and with different outcomes. Most of the resource literature derives from the field of economy, where models have been tested on many units and over a period time, using a quantitative approach[8] and concluding with contrary results (see chapter 2.2.). The divergent outcomes indicate that it is somewhat unclear if the causal relationship between natural resource abundance and development is a positive or negative one. Hopefully, this thesis will a), contribute to this uncertainty and b), shed light on a broader understanding of development and governance, as part of this relationship.

A causal relation might better be explained when we study a phenomenon over time (Stock and Watson 2007: 349). Causal hypotheses deals with changes, for instance that one occurrence is predicted to generate other occurrences. Data that describes changes over a period of time can bring us closer in understanding and explaining the causal mechanism (Skog 2010: 324). More specifically, when studying changes in the dependent variable over time, we can be able to eliminate the effect of overlooked (omitted) variables that differ

[8] One of the exceptions is McNeish and Borchgrevink (2015), who base their conclusions on qualitative case studies.

across countries but are constant over time(Stock and Watson 2015: 398). Time-series-cross-section data (TSCS), also referred to as panel data, "are characterized by repeated observations on fixed units such as states or nations" (Beck and Katz 1996: 2). The numbers of observations in panel data are the number of country-year observations (Stock and Watson 2015: 397). There are upsides and downsides with using panel data[9], but the advantage of studying the relationship over time, controlling for country specific effects and reducing the risk of omitted variable bias compensate for the potential downsides.

3.2 From concepts to variables in analysis

The following section contains a justification and explanation of the operationalization of variables that are included in the analysis. King, Keohane and Verba (1994: 9) emphasize the importance in research of following the "rules of inference which its validity depends". One important step is to assure a proper operationalization of the phenomenon we study. This essentially implies that the measurement we adopt reflects the concept we study in a meaningful way (Adcock and Collier 2001). As previously seen, the choice of indicator has been a debated in the resource literature, concerning both the measurement for natural resources, development and other variables that might influence this relationship. In relation to the latter, I include relevant control variables to minimize the risk of omitted variable bias that can influence the analysis (King, Keohane, and Verba 1994)[10].

3.3 Social development as the dependent variable

Social development was in section 2.1 defined as achieving a basic level of well-being, based on a multidimensional understanding including dimensions such as health, education, employment, access to basic services and poverty. The concept was defined based on previous theory and research, and in specific, inspired by the works of Sen (1983), the Human Development Programme and ECLACs multidimensional poverty measure. Based on the literature, but also availability and quality of data, my conclusion is that the existing measurement that best reflects my definition of social development is *the Human Developments Index (HDI*).

---

[9] For instance, panel data is considered to be "dependent", since observations from one year, often depend on previous year's values. Autocorrelated residuals are therefore often an issue. More details on the potential risks are reviewed in section 4.2.

[10] The correct model specification is important in regard to omitted variables. This will be discussed further in chapter 4.

While theories consist of concepts, the data we use for statistical analysis consist of indicators (Christophersen 2013: 179). Social development is a multidimensional concept which indicates that one indicator alone, cannot account for all the aspects that social development involves. I considered the creation of an index, combining “the best” from the HDI and ECLAC’s multidimensional poverty index. The following table illustrates a possible way to construct a social development index. “HDI” and “ECLAC” in front of the name of indicator signal that the indicator is used in either the HDI or in ECLAC’s multidimensional poverty measure.

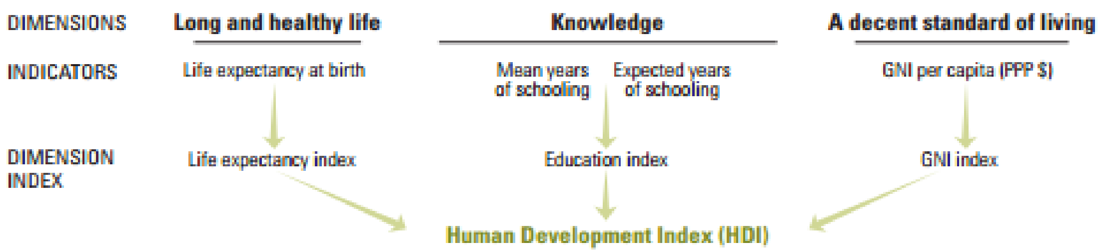
Table 3.1: From systemized concept to indicators

Systematized concept	***Indicators***
Social development as achieving a basic level of well- being, including indicators such as poverty, health, education, employment and access to basic services	1. HDI: Gross national income per capita (GNI PPP), *WB 2. HDI: Life expectancy at birth, total years, *WB 3. HDI: Average years of schooling, *QoG 4. ECLAC: Unemployment/employment, *WB 5. ECLAC: Access to electricity, water and sanitation,* WB

*indicates the source of data

By investigating these data[11] more closely, the creation of a new index, proved to be less suitable than expected. Taking into consideration problems such as missing data, intercorrelation, skewness and different types (and range) of values, I decided that the HDI-index would be a better choice of measure. It was especially the missing educational data that led to this decision. The HDI-index has the upside of complete coverage (no missing data) for all the country year observations and is, as we already now, a measure developed with direct reference to the works of Amartya Sen. In short; the HDI-index covers the majority of the dimensions I included in my conceptualization[12].

Figure 3.1: The human development Index



Source: Human Development Programme[13]

[11] The table also shows the source of the data, where ”WB” stands for ”The World Bank” and ”QoG” stands for Quality of Governance Institute
[12] Employment and access to sanitation/water are the dimensions that the HDI does not cover.
[13] See http://hdr.undp.org/en/content/human-development-index-hdi

The HDI-index focuses on three basic dimensions of human development: living long and healthy life, access to knowledge and a decent standard of living (UNDP 2015: 3). The indicators that are used to measure these dimensions are life expectancy at birth, mean years of schooling, expected years of schooling and Gross National Income (GNI) per Capita. As figure 3.1 illustrates, these indicators form three different indices, and by using mean of each one, added together, the HDI-index is created[14]. The index has a score from 0-1, where higher score indicates better achievement. The HDI as not a perfect match to my definition of social development, but the essential idea that development is concerned with the well-being of people is well preserved by using the indicator. “The basic objective of development is to create an enabling environment for people to enjoy long, healthy and creative lives” is the rationale behind the HDI-index (UNDP 1990: 9). Nevertheless, as a precautionary measure, I still include the dimensions from the ECLAC’s multidimensional poverty measure not covered by the HDI in evaluations of model specifications and stability of the results.

3.4 Independent variables

3.4.1 Natural Resources

Chapter 2.2 discussed different attributions to the resource literature, using Sachs and Warner (1995) as a starting point for discussing whether or not natural resources are a curse for a country’s development. A main objective for this thesis is to investigate the relationship between riches from natural resources and social development. To evaluate hypothesis 1 and 2, I argue that *Total natural resources rents (% of GDP)* is a suiting indicator. I also seek to distinguish between types of natural resources to answer the third hypothesis. Therefore, I use a second operationalization of natural resources, based on production data for different types of natural resources. In the following I will explain why I choose to use these two measures.

Sachs and Warner (1997, 200) have in many ways “set out the course” in how the relationship between natural resources and economic development are investigated, at least amongst economist. They have used "the share of primary exports", either as % of GDP (Sachs and Warner: 1997) or gross domestic product (GNP) as % of GDP (Sachs and Warner: 2001). When Mehlum et al. (2006) challenges Sachs and Warner's findings, they also use "the share of primary exports in GNP" as their indicator for resource abundance. However, scholars like

[14] For further details on the construction of the HDI-index, see the technical note accompanying the Human Development report. 2016: http://dev-hdr.pantheonsite.io/sites/default/files/hdr2016_technical_notes_0.pdf)

Wright and Czelusta (2004), Maloney (2002) and Stijns (2003) have raised concern about the use of export measures alone, and therefore promote the use of reserves data and exports per worker. In more recent studies it is more common to focus on government revenues, since the availability of data in this "field" has improved[15], and some argue that the previously used export measures are a better indication of resource dependency, than resource wealth (Brunnschweiler and Bulte 2008)[16]. Especially studies that investigate the relationship between natural resource wealth and political indicators, such as democracy, institutions, corruption etc., suggest the use of measures that better reflect actual resource wealth, not its share of exports (Aslaksen 2010; Haber and Menaldo 2011; Ross 2001). Ultimately, that implies using resource rents or revenues as percentage of GDP as indicator for resource wealth. In order to answer my research question, the use of resource rents or revenues will lead to an operationalization that reflect "riches from natural resources" in a better way, than for instance the use of export- or reserve measurements. Based on data availability (time period and countries), using natural resource rents as % of GDP came down as the best option[17], using the indicator provided in the QoG Standard dataset, based on World Bank Development indicators (Teorell et al. 2016: 740). Resource rents are calculated based on the difference between the production of the resource and the price, and is therefore a quite precise measurement on riches from natural resources[18]. The indicator is the "sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents, and forest rents" as percentage of GDP (Teorell, Jan et al. 2016: 740).

In the resource literature, it has been common to focus on fuels (petroleum and natural gas), minerals, with the exception of Sachs and Warner (1995), who include all type of resources. Still, the most debated natural resources have been the non-renewable fuels and minerals. In order to complement previous research, I ask if the production of different types of natural resources could lead to different outcomes on the dependent variable. In light of previous research and objectives of this thesis, I will focus on fuels and minerals (not agriculture). The different arguments raised in chapter 2.2, and in the previous paragraphs, indicate that the use of resource rents (for each type)[19], production data or reserve data are relevant options. In

[15] Ross (2001, 2004) and Haber and Menaldo (2011) are identified as important contributors to the improvement of data. For more info, see Prichard et al. (2014: 14).
[16] Brunnschweiler and Bulte (2008) suggest the use of resource stock-based measures
[17] Resource revenues based on Haber and Menaldo (2011) was evaluated, but their time limit ends in 2006
[18] Details on these calculations can be found in World Bank (2001).
[19] The use of separate rents measures would be an option, but I did not encounter an indicator for Gas rents (in the QoG Standard dataset), only mineral rents, oil rents and coal rents

order to investigate how the production of different types of natural resources affect social development, it seems more logical to focus on production data since, data on reserves do not give us precise insight of the actual activity in the sector, it reflects the potential size. Hence, production data can give us more information than data on reserves. Another important aspect is that reliable data on natural resource reserves in Latin American countries are not easy to find[20].

In light of the mentioned arguments, I natural resource production is measured using production data on minerals and fuels. CEPALSTAT[21] provides production data on important minerals[22], as well as natural gas and petroleum. In order to detect the differences from one type of resource production to another, the creation of dummy-variables or a categorical variable, is an options. These could inforce, or at least clarify the differences in outcome on the dependent variable between the two main categories of natural resources. However, by creating dummies or a categorical variable, we lose important information. For instance, we would not be able to say anything about differences in the scale of production and its relationship with the dependent variable. I want to extract as much information from the data retrieved from CEPALSTAT as possible. Therefore, natural resources are operationalized as shown in the following table:

Table 3.2: Operationalization of natural resource production

Natural gas production	*Natural gas production in petajoules*
Crude Petroleum production	*Crude petroleum production in thousands of metric tons*
Gold and Silver production	*Gold and silver production in tons.*
Mineral production	*All other minerals in thousands of tons*[*]

*copper, coal, zinc, tin, iron, wire rod, pig iron, aluminum

In the analysis, these variables are log transformed to prevent extreme values to lead the results (see chapter 4).

[20] If we were only to study petroleum and gas, we could have used reserve data from for instance U.S. Energy Information Administration (eia). If we were to follow in the footsteps of Sachs and Warner (1995), we could use World Bank data on Fuel, Ores and Metals export as percentage of total exports.
[21] CEPALSTAT data is provided in Excel-format. Reshaping the data from “long” to “wide” is therefore necessary in order to use the data for time series analysis in Stata. This also applies for World Bank data.
[22] CEPALSTAT provides data on copper, coal, zinc, tin, iron, wire rod, pig iron, aluminum, gold and silver

3.4.2 Quality of governance

As we have seen in previous sections, several arguments have been raised in the resource literature concerning how government policies[23] and institutional quality[24] influence the effect natural resource riches have on development. This is the rationale behind the second hypothesis of this thesis, assuming that a positive effect of natural resource riches on social development is conditioned by high governance quality. In order to investigate this relationship I use "The International Country Risk Guide's (ICRG) Indicator of Quality of Governance" as my measurement for quality of governance. As the following section will demonstrate, this indicator is an adequate measure in light of my defining of governance, and in addition, has shared features with the institutional indicators used in Sachs and Warner (1995) and Mehlum et al. (2006)

Since Governance is such a complex concept, it is important that the operationalization we choose is valid and reliable. In chapter 2.3, the quality of governance was, defined as the state's ability to enforce rules and provide services for its population This definition is based on Fukuyama (2013), Johnson (1982) and Leftwich's (1993) state-centric view on governance, implying that the main actor and enforcer of power is the state. Since the theorized concept is based on the idea that governance itself matters[25], not the way the rulers are selected, democracy is not regarded as a dimension of governance quality. Democracy is more relevant in the analysis as a control variable[26]. The dimensions concerned with the government's authority and ability to provide services are more in line with the theorized concept I defined.

Based on my definition, governance quality can be operationalized focusing on three leading dimensions or capacities[27]:

1. The government's authority over resources
2. The government's authority of coercion
3. The state's ability to provide services

---

[23]See Lederman and Maloney (2007), Mikesell (1997) or van der Ploeg (2011)
[24]See Mehlum et al. (2006) or Wright and Czelusta(2004)
[25] See Fukuyama (2013), Johnson (1982) or Leftwich (1993)
[26] Democracy a relevant control variable in light of studies by for instance Ross (2001), Haber and Menaldo (2010,2011) and Aslaksen (2010), but also due to arguments of the relationship between democracy and development (see Bevir 2008 and Leftwich 1993), and democracy as "good governance" (Rothstein and Teorell 2008).
[27] See Hanson and Sigman (2013)

The first dimension can be understood as the government's authority over taxing, spending and legislating. The second, as the government's monopoly on the use of force (Hanson and Sigman 2013). These two dimensions are a bit slippery, since both can be linked to general "law and order" in the literature. The third dimension relates to the quality of the bureaucracy.

Finding good data on these dimensions (for at least two decades), is not necessarily easy since it relies on countries to report on political and institutional indicators. Said in other words, no data at all, or missing data for countries or certain time periods are frequent problems. The Quality of Governance Institute (QoG) Standard Dataset can be a good source when studying political/societal country level indicators, since they provide time series on country level indicators. I aimed at finding indicators that reflected the three dimensions of the concept and had available data on Latin American countries in the same period as the HDI-index (1990-). I especially considered the variable "Government Authority over taxing, spending or legislating"[28] as well as the variable "Monopoly on the use of force"[29]. At a first glance, they are both well-founded in light of my definition of quality of governance; however the last dimension regarding quality of bureaucracy is not a part of either of the indicators (or indices). The other, and also severe issue, is that "Monopoly on the use of force" has no data before 2006 and "Government Authority" has no data at all for seven of the countries to be studied[30]. Using either of those, or both, would therefore not be very efficient.

Another option is to use the same or similar indicators as the ones we find in the resource literature, like the institutional quality indicators used by for instance Mehlum et al. (2006) and Sachs and Warner (2001). In both studies, The ICRG indicator "Rule of Law" is used, either separated[31] or as part of the institutional quality index[32]. Three, out of the five, indicators used by Mehlum et al. (2006) forms the (ICRG) Indicator of Quality of Governance[33] found in the QoG Stardard Dataset. This index is constructed using mean values of the ICRG variables "Corruption", "Law and Order" and "Bureaucracy Quality", scaling from 0-1, where higher values indicate higher governance quality (Torell et al. 2016:

[28] called dpi_author in the QoG Stardard Dataset codebook
[29] called bti_muf in the QoG Stardard Dataset codebook
[30] Government authority has no data on Bolivia, Honduras, Panama, Paraguay, Peru, Uruguay and Nicaragua.
[31] Sachs and Warner (2011) uses this as one of many a control variables
[32] Mehlum et. al uses an institutional quality index (Knack and Keefer 1995) as independent variable and as part of an interaction term
[33] ICRG collects political information and converts them into risk point. For more info: https://www.prsgroup.com/about-us/our-two-methodologies/icrg

348). The ICRG indicator has complete coverage from 1991-2012 for all countries[34], providing a much better practical starting point than the other options. Using this indicator implies that I have to adjust the former operationalization, and clarify what I am actually measuring.

The quality of governance will be operationalized using the ICRG Indicator of governance including the following indicators:

1) The degree of corruption in the political system
2) The strength and impartiality of the legal system and the popular observance of the law
3) Bureaucracy quality and institutional strength

Measuring governance using the ICRG indicator implies that the dimension “government authority”, used in the initial operationalization, is left out of our measure[35].

3.5 Control variables:

To make sure that we do not exclude relevant variables from our analysis we include control variables. This is often referred to as minimizing the risk of omitted variable bias. More specifically this bias occurs when or if variables that have an effect on the dependent variable, and simultaneously are correlated with one or more of the independent variables, are not included in the analysis (King, Keohane, and Verba 1994: 169). Not taking into account the risk of omitted variable bias can lead us to under- or overestimate the results and of course, raise questions of spurious effects. This does not imply that we should include “all we can think of” in our analysis. King et al. (1994) actually advise against it, mainly because it can make the estimate of the independent variable less efficient and give us problems with multicollinearity (this will be discussed further in the next chapter). “The implication should be clear: by including an irrelevant variable, we are putting more demands on our finite data set, resulting in less information available for each inference” (King, Keohane, and Verba 1994: 183). This is why the control variables should be selected carefully, with a clear reference to the literature on the causal relationship we study. I have chosen control variables based on variables that, in the reviewed literature, have been argued to have an effect on development.

[34] The exception is Nicaragua
[35] “Authority” is kept in the dataset and used in the imputation model.

Level of economic development

Economic development has been used as the dependent variable in many of the mentioned contributions to the resource literature, where GDP growth or GDP per capita are commonly used measures. In order for us to carry out standard comparisons across time, it preferable to use a GDP measure with constant dollars (Strand, Hegre, Gates, and Dahl 2012). The World Bank provides "GDP in constant US dollars (2011)", based on purchasing power parity. In the analysis, this variable is log transformed. In testing the stability of the results, GDP will be used as dependent variable.

Democracy

Democracy has been prompted by western countries and organizations as the key to successful development (see section 2.3). There are also those who claim that democracy is a necessary dimension to include whilst studying the quality of governance (Rothstein and Teorell 2008). One can also fin contributions to the resource literature that study the relationship between resource wealth and democracy[36]. I therefore regard democracy as a theoretically relevant variable. Democracy itself is a complex concept and one can find a variety of understandings of the concept in the field of political science. I will not pursue a new theoretical discussion. The reason is quite simple, the previously mentioned Rothstein and Teorell (2008) argue that democracy is important when you study governance and therefore I choose an indicator that the two scholars themselves have provided to the QoG Standard dataset. Democracy is therefore operationalized with a dichotomous measure, where a country is defined as democratic if it satisfies the conditions for both contestation and participation (Teorell, Dahlberg, Holmberg, Rothstein, and Svensson 2016: 109). This measure was developed by Charles Boix, Michael K. Miller and Sebastian Rosato, first published in "A complete dataset of political regimes 1800-2007" (Boix, Miller, and Rosato 2012).

Population

The size of a population is an important control variable. Especially since the countries in Latin America vary enormously in size of population. Another aspect connected to the resource literature, is the claim that measures of natural resources should take into account the size of the population (Wright and Czelusta 2004). Our operationalization of natural resources does not do this, which strengthens the rationale of including of population as a control

[36] See Ross (2001), Haber and Menaldo (2010,2011) and Aslaksen (2010)

variable. The data is collected from The World Bank and based on the de facto definition of population, which includes all residents (2016). Population is log transformed in the analysis in order to prevent extreme values from leading the results.

Time

After decades with slow economic growth, Latin America experienced a historical economic growth period between the years 2004-2007, partly explained by sky rocking prices on oil, cobber and gas (Bull 2015: 58). The resource dependency has decreased in the region in the period of study, and in recent years, in many countries the state has seized a greater role administrating natural resources (Bull 2015; Sinnott, Nash, and Torre 2010). Positive increase in several of the indicators in the analysis, and the mentioned arguments, makes it relevant to include the variable “Time”, to capture this linear time trend. A dummy variable “splitting” the time period in 1991-2001 and 2002-2012 will be used when testing the stability of the results.

Additional (potential) controls

There are several other variables that could be relevant to include as controls, such as natural resource prices or indicators that were discussed in relation to the operationalization of the independent variables. In the model using production data, it might be relevant to use one of the World Bank indicators with world prices on natural gas, oil or minerals in constant 2010 dollars (The World Bank 2016). Since I decided to use the HDI-index as the dependent variable, indicators such as unemployment rates and access to water sources and electricity are not part of operationalized social development. Employment and access to basic services are dimensions that are often included in poverty measures and plays a central part in studies of poverty in Latin America (see chapter 2.1). Measuring governance using the ICRG indicator implies that the dimension “authority”, used in the initial operationalization, is left out of our measure. King et al. (1994) advice against overfitting the model using “all potential variables”. I risk potential omitted variable bias, multicollinearity and loss of efficiency by adding these variables as controls. As will be discusses in the missing data section, these variables can still serve a purpose in the imputation model. As a precautionary measure, I test the use of these variables as controls in light of omitted variable bias, as well as using them in extensive models when evaluation the stability of my results.

3.6 Methodological challenges

As mentioned earlier, there are certain obstacles that we come across using panel data. Some of them are important to give account for before we start the analysis. Hence, missing data will be addressed in the following section. Other features that are typical for time series analysis, like autocorrelation and multicollinearity will be discussed in chapter four in conjunction with regression assumptions.

3.6.1 Missing data

> *"Problems of missing data are pervasive in empirical social science research (…) A relatively few absent observations on a handful of variables can quickly reduce the effective N"* (Allison 2001: v).

We want our results to be accurate and at the same time, use as much as possible of the available information. When using panel data, missing data can pose considerable challenges. (Little and Rubin 2002: 3). As can be seen in table 3.4, the number of observation for each variable vary, telling us that some variables have moderate amounts of missing observations, while others have complete coverage. The question is if, and how, as much of this information can be preserved? There is not "a one and only cure" to address missing data, however one can encounter medicine that will reduce the severity of problems and strengthen the analysis. Like humans, datasets and the analytical models differ, and therefore the medicine for missing data should be chosen carefully. I will start this section by describing the different types of missing data and thereafter I will justify my choice of medicine: multiple imputation.

Based on the mechanisms that leads to missing data, one can distinguish between the three types of missing data: 1) data missing completely at random (MCAR), 2) data missing at random (MAR), and 3) data not missing at random (NMAR) (Little and Rubin 2002: 11–12). "Random" can be a quite confusing term in this setting since it is being used in a statistical context. "Random" in this sense indicates that a process is probabilistic rather than deterministic (Schafer and Graham 2002: 152). When "missingness does not depend in the values of the data" it is referred to as MCAR (Little and Rubin 2002: 12). This occurs when there is no specific cause that has led to the missing values and the missing data is not related to either the variable itself or other variables in the model (Christophersen 2013: 81). If data is missing completely at random, we will not find any pattern or logic in the missing values, since the missingness is not caused by other variables in our data. MAR on the other hand, occurs when missing values on one variable depends on other variables in the analysis (Little and Rubin 2002: 12). In other words, when data is MAR it can be explained by other observed values in the data. At last, if missing values are neither MCAR or MAR, they are missing due to variables that are not included in the data (NMAR). This type of missing data can be the most difficult to handle (Christophersen 2013: 82). We have to start by identifying which mechanism of missing data we are dealing with, which according to (Høyland and Nygaard 2011: 2–3) is very difficult to be entirely confident about. So, in sum, one could argue that there is no single right answer (or medicine) on how to handle missingness.

Table 3.4: Summary statistics before imputation

Variable	Mean	Std. Dev.	Min.	Max.	N
HDI-index	.6749064	.0702851	.484	.831	374
Governance	.4624961	.1224177	.1666667	.7777778	352
Natural res. rents	7.225785	8.12085	.2616182	44.10928	374
Aluminum	1419.681	5413.75	0	34078	329
Copper	3531.344	12781.52	0	89024	346
Coal	356.3006	1106.783	0	5557	321
Petroleum	26900.63	49667.43	0	178280	319
Gold	9.106383	16.90475	0	103	235
Iron	20721.78	79905.22	0	526750	302
Natural gas	292.0325	522.074	0	1950	338
Pig iron	485.2962	1531.725	0	10110	290
Silver	293.3152	834.0287	0	5358	276
Tin	2.836299	8.292564	0	42.145	355
Wire rod	435.2163	1255.369	0	7673.8	311
Zink	115.0528	295.2819	0	1602.6	301
Gold_Silver	376.0688	938.1714	0	5461	218
Minerals	15880.81	75302.45	0	538637.4	184
Democracy	.8794118	.3261281	0	1	340
GDP constant	9465.189	4031.023	3219.128	21141.98	352
World prices oil	45.68806	27.75922	15.8991	97.5972	374
World prices min.	63.19036	24.95056	37.23676	113.3027	374
World prices gas	6.839091	3.324477	3.674252	15.38487	374
Population	374	2.89e+07	4.50e+07	191127	374
Year	2001.5	6.352787	1991	2012	374

A common technique to handle missing data is by deletion, either listwise or pairwise (Peugh and Enders 2004). *Listwise deletion* has been the common way to handle missing data in statistical analysis. This method simply excludes all missing values and only keep the units with complete coverage of data, with the aim of securing consistent results. Another common method, *pairwise deletion*, excludes units with missing data only when the variable is used in calculations. These two methods do not take into account the scale of missing values on specific variables, the number of units with missing values nor the amount of missing values on variables with different measurement levels (Christophersen 2013: 82). An important assumption that has to be met in order to use these methods of deletion is that the data is MCAR, hence missing data on the dependent variable is unrelated to other independent variables and the dependent variable itself (Peugh and Enders 2004). This assumption is likely not be met when dealing with socio-economic variables, were missing data often is consequential (Honaker and King 2010: 562). Using listwise deletion can in these cases induce serious bias on parameter estimates and lead to narrow confidence intervals (Høyland and Nygaard 2011: 2). With a decreasing number of observations, the efficiency of the model will also be reduced using listwise deletion. This becomes quite clear in our case by running a simple model with the production data using listwise deletion, leaving us with only 132 complete observations[37]. In light of the previous mentioned aspects of social-economic variables, it is not very surprising that listwise deletion reduces the efficiency of our model[38]. In short, it seems that liswise deletion is not the best medicine to handle missingness.

Another way to handle missingness is to use an imputation method to replace missing values, with the aim of enforcing the effects that already exist in the model. A single imputation model, for instance a linear imputation model, would use information from known observations (prior or subsequent) when replacing values, assuming a linear trend in the data (Pedersen 2013: 59). Linear imputation would not discard missing values (as listwise deletion does), but treat imputed values as known values in the analysis (not taking into account the imputation uncertainty). By doing so, the variance of estimates can be underestimated leading to overstating the precision of confidence intervals, which in turn can make the interpretation of significance test too optimistic (Allison 2001; StataCorp 2013: 3)[39]. To overcome the shortcomings of linear imputation, especially when dealing with panel data with larger gaps between observed data, many researcher, especially in more recent times, choose multiple

[37] The results can be sees in appendix 1.
[38] The model includes socio-economic variables, such as the Human Development Index, natural resource data and a quality of governance indicator. For more information on socio-economic factors and the use of listwise deletion, see Høyland and Nygard (2011:4)
[39] For a more detailed comparison of imputation methods see Allison (2011) and Little and Rubin (2002).

imputation as their method to handle missing data (Høyland and Nygaard 2011: 2). As the term indicates, the multiple imputation method creates multiple sets of plausible values, and is a "simulation-based statistical technique for handling missing data" (StataCorp 2013: 3).

Multiple imputation is a time-consuming method, but in light of the potential loss of efficiency[40] (listwise delition) and possible misleading results (linear imputation), multiple imputation seems to be the best medicine in our case. Before I go into detail about the conducting of multiple imputation, it should be noted that the main goal of multiple imputation is not to produce values as close as possible to the true ones, but, according to the creator of the method D.B. Rubin, the purpose is "to handle missing data in a way resulting in valid statistical inference"(StataCorp 2013)[41]. One has to assume that the data is MAR when using multiple imputation, since the imputed values are based on the observed data (Honaker, King, and Blackwell 2011: 50–51; Pedersen 2013: 60). As a response to the MAR-assumption, one should include variables in the imputation model that may contain information about the missing data (variables that might in fact be reason for missingness, such as political or economic variables for instance) (Høyland and Nygaard 2011: 8; StataCorp 2013: 118) (Stata 118). By doing so, the MAR assumption will be more plausible and in turn, the imputation quality will increase[42]. Following that suggestion, 14 variables were included in the imputation model. The appendix 1 provides an outline of this model.

The pattern of missing data can either be monotone or arbitrary, and the choice of imputation model should take into account the pattern. In our case, it is arbitrary[43]. I tested different imputation models before deciding on using the predictive mean matching (PMM) imputation method[44]. PMM "imputes missing values by means of the nearest-neighbor donor with distance based on the expected values of the missing variables conditional on the observed covariates" (Vink, Frank, Pannekoek, and Buuren 2014: 62)[45]. The model can be of preference when the distribution of a variable is skew[46], which is specially the case for natural

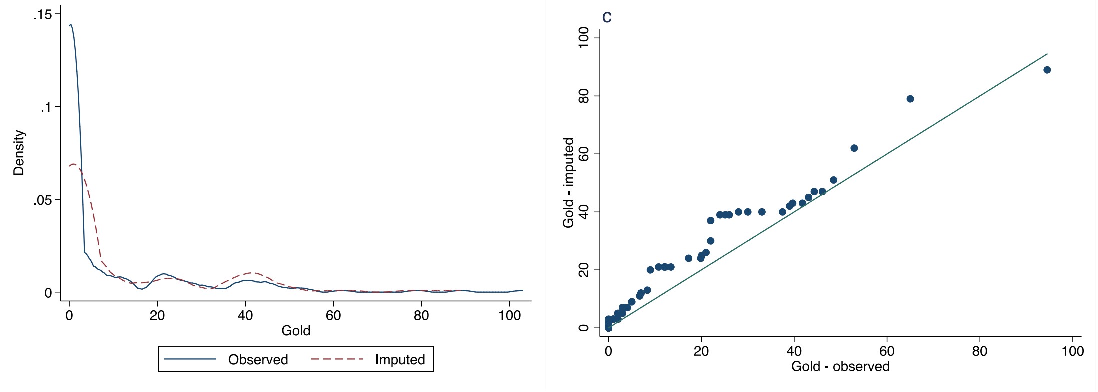
[40] It is especially the model using production data that will experience a loss of efficiency.
[41] For more information, see Rubin (1987)
[42] See (Meng 1994) for more information on the accordance between imputation and complete-data models.
[43] The results of the command "mi misstable nested" indicated that there were 10 different "statements" describing the missing data pattern. The arbitrary pattern was also confirmed when testing an imputation model for monotone missing data.
[44] Several imputation models were tested. Both multivariate models and the linear model produced "impossible" values for several variables.
[45] For a more detailed description on PMM see Rubin (1987: 158) or Vink et al. (2014)
[46] The skewness of the variables at hand might be the reason for why other imputation models failed to produce likely values. For more details se (StataCorp 2013: 245–246)

resource production data. The model also allows us to restrict imputed values within the range of observed values (Statacorp 245-46), an advantage in our case compared to other models that imputed negative values on positive variables. When specifying the numbers of neighbors in the model, I followed the arguments of (Morris, White, and Royston 2014) and (Allison 2015), using 5 neighbors. Lastly, I choose to create 20 imputed datasets for each variable[47]. Since the imputation model has to both backcast and forecast values, imputing such a large number of imputed datasets is done as a precautionary measure.

The next question is, how well did the medicine work? In order to determine how well the model worked, I investigated several types of imputation diagnostics and estimations[48]. Graphical comparison of the imputed and observed values is a commonly recommended diagnostic (Nguyen, Carlin, and Lee 2017; White, Royston, and Wood 2011). In figure 3.2 contain graphical comparisons of imputed and observed values using Kernel density plots and quantile-quantile plots for the variable "Gold" (that initially had the largest amount of missing values)[49]. As we can see in the density plot, the distribution of both observed data and the imputed data are both left skewed. An alternative way to evaluate the distribution is by a quantile-quantile plot that illustrates if the imputed values follow the same distribution pattern as the observed one. Both comparisons indicate a quite successful imputation model, bearing in mind that the aim of multiple imputation is of analytical purpose. A closer look indicates that some unusual values might have been produced by the imputation model. This will be reviewed in section 4.2, in the general investigation of influential values. Additional diagnostics are presented in appendix 1 and summary statistics are displayed in section 3.7. Even though there is uncertainty attached to the estimates the imputed data, the side effects, such as bias with other solutions are greater. I will therefore proceed to the analysis using imputed data.

[47] PMM is not a multivariate imputation model, and 14 different imputation models were carried out.
[48] Stata is very sensible concerning the number of observations in each estimation model and does not except estimation models with different numbers of observations. The estimation process and following diagnostics was therefore conducted in separate datasets for each imputation model (variable).
[49] See appendix xx for graphics on all variables

Figure 3.2: Comparison of imputed and observed values for “Gold”.



3.7 The dataset

The data used in this thesis consist of comparable time series observed on 17 Latin American countries from the year 1991 to 2012. The panel is balanced, meaning that the panel has the same number of observations per year and there are no “gaps”. An overview of the countries included in the panel is presented in the appendix. The time period is restricted to this period based on a theoretical argument and a practical one. The resource curse debate was revived in the middle of the 1990s, using data dating up to around 1990. Of course, more recent publications include more recent data. It is still interesting to study the period from 1990 until the most recent available data for other reasons. Many of the countries have experienced transition to democratic ruling in this period, as well as improvement in growth rates, policy implementation and reduced poverty (Grynspan and López-Calva 2011: 2). Consequently, studying this period might give us new insight to the relationship between natural resources, quality of governance and social development Still; data availability does to some degree determines the time limit. The cut of point at 2012, and not later, is due to available data on the Quality of governance indicator[50]. Relatively short time series are ordinary in social sciences, and trends in the data can therefore be a more serious issue, compared to datasets with longer time span (Skog 2010: 327).

The dataset is constructed with data retrieved from four main sources: World Bank Data, CEPALSTAT, United Nations Development Programme[51] and The QoG Standard dataset. The latter already exists in panel data format for Stata, which is time saving, whilst the other three had to be reshaped (and sorted) before merging. In total the dataset consist of 374

[50] The data from the QoG Standard dataset 2016 was downloaded by the writer in spring 2016. The 2017 version of the dataset does not include any additional years.
[51] The data for the HDI-index is dowlaoded directly from UNDP

country-year observations (N). Summary statistics of the variables in the model after imputation is presented in the following table.

Table 3.5: Summary statistics after imputation

Variable	Mean	Std. Dev.	Min.	Max.	N
HDI-index	.6749064	.0702851	.484	.831	374
Governance	.4624961	.1224177	.1666667	.7777778	352
Natural res. rents	7.225785	8.12085	.2616182	44.10928	374
Aluminum	1471.217	5252.819	0	34078	374
Copper	323.114	1045.801	0	5557	364
Coal	3602.081	13001.4	0	89024	370
Petroleum	27918.63	51325.73	0	178280	355
Gold	8.740741	16.20206	0	103	351
Iron	20453.95	76036.28	0	526750	354
Natural gas	298.2378	518.9686	0	1950	370
Pig iron	915.0296	2190.984	0	10110	361
Silver	271.1918	780.1526	0	5358	365
Tin	3.246118	8.943788	0	42.145	374
Wire rod	415.9034	1182.654	0	7673.8	366
Zink	107.9259	274.5561	0	1602.6	366
Gold_Silver*	289.3219	804.2306	0	5461	351
Minerals*	27881.89	86183.55	0	576020	348
Democracy	.8783784	.3272909	0	1	370
GDP constant	9465.189	4031.023	3219.128	21141.98	352
World prices oil	45.68806	27.75922	15.8991	97.5972	374
World prices min.	63.19036	24.95056	37.23676	113.3027	374
World prices gas	6.839091	3.324477	3.674252	15.38487	374
Population	374	2.89e+07	4.50e+07	191127	374
Year	2001.5	6.352787	1991	2012	374

*New indices based on imputed values were created

3.8 The statistical model

Two main models are developed in order to investigate the hypotheses deducted in chapter two. I call these models "The resource rents model" and "The production model". The latter includes natural resource production data on different types of natural resources and the former includes "Natural resource rents as percentage of GDP". Panel data should be perceived as dependent, since repeated observations of one variable are usually more similar for the same unit then for different units (Christophersen 2013: 159). As with other methodological questions, designing the analytical model has to address the data at hand and the research question. Different models have been developed that take into account the nature of panel data, like the fixed effects and random effects models (Stock and Watson 2015: 396) Other models such as OLS using panel corrected standard errors (OLS PCSE), also referred to as pooled OLS, has also increased in popularity, since this model takes account of autocorrelation and heteroscedasticity in the residuals. Since we are dealing with different countries, using a shared intercept as the PCSE model does, might not be the best fit. In our case, a model that is based on unit specific intercepts, such as the fixed effects model might be the best solution (Stock and Watson 2015: 411). Running the Hausman test and comparing of models (random versus fixed) supports the use of fixed effects[52].

The models are developed in order to investigate the research question, and consequently the hypothesis. The resource rents model is linked to the first hypothesis (*Riches from natural resources have a positive effect on social development)* The interaction term included in the following model has the aim of testing the second hypothesis (*Riches from natural resources have a positive effect on social development, when the score on governance is high*) and the third model is developed in order to study if production of different types of natural resources affect social development differently (hypothesis 3). The third model will be referred to as the "production model", since production data on different types of natural resources are included.

In the written models "αi" is the unknown intercept for each entity (n entity-specific intercepts), which characterized the fixed effects model. $B_{i,t}$ is the coefficient for each independent variable, where i = country and t = year (the same goes i,t in the dependent variable). The $e_{i,t}$ is the error term (residual).

[52] The null hypothesis "the coefficients for all years are jointly equal to zero" is rejected in our case, and time effects are needed (Stock and Watson 2015: 407–408). See appendix 2 for results of the Hausman test and comparing of FE and RE..

Resource rents model:

$$HDI\ (i.t) = \beta 1 Rents\ (i,t)\ \beta 2 Gov\ (i,t) + + \beta\ 3 Dem(i,t) + \beta 5 GDP\ (i,t) + \beta 6 Time(i,t) + ai + e(i,t)$$

Resource rents model with interaction term:

$$HDI\ (i.t) = \beta 1 Gov\ (i,t) + \beta 2 Rents\ (i,t) + \beta 3 GovRents\ (i,t) + \beta 4 Dem(i,t) + \beta 5 GDP\ (i,t) + \beta 6 Time(i,t) + ai + e(i,t)$$

Production model:

$$HDI\ (i,t) = \beta 1 Gov\ (i,t) + \beta 2 Petro\ (i,t) + \beta 3 Gas\ (i,t) + \beta 3\ Gold\_\ Silver\ (i,t) + \beta 4 Minerals\ (i,t) + \beta 5 GDP\ (i,t) + \beta 6 Time(\beta) + ai + e(i,t)$$

4. Empirical analysis

In this chapter I present the results of the empirical analysis. I will explore if the different hypotheses deducted in chapter 2 can be discarded or not, based on the research design outlined in the previous chapter. This also involves making necessary adjustment to the model, in order for OLS regression to be optimal. In the last part of this chapter I perform a series of robustness test, aiming to ensure that the conclusions I draw are not driven by any arbitrary specifications.

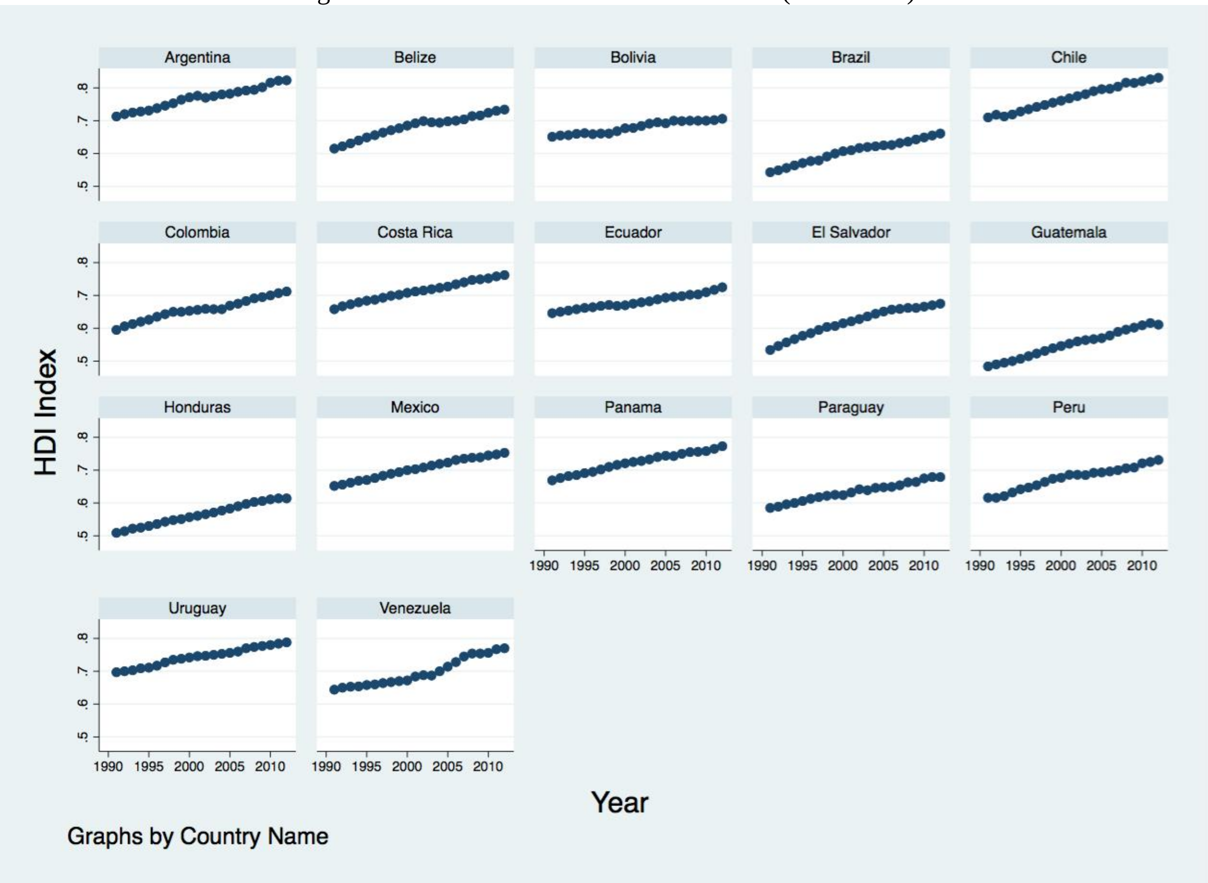
4.1 Descriptive statistics

It is useful to inspect some basic characteristics of the variables at hand before moving on to the regression analysis. Table 4.1 presents summary statistics, including average scores. In the literature review, the contrast of Latin America was mentioned. When looking at Figure 4.1, that graphically displays the HDI scores by country, this becomes evident. Guatemala and Paraguay has the lowest scores, while Chile and Argentina has the highest scores. The similarity however, is the positive trend.

Table 4.1: Descriptive statistics

Variable	Average score	Std. Dev.	Min	Max	Number of obs.
HDI Index	0.675	0.0702	0.484	.831	374
Natural resource rents (% GDP)	7.226	8.120	0.262	44.109	374
Petroleum (log)	5.392	4.919	0	12.091	349
Gas (log)	2.708	2.961	0	7.576	361
Gold and silver (log)	2.121	2.580	0	8.193	348
Minerals (log)	5.124	4.495	0	13.264	347
Governance	0.462	0.122	0.167	.778	352
Democracy	0.878	0.327	0	1	370
GDP (log)	9.062	0.440	8.077	9.959	352
Resource rents*Governance	3.084	3.574	0.116	18.645	352
Population (log)	16.238	1.485	12.160	19.125	374
World prices Crude Oil (log)	3.638	0.605	2.766	4.580	374
Time	10.5	6.352	0	21	374

Figure 4.1 HDI for Latin American countries (1991-2012)



4.2 Pre estimation assessments

Before proceeding to the analysis I investigate if any unusual observation can influence the regression and test if the assumptions for the regression model are met. Since I use panel data, I should be especially aware of heteroscedasticity and autocorrelation in the error term (residuals), and consider if multicollinearity can pose potential problems. If we were to ignore this, our results could be misleading and we risk drawing the wrong conclusions. In many of the cases, we could adjust the models or transform variables to take account for these problems[53].

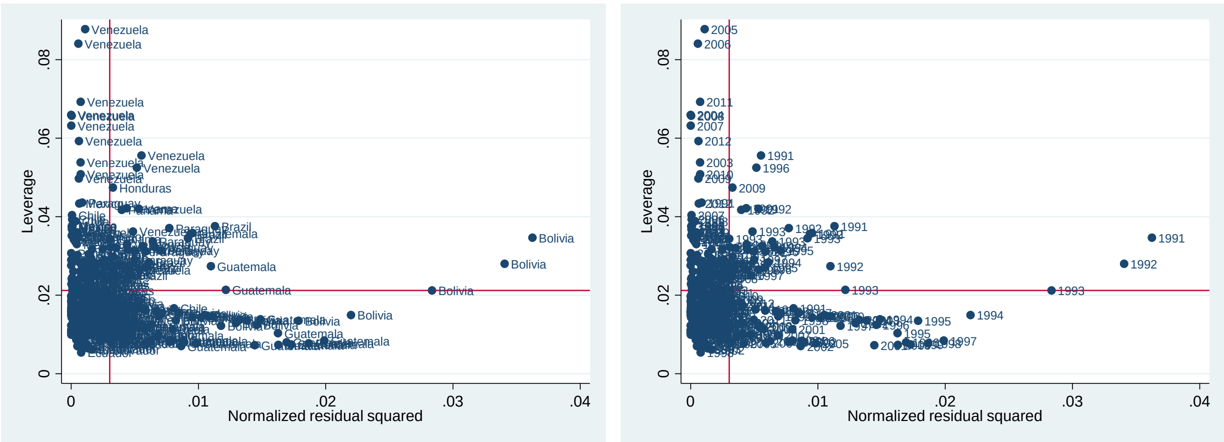
4.2.1 Unusual values

Unusual values are values that the model is not able to "digest", and in smaller samples, like the one I use, unusual values can be of critical or influential importance to regression coefficients (Stock and Watson 2012: 167). A model where regression coefficients depend (or

[53] For the two main models "resource rents model" and "production data models" results of all tests/diagnostics are presented either in the appendix or in between texts. The resource rents model with the interaction term is very similar to the first model, so it is only multicollinearity were the result for that model separately is included.

rely) on a single or few values, is not a very trustworthy one. I have therefore investigated potential outliers and high leverage points, and the influence unusual values have on coefficient estimates. The leverage versus residual plot can tell us if we have observations with high leverage points, detect outliers in addition indicate which observations that might influence our results (by jointly having high leverage points and being an outlier). For the first model, the natural resource rents model (4.2), Venezuela stands out with high leverage points. Influential values would present themselves in the upper right corner of this plot. Bolivia is the only country that has values that might fall in this category. Cooks D versus residual plots indicate that natural resource rents have some outlier values for Bolivia[54].

Figure 4.2: Leverage versus residual squared plots for natural resource rents model



In order to investigate high leverage points and outliers in more detail and determine if those values influence the analysis, I create a DF Beta plotS, where influential values will be seen above or below the line (figure 4.3). DF Beta is created for each independent variable and the calculations asses how each coefficient is changes by deletion of an observation (Chen, Ender, Mitchell, and Wells 2003). The red line indicates the conservative cut of point (+/- 0.110) [55] and observations outside this line should be given more attention. However, less conservative rules set the cut of point at 1 (Chen, Ender, Mitchell, and Wells 2003; Christophersen 2013: 119) and at .5 (Menard 2010: 136–139) Additional plots with observation by year are presented in appendix 2. In figure 4.3 we can see that, Democracy and Governance values for Bolivia in 1991 are outside the conservative cut off point, the same goes for Governance values in Guatemala in 1991 and Paraguay in 1998. By investigating the

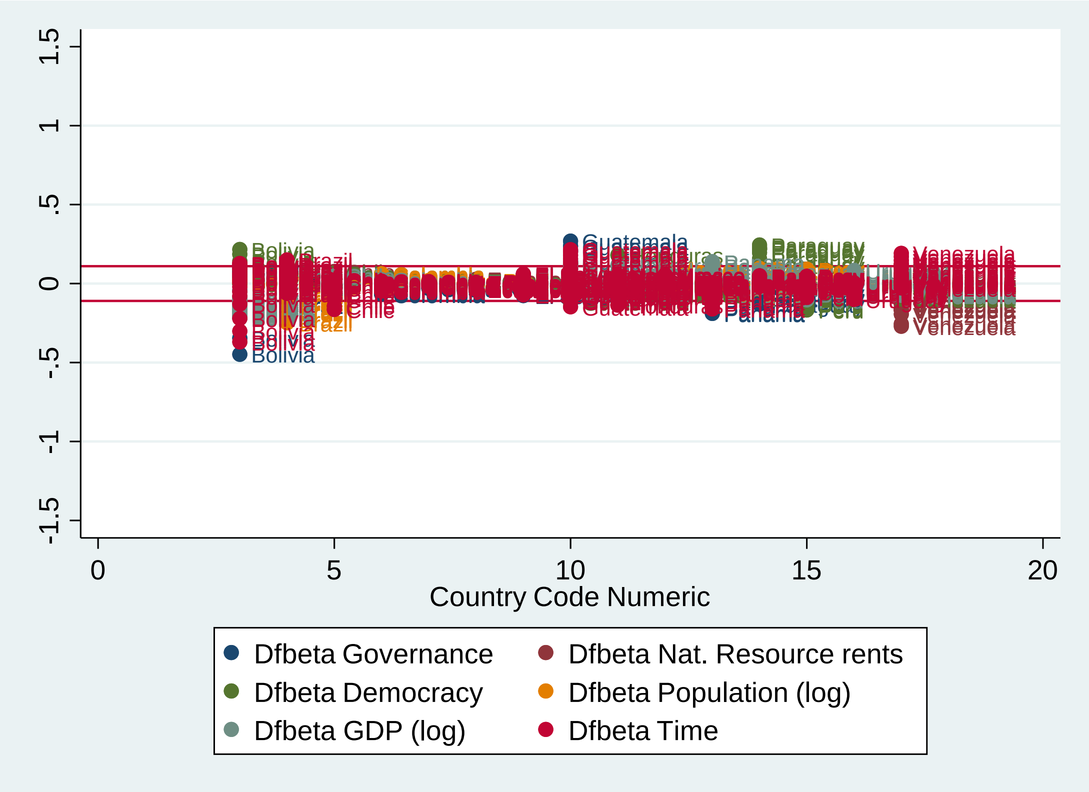
---

[54] I checked if this was reduced by log transformation, but it does not make a lot of difference, since natural resource rents is in percentage points (0-100). See appendix 2.

[55] The cut off point for DF Beta is 2/sqrt(N), where N is the number of observations.

values closer, we can see that Stata is to some extent warning us that values are changing for a country-year observation[56]. Deleting those observations is not a good idea since the observations represent a natural part of the sample. Transforming a dichotomous variable and an index makes little sense (extreme values are not the case and log transforming a variable were cero I a frequent value (democracy) would produce missing values), so one preferable option is to estimate the model with and without those observations, and comparing the results (Christophersen 2013: 119). This is carried out in section 4.4.

Figure 4.3: DF Beta Plot of influential values in the natural resource rents model

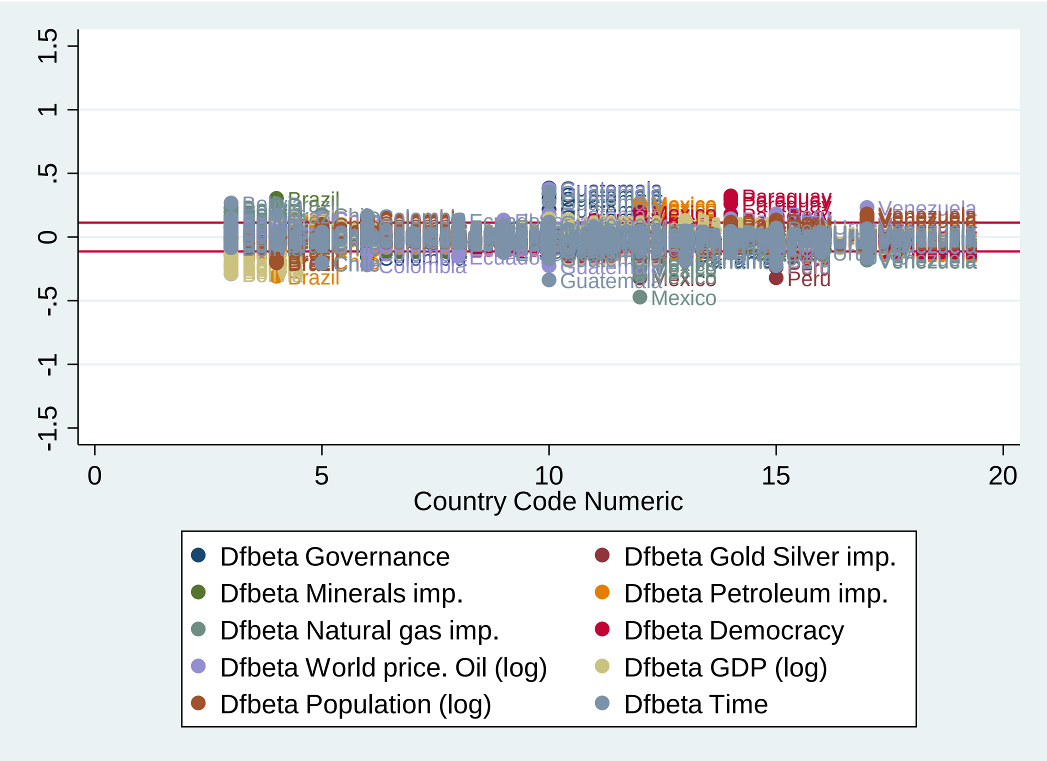


For the second model, the production data model, these test and plots helped me detect some extreme values caused by the imputation method. The leverage versus residual plot (appendix 2) indicate that Brazil have high leverage points, Guatemala has outliers (as in the resource rents model), and Mexico, Brazil, Bolivia and Peru might have some influential values. The DFbeta plots presented in appendix 2 confirm this, and in addition points out Venezuela's influential values. Comparing imputed values with observed values confirm that the imputation method has led to some extreme values for Petroleum, Natural Gas, Minerals and Gold_Silver in the beginning of the time series. The DF Beta plot in figure 4.4 illustrate that

[56] By investigating the unusual values in specific, we see that Bolivia and Guatemala have increasing values on the ICRG quality of governance indicator in 1991-1992 and Paraguay experience a decrease from 1998-1999.

removal of the extreme values caused by the imputation model reduces the number of influential values and leaves us with no DF Beta values over 0.5. As in the resource rents model, Bolivia has values with high Cooks D (see appendix 2, figure 6.6). In section 4.4, I will compare regressions with and without influential units (countries).

Figure 4.4: DF Beta Plot (B) of influential values in the production model[57]



4.2.2 Assumptions for OLS to be optimal

If we ignore testing the assumptions for the model at hand, we risk getting misleading results and drawing wrong conclusions about the relationship we study. Researches using panel data often focus on heteroscedasticity and autocorrelations in the residuals, and use a model that corrects the standard errors for the model, to take account of this violation of assumption. The other assumptions for OLS to be optimal (BLUE[58]) is still important when using fixed effects models, logically since this model is an OLS model with country specific dummies implemented. Testing the assumptions is not always as “straight forward” as it might seem. A test of the violation of one assumption can be influences violations of other assumptions that in turn can make it difficult to distinguish one violation from another. Second, the possibility

[57] Additional DF Beta plots comparing before and after removal of influential imputed values are displayed in appendix 2.
[58] Best Linear Unbiased Estimator, for more info, see Stock and Watson (2015: 115)

to uncover a violation is bigger when the sample size increases, even though the consequences are more severe in small samples (Cameron and Trivedi 2010: 93; Midtbø 2012: 106). Regression diagnostics, such as residual plots, are therefore a supplement, and not a replacement, to the theory driven analysis. In the next sections, I test the following assumptions:

1. The relationship between the independent variables and the dependent should be linear
2. Residuals should be normally distributed, homoscedastic and dependent
3. There is no perfect multicollinearity
4. There is no omitted variable bias

The relationship between the independent variables and the dependent should be linear

Regression analysis assumes a linear relationship and is therefore important to investigate is the relationship between independent variables and the dependent is actually linear. Linearity is the central mathematical assumption in the model (Gelman and Hill 2006: 46). An effect can never be perfectly linear and often sufficiently linear is what one could hope for (Agresti and Finlay 2008: 283). If the functional form is nonlinear, and we do nothing about it, our regression coefficients can tell us a misleading story (Pennings, Keman, and Kleinnijenhuis 2017: 26). In consequence, this can lead to serious bias estimates, not only bias standard errors. I have investigated scatter plots of each independent variables versus the dependent one (results can be seen in appendix 3).

As expected, none of the relationships are perfectly linear (the time variable is very close to being perfect). In the resource rents model, the variable “Natural resource rents” and “the Governance indicator” deviate a bit from the regression line, but not enough to speak of a curvilinear relationship. To adjust this, an option could be to transform the variables, but these variables are, as mentioned earlier, not very suitable for transformation, and I therefore continue without transforming these. The variables “Population” and “GDP per Capita” are already log transformed[59]. In the case of population, the relationship between population and HDI deviates considerably more than others (which will be discussed further in later

[59] One of the reasons for transforming population and GDP was early testing of influential values (due to a big range in these variables, and therefore high leverage points and outliers). In addition, it is common to log transform these types of variables. It is on the other hand not common to log transform variables that already are expressed in percentage, such as Natural resource rents or variables with small range, as Governance (Midtbø 2012: 135)

sections). In the production model, there is some deviance from the linear regression line, especially for natural production variables. Log transforming[60] Petroleum, Natural Gas, Gold_Silver and Minerals improves this (see appendix 3).

The residuals should be normally distributed

The normality assumptions, as well as the other residual assumptions, relate to the testing of significance, not the estimates. If the normality assumption is not met, hypothesis testing might not be valid, but our estimates will still be unbiased (Chen, Ender, Mitchell, and Wells 2003). Of the residual assumptions, this is regarded as the least important one, and some argue that this assumption is superfluous (Gelman and Hill 2006: 46). Since the sample at hand is quite small, and skewness can have a bigger impact, I take a closer look at this assumption (Midtbø 2012: 114). In the resource rents model density plots indicate that the residuals are distributed quite normal, but the “peek” might be caused by outliers[61]. A similar distribution is found in the production model[62]. I test the normality assumption by using the Shapiro-Wilk W test for normality (Chen, Ender, Mitchell, and Wells 2003). In both models, the p-values is 0.00, indicating that we have to reject the normality assumption. In short, the test and graphics indicate that the residuals not normally distributed. In order to see if the violation of the normality assumption influences testing of significance, I use a specified bootstrapping model in testing the robustness of the results.

Heteroscedasticity and autocorrelation

Two of the three OLS regression assumptions regarding the residuals are often not met when using panel data; that the residuals are homoscedastic (have the same variance) and are independent of each other. Independence in the case of panel data means that residuals for one unit, in our case country, at one time has to be unrelated to residuals for that country at all other times (no serial correlation) and that the residuals for one country are unrelated to the residuals for other countries (no spatial correlation) (Beck and Katz 1995: 636; Stock and Watson 2015: 411–412).This type of residual correlation in panel data is often referred to as autocorrelation, and can lead to biased standard errors.

---

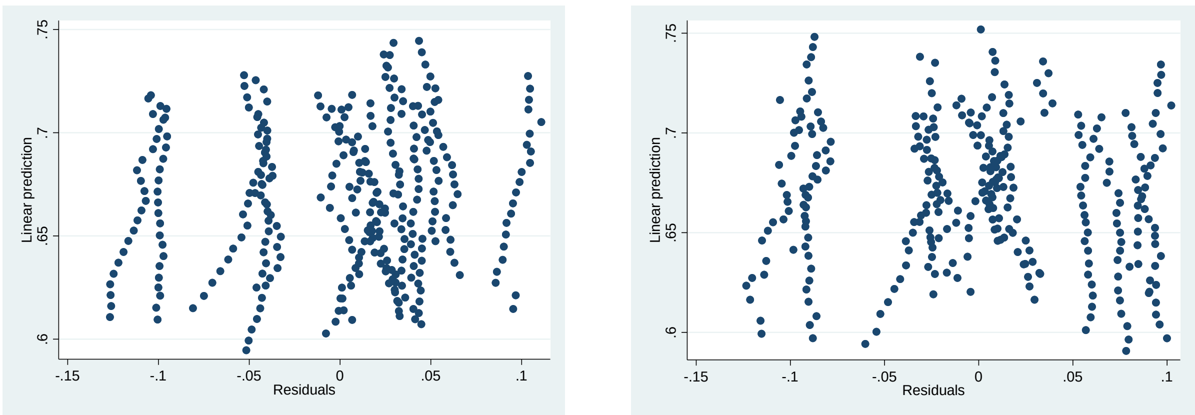
[60] Since zero is a common value for these variables, I use log+1 in order to prevent 0-values to be declared as missing.

[61] The Kernel Density plots of standardized and studentized residuals are shown in appendix 3. Inter-quartil-ranges using “iqr” in Stata indicates that the model includes 24 severe outliers. Dropping out Bolivia in an additional test gives is 14 severe outliers.

[62] The “icr” detects 23 severe outliers in the production data model.

As expected, heteroscedasticity is visible in both models graphically presented in the figure below, as clusters[63]. Testing for heteroscedasticity also give significant results, indicating that we should adjust the model[64]. Autocorrelation in the residuals is confirmed using the Wooldridge test, a test for autocorrelation in panel data models (Wooldridge 2010). For OLS to be optimal in the case of autocorrelation and heteroscedasticity, correcting the standard errors in the model, using robust, clustered or panel-corrected standard errors are common solutions (Stock and Watson 2012: 33). Clustered errors allow arbitrary correlation between residuals within identified clusters (Long and Freese 2006). I fit the models using cluster robust errors identified by each unit (country). In evaluation the robustness of the results, I also test the models using panel-corrected standard errors and bootstrapped standard errors.

Figure 4.5: Investigating heteroscedasticity with residuals versus estimates plot
left plot: Resource rents model right plot: production data model



Multicollinearity

Multicollinearity implies high (but not perfect) correlation between two or more independent variables (Kennedy 2003: 192–193). Perfect multicollinearity indicates that one of the independent variables is a combination of one or more of the other independent variables (Christophersen 2013: 77). As long as the correlation between independent variables is not perfect or exact, this does not violate any of the assumptions for using OLS-models, but it can lead to estimation problems, such as high standard errors (M. Wooldridge 2009: 97). High multicollinearity can make the variance in the regression coefficient large and so forth, more difficult for us to draw valid conclusions (Christophersen 2013: 77). To determine if

[63] The clusters in figure 4.5 are individual for each country (residuals for one country has the same pattern)
[64] The Breusch-Pagan / Cook-Weisberg test for heteroscedasticity is significant for both models. Results are presented in appendix 3.

multicollinearity poses a problem in our analysis, I produce collinearity diagnostics that are presented in table 4.2-4.5[65]. Estimation of the Variance Inflation factor (VIF) and tolerance values can give us a clue to how problematic multicollinearity is in our case. A "liberal" rule use VIF values above 10 as guideline for concern, while more conservative rules states that VIF values should not exceed 4 (J. M. Wooldridge 2009: 99). I use the latter as a guideline. In the resource rents model GDP (log) has the highest VIF value (1.67), which is below any of the mentioned "rules". The tolerance values tell us the same story: GDP (log) has the value closest to cero, with 0.5972, where values close to 0 suggest that multicollinearity may be a threat; whereas a tolerance values close to 1 indicates little multicollinearity. In light of these diagnostics, we can assume that multicollinearity does not pose a big threat in the resource rents model.

Table 4.2: Collinearity diagnostics natural resource rents model

Variable	**VIF**	SORT VID	Tolerance
Governance	**1.36**	1.17	0.7327
Natural res. rents	**1.22**	1.10	0.8213
Democracy	**1.13**	1.06	0.8885
Population (log)	**1.18**	1.09	0.8440
GDP (log)	**1.67**	1.29	0.5972
Time	**1.20**	1.10	0.8318
Mean VIF	**1.29**		

Table 4.3: Collinearity diagnostics natural resource rents model with interaction term

Variable	**VIF**	SORT VID	Tolerance
Governance	**2.61**	1.62	0.3833
Natural res. rents	**10.79**	3.28	0.0927
Nat.Rents_Governance	**11.41**	3.38	0.0876
Democracy	**1.18**	1.09	0.8473
Population (log)	**1.18**	1.09	0.8440
GDP (log)	**1.70**	1.30	0.5893
Time	**1.20**	1.10	0.8316
Mean VIF	**4.30**		

[65] The package «collin» can be downloaded for Stata (15) and produces collinearity diagnostics, such as VIF-values and tolerance values for all variables in a model.

When the interaction term is included in the resource rents model, it is inevitable that multicollinearity presents itself, since the interaction is the product of two other variables in the model[66]. This can lead to inflation in the standard error and p-values, but removing the interaction term is a misguided remedy (insert fox 1991:15). Several scholars defend the inclusion of interaction terms and argue that the concern about multicollinearity in interaction models have been overstated (Brambor, Clark, and Golder 2006). The standard errors will still be the "true" standard errors when multicollinearity is high. The high errors just imply lack of information and do not bias the coefficients. Therefore I will include the interaction term, despite the threat of multicollinearity the high VIF values express.

Table 4.4: Collinearity diagnostics production data model (all variables)

Variable	VIF	SORT VID	Tolerance
Petroleum	5,33	2,31	0,1877
Natural gas	5,08	2,25	0,1968
Gold_Silver	2,13	1,46	0,4704
Minerals	4,03	2,01	0,2482
Governance	1,76	1,33	0,5678
Democracy	1,13	1,06	0,8829
W. Prices oil	8,07	2,84	0,1239
GDP (log)	1,98	1,41	0,5046
Population (log)	5,79	2,41	0,1728
Time	8,13	2,85	0,123
Mean VIF	4,34		

The collinearity diagnostics for the production data model implies that "World prices Oil" (8.07), "Population" (5, 79) and "Time" (8.13) might pose multicollinearity issues (table 4.4). Petroleum, Natural gas and Minerals do also have high VIF values (between 4 and 6). Based on the evaluation of potential omitted variable bias (next section), I produced additional collinearity diagnostics for the model as it will be used in analysis (table 4.5). As we can see, the VIF values are lower in this model. The highest VIF-score is now "Petroleum" (4.17) and "Natural Gas" (4.80). They are both in a "grey zone", but I accept the potential risk, since they are essential in the model.

[66] The interaction term has 11.41 as a VIF-value and Natural resource rents 11.41.

Table 4.5: Collinearity diagnostics final production data model

Variable	VIF	SORT VID	Tolerance
Petroleum	**4.17**	2.04	0.2399
Natural gas	**4.80**	2.19	0.2085
Gold_Silver	**2.00**	1.41	0.5011
Minerals	**2.69**	1.64	0.3715
Governance	**1.64**	1.28	0.6102
GDP (log)	**1.95**	1.39	0.5140
Time	**1.24**	1.11	0.8076
Mean VIF	**2.64**		

Omitted variables and model specifications

There should not exist underlying variables that are the cause of the dependent variable and furthermore correlated with the independent variable in the model, also referred to as omitted variable bias (Skog 2010: 237). In short, this deals with whether or not we have specified the model with the inclusion of relevant (or not). The fixed effects model control for omitted variable bias, since FE controls for variables that varies between units, but are constant for each unit through time (Christophersen 2013: 174). However, omitted variable bias can still occur if relevant variables that varies through time for each unit is not part of the model.

The question of omitted variable bias is first and foremost a question that has to be answered based on existing theory on the relationship we study (Skog 2010: 271–72). Statistical tools might however help us in identifying variables that should be given more attention. An option is using Regression Specification error test (RESET), such as ovtest and ovtest rhs, who some argue are general model specification tests (Lee and Hill 2008: 160; J. M. Wooldridge 2009: 304-05), or linktest. Midtbø (2012:131) on the other hand suggest that these test are more useful in light of the linearity assumption. In the resource rents model, the two types of RESET-test indicate a model specification error, but when I run the linktest I get conflicting results. Changes in the results indicate that removing population and adding access to sanitation[67] might improve the model[68]. Another way to investigate omitted variable bias is by comparing r^2 and adjusted r^2 for models with and without different variables/ variables we are

[67] Access to water and sanitation was tested in the model in light of multidimensional social development measures including these dimension, such as ECLAC’s measurement (see section 3.4 and 3.5)

[68] Ovtest was resulted not significant when removing population (see appendix). I also checked other potentially relevant variables, such as access to water and unemployment (see do-file), that did not improve the model.

concern about[69]. I pay most attention to the adjusted r2 (due to small sample size). In the resource rents model, the adjusted ^{r2} decreases when population is added (the same goes for access to sanitation), implying that population does not give the resource rents model more explanatory power. Both previous concerns (linearity, normality, outliers), and the tests I have conducted, point in the direction that population might cause issues in the resource rents model. From a theoretical point of view, population is one of the control variables with least importance in light of the research question, so I rather continue without it to prevent biased results. Another point worth mentioning is that the comparisons of models confirm that including a time trend strengthens the model.

In the production data model, I first perform the statistical tests without Population and World prices[70]. The ovtest rhs is the only test that results significant (figure 7.5, appendix 3). As a precautionary measure, I investigate if including Population and World prices changes the results in these tests. By including Population, all tests turn out significant and when the same is conducted with the inclusion of World Prices; two out of three tests turn out significant. Including both variables at the same time also results in significant ovtest, ovtest rhs and linktest, pointing in the direction of exclusion of these variables in order to have a better specified model[71].

When comparing the model's explanatory power, another possible irrelevant variable presents itself: Democracy. When Democracy is included to the model the adjusted r^2 decreases. The same occurs when population is added to the model. When World Prices is added on the other hand, the explanatory power increases, but the change is trivial (appendix 3). Based on the same assessment as the resource rents model, I will conduct the main analysis without population as a control variable. The rationale behind including world prices on oils as a control was partly to compensate for the production data's lack of ability to describe the value of natural resources. World Prices for oil, is not satisfactory in actually doing this, and in addition increases the collinearity, I will exclude the variable from the analysis. The aim of the production model is to see if production of different types of natural resources leads to different outcome on social development. In this model, the governance perspective is not as

[69] A regression models explanatory power is often described by how much of the variations in the dependent variable that is explained by the independent variable. This is measured by r2 ("sum of squared errors"), see Skog 2010 (224).
[70] One reason is that these variables produces high VIF scores when investigating collinearity
[71] The results of additional ovtest and linktest are not part of the appendix, but is reported in the do-file.

important as in the resource rents model, and I therefore exclude Democracy from the production data model. The VIF scores for the "final" model can be seen in table 4.5. The excluded variables will experience a "comeback" in expanded models when investigating how resilient the results are.

4.2.3 Summary

This section has investigated unusual values that might influence the results and tested the OSL assumptions, with specific focus on characteristics that are common for panel data. Due to heteroscedasticity and autocorrelation in the residuals, I use cluster robust standard errors. Transforming variables was seen as a useful, both in light of linearity, normality and influential values. Multicollinearity is common in panel data, but fortunately, proved to be minor in the resource rents model. This was however more prominent in the production data model. By investigation of omitted variable bias (and collinearity), I concluded on the exclusion of "Population" from the models, as well as "World Prices" and "Democracy" from the production data model.

4.3 Regression analysis

The main focus of this thesis is to investigate how riches from natural resources affect social development in countries in Latin America, with specific focus on governance as an intermediate variable that can influence this relationship. An obvious objective is to increase the possibility to of obtaining valid conclusions, a reason for devoting considerable effort in defining and operationalizing the relevant concepts and adjusting the model to reduce the risk of omitted variable bias. In this section I present the results of fixed effects analysis with clustered standard errors, structured according to the three research questions and related hypotheses. As a guideline in testing of statistical hypotheses, I will follow what is customary in the social sciences and use 5% significance level as a basis for either rejecting or accepting the null hypothesis in each case[72].

[72] 10 %, 5% and 1% percent significance level are usual in social sciences, and according to Stock and Watson (2015: 124) as well as Christophersen (2013: 29) 5% is considered the most common of the three.

4.3.1 Does natural resource riches have a positive effect on social development?

The rationale behind the first research question and hypothesis is to study how riches from natural resources affect social development. Some researchers argue that there exists a negative relationship between natural resources and GDP[73], whilst other claims the opposite[74] or indicate that the relationship is at least not as direct as previously assumed[75]. Since previous research has concluded differently about this relationship, it is interesting to study it further, in a way that still makes comparisons possible.

I have shown in earlier sections that there are many ways to define development, and at least as many methods to measure it. The Human Development Index (HDI) is a well-established measurement that coincides with this thesis' theoretical definition of social development. By using the HDI as a dependent variable, I use an indicator that includes more dimensions than indicators based on GDP growth, as the ones used in the studies of Sachs and Warner (1995) and Mehlum et al. (2006)[76].

When investigating the relationship between natural resources and economic development, scholars have used export-based measures[77], reserve data[78], and more recently, resource revenues or rents as share of GDP[79]. Based on the theoretical discussion, I have chosen to use natural resources rents as percentage of GDP as indicator in the analysis. One of the benefits by using this indicator is that it allows me to compare results with previous research. In addition to natural resource rents, the most important intermediate variable in the analysis is "Quality of Governance". The basis for this indicator is a state centric view of governance and as explained in section 3.4.2., this indicator covers many of the same dimensions as the institutional indicator used by Mehlum et al. (2006). In light of the literature review, the logical assumption would be that the higher quality of governance, the higher the score on social development[80]

---

[73] See Sachs and Warner (1995, 1999, 2001), Auty (1990) or Gelb (1988)
[74] See Mikesell 1997 or Brunnchweiler and Bulte (2008)
[75] See van der Ploeg (2011), Mehlum et al. (2006), Lederman and maloney (2007) or Stijns (2005).
[76] See section 2.1 and 3.3 for more info
[77] See Sachs and Warner (1995, 1999, 2001) and Mehlum et al (2006)
[78] Maloney (2002), Wright and Czelusta (2004) and Stijns (2003) have promoted alternative measures to the export-based ones
[79] Aslaksen (2010), Haber and Menaldo (2011) and Ross (2001)
[80] Based on the arguments of for instance van der Ploeg (2011) , Lederman and Maloney (2007) and Mehlum et. al. (2006) the quality of governance or institutional quality can have a positive impact on social development, also in light of how natural resources are managed by the Government.

With the necessary adjustments of the model in order (see section 4.2), I can start investigating the results closer and testing the first hypothesis: *Riches from natural resources have a positive effect on social development.* In order to investigate the individual effects of each variable more carefully, I introduce one variable at the time to the regression. In the following table the results are presented:

Table 4.6: regression models, resource rents mod

Variable	*FE (1)* *HDI*	*FE (2)* *HDI*	*FE (3)* *HDI*	*FE (4)* *HDI*	*FE (5)* *HDI*
Natural resource rents (% GDP)	0.003**	0.003**	0.003**	0.000	-0.000
	(3.93)	(3.10)	(3.41)	(0.39)	(-0.84)
Governance		-0.045	-0.046	-0.083+	0.003
		(-0.65)	(-0.66)	(-1.99)	(0.17)
Democracy			0.012	0.001	-0.006
			(0.79)	(0.13)	(-1.21)
GDP (log)				0.196***	0.014
				(8.91)	(0.90)
Time					0.005***
					(10.23)
Constant	0.655***	0.672***	0.661***	-1.072***	0.499**
	(128.31)	(18.07)	(15.14)	(-5.56)	(3.85)
Number of obs.	374	352	350	330	330
r2	0.133	0.149	0.158	0.752	0.944
r2 adjusted	0.131	0.141	0.150	0.749	0.943

Notes: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table report regression coefficients with clustered robust standard errors in parentheses.
The number of obs. decreases in model two due to missing Governance values for Belize
The number of obs. decreases in model three due to missing GDP values for Argentina

Model 1 consists only of the dependent variable (HDI) and the independent variable natural resource rents. The regression coefficient, 0.003, indicates a weak positive effect between increases in natural resource rents on HDI. Even though the coefficient is significant at a 1% significance level, this is not sufficient to draw inference since no control variables are included in this model. The effect is very close to zero (no effect), which might indicate that my results point in the same direction as for instance the works by Lederman and Maloney (2007), who concludes on “neither curse nor destiny”.

In model two, Governance (ICRG Quality of Governance indicator) is added to the regression. The positive effect of natural resource rents on HDI remains significant when introducing Governance. A surprising finding is the negative coefficient for governance on HDI. In light of studies by for instance Mehlum et. al (2006), van de Ploeg (2011) and Maloney (2007), I expected that higher scores on governance quality would lead to higher HDI scores. This is not what my results indicate. It is never the less a very weak negative effect and it is not significant (not even at a 10% level).

In model three, the control variable Democracy, is introduced to the regression. As we can see, the positive effect of natural resource rents on HDI remains significant, and unchanged, when controlling both for Governance and Democracy. The effect of Governance on HDI on the other hand, is slightly strengthened between model 2 and 3. However, this effect is still not significant. Democracy has a small positive effect (0.012) on HDI in this model, but the effect is not significant. Democracy has been argued to be a key for successful development[81]. My results seem to draw more in the direction of no relationship between democracy and social development.

In model four, when GDP (log) is added to the regression, some rather noteworthy changes occur. The effect of natural resource rents on HDI changes from a positive effect to no effect, and the coefficient is no longer significant. The coefficient for Governance on HDI is strengthened between model 3 and 4 (from -0.046 to -0.083). This coefficient is not significant at a 5% level, but is significant at a 10% level. The effect of Democracy on HDI is reduced from model 3 to model 4, indicating a close to zero effect, but positive effect. As in model 3, the coefficient for democracy is not significant. GDP has the strongest effect in model 4 (0.196), an effect that is significant at 0.1 % level. Model 4 is a quite good illustration of how a variable that is very influenced by previous values (country-year observations) can cause problems in the model[82]. The results in model 4, especially compared to model 5, display that the inclusion of a time trend is a better fit for the data at hand.

In model five, time is included in the regression. By taking into account a linear time trend, the effect of GDP (log) on HDI is no longer significant and is reduced to a small positive

[81] See Bovir (2008) for instance
[82] The constant term (intercept), which is seldom relevant, has a negative value in model 4, and deviates clearly from the other models. Still, the intercept coefficient has no (or trivial) interpretation, especially not in FE models.

effect (0.014). The effect of natural resource rents on HDI has in model 5 changed proceeding sign, indicating a trivial negative effect. Trivial, because the regression coefficient is -0.000 and it is no longer significant. The estimated effect of Governance on HDI has changed from a negative, to a slightly positive effect (0.003), and is not significant. In model 5, the regression coefficient for Democracy's effect on HDI has changed proceeding sign, from a very small positive effect to a small negative effect (-0.006). As in all the other models, the regression coefficient for Democracy on HDI is not significant. Time is the only independent variable with a regression coefficient that is significant at a 0.01 % level in model five. Not surprisingly, time has a small (0.005) positive effect on HDI.

The results presented in this section illustrates that the effect of natural resource rents on HDI is very small. However, differences in proceeding signs in the estimated effects for natural resource rents on HDI in model 1-4 and 5 makes concluding on either rejection or acceptance of the first hypothesis uncertain. In continuation, I find it surprising that Governance has a negative effect on HDI in model 2-4, and only a slight positive effect in model 5. I expected a positive effect in all models. In determining how robust these results are, one could add more variables to the model (but this reduces the efficiency). Another aspect, discussed in previous sections, is the option of lagging either the dependent or independent variables and see if the results change. Influential values (see section 5.2.1), especially for Bolivia, could be conflicting the model, as well as the choice of model (Fixed effects). As a precautionary measure, I create extensive models in order to test the robustness and sensitivity of the results, before I draw conclusions concerning the relationship between natural resource rents and social development. The results of the extensive models are presented in section 4.3.4.

4.3.2 Does the quality of governance condition the relationship between riches in natural resources and social development?

Scholars such as Leftwich (1993) and Johnson (1982) have argued that a strong state is essential for a country's development. Mehlum et al. (2006) has pointed out that the effect of natural resource abundance is determined on factors such as the quality of institutions, and other point in the direction of government policies or management of natural resources as important elements (van der Ploeg 2011, Lederman and Maloney 2007). This makes is relevant for me to investigate if a positive effect of natural resource rents on HDI is conditioned by high scores on governance. In the following, I will investigate the second hypothesis with the assumption that *riches from natural resource have a positive effect on social development, when the score on governance is high.*

In order to study this hypothesis, I extend the model by including an interaction term between Natural resource rents and Governance. Interaction effects are often both empirically relevant and interesting, but they can be tricky to interpret (Jaccard and Turrisi 2003; Kam and Franzese 2007). "An interaction effect is said to exist when the effect of the independent variable on the dependent variable differs depending on the value of a third variable, called the moderator variable" (Jaccard and Turrisi 2003: 3). In my case, a natural resource rents is the independent variable with main "focus" and Governance is the "moderator"[83] that potentially influences natural resource rents effect on HDI. The results of including an interaction term are displayed in the following table:

Table 4.7: Regression with interaction term

Variable	*FE (5)* *HDI*	*FE (6)* *HDI*
Natural resource rents (% GDP)	-0.000	-0.000
	(-0.84)	(0.38)
Governance	0.003	0.011
	(0.17)	(0.55)
Democracy	-0.006	-0.005
	(-1.21)	(-1.22)
GDP (log)	0.014	0.018
	(0.90)	(0.98)
Time	0.005***	0.005***
	(10.23)	(9.23)
Interaction term (Nat*Gov)		-0.001
		(-0.68)
Constant	0.499**	0.454**
	(3.85)	(2.81)
Number of obs.	330	330
r2	0.944	0.944
r2 adjusted	0.943	0.943

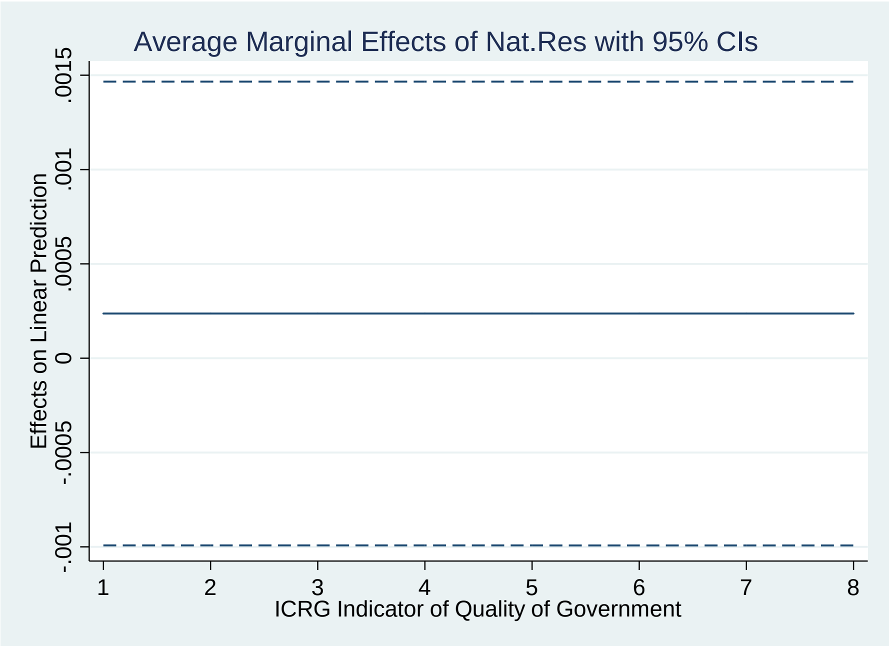
Notes: $^{+}$ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001
The table reports regression coefficients with cluster robust standard errors in parentheses.

With one scale unit increase in governance, the effect of natural resources on HDI decreases by 0.001. The coefficient for the interaction term is almost as close to zero effect that one

[83] To simplify the interpretation of interaction effects, Jaccard and Turrisi (2003:3) suggest to distinguish between focus and moderator variable in the interaction term.

could get, but it is still negative, pointing in the direction that the higher the on governance, the weaker effect of natural resource rents on HDI[84]. Brambor et al. (2006) suggest using margin plots to evaluate the interaction effects, since the interaction coefficients only show us a fragment of the picture. Margin plots can for instance help us to determine if different values on the moderator (governance) influence the factor's effect (Natural resource rents) on the dependent variable (HDI) differently. The margin plot, figure 4.6, confirms that the interaction term has a zero effect and in addition tells us that the effect is the same throughout the whole governance scale.

Figure 4.6: Interaction affect between Natural resource rents and Governance



Since the coefficient is negative, very small, and not significant, we cannot give any preliminary support to the second hypothesis that assumes that an increase in governance quality, leads to a positive effect for natural resources on HDI. The coefficient for the interaction term is not significant, so rejecting the hypothesis is also difficult. It is still quite surprising that higher scores on Governance alone, has such a small effect on HDI, and that the results gives us little evidence of governance being important as a moderator for natural resources effect on HDI. That is the inference the results so far let us determine, the one of little or no effect. The interaction term will also be a part of the extensive models in section 5.3.4.

[84] This interpretation in based on Midtbø (2012: 137).

4.3.3 Can production of different types of natural resources affect social development differently?

The previous models have investigated how riches (rents) from natural resources affect social development. The additional question this thesis seeks to answer is if production of different types of natural resources leads to different outcomes on social development, with the diverging results from studying minerals and fuels separately as a backdrop[85]. In the following I investigate the third hypothesis that assumes that: *Production of different types of natural resources effects social development differently*. The results are presented in table 4.8

Table 4.8 regression models: Production data

Variable	*FE (1)* *HDI*	*FE (2)* *HDI*	*FE (3)* *HDI*	*FE (4)* *HDI*	*FE (5)* *HDI*	*FE (6)* *HDI*	*FE (7)* *HDI*
Petroleum	0.007*** (4.79)	0.006*** (4.58)	0.006*** (4.26)	0.006*** (4.15)	0.006** (3.77)	0.002 (1.75)	-0.000 (-0.36)
Natural gas		0.012*** (5.54)	0.011*** (4.39)	0.011*** (4.58)	0.011*** (4.36)	0.002 (0.83)	0.001 (0.52)
Gold & Silver			0.000 (0.28)	0.000 (0.03)	0.001 (0.52)	0.000 (0.39)	0.000 (0.00)
Minerals				0.003* (2.91)	0.003** (3.03)	0.001 (0.81)	0.001* (2.16)
Governance					-0.064 (-0.84)	-0.075^{+} (-1.84)	0.017 (0.70)
GDP (log)						0.189*** (8.95)	0.006 (0.43)
Time							0.005*** (12.74)
Constant	0.638*** (82.48)	0.610*** (63.41)	0.610*** (57.92)	0.597*** (55.22)	0.623*** (18.86)	-1.028*** (-5.57)	0.546*** (4.38)
Number of obs.	349	344	327	311	311	311	311
r2	0.114	0.192	0.177	0.199	0.219	0.781	0.945
r2 adjusted	0.112	0.187	0.170	0.189	0.207	0.777	0.943

Notes: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with clustered standard errors in parentheses.

The number of obs. decreases in model three due to missing Gold_Silver values for Belize

The number of obs. decreases in model 4-7 due to exclusion of Argentina and Belize

(missing values on GDP, Governance, Minerals and Gold_Silver)

[85] See for instance Mikesell (1997), who argue that mineral exporting countries are the most visible victim of the resource curse.

In the first model, Petroleum is the only independent variable in the regression. When introduced alone, the coefficient for Petroleum is positive and significant at a 1% level (0.007). It is however a small positive effect. When a variable is log transformed, the interpretation of the regression coefficients change. Instead of referring to a "one unit scale increase" in the independent variable, one refers to a "one percent increase" in the dependent variable (Midtbø 2012: 135). So, when Petroleum production increases with one percent, the HDI is expected to increase with 0.007. Since the effect of petroleum on social development is positive, it represents a contrast to studies claiming that there exists a resource curse[86].

In the second model, Natural gas is added to the regression. The coefficient for Petroleum remains significant in this model as well, but with a very small decrease in power, compared to model 1. The coefficient for Natural Gas on HDI is also positive and significant at a 0.1% level. This indicates that an increase in the production of Petroleum or Gas, has a positive effect on HDI (without the control of other variables). So far, the similarity in estimated effects indicates that the third hypothesis (assuming difference in effects), might be rejected.

In the third model, Gold & Silver is introduced to regression. The positive, and significant, coefficients for Petroleum and Natural Gas on HDI still hold in this model. In the case of Petroleum, there is no change compared to model 2, and Natural Gas only experience a minor reduction of power (from 0.012 to 0.011). The estimated effect for Gold & Silver on HDI is 0.000, and it is not significant. This is a slight change from the small positive effects of Gas and Petroleum on HDI, but as long as the coefficient for Gold & Silver is not significant, drawing inference of differences in estimated effects is uncertain.

In model 4 Mineral production[87] is included in the regression. The regression coefficients for Petroleum (0.006) and Natural Gas (0.011) sustains when Minerals is added to the model. Petroleum, as well as Natural Gas, still has a significant effect on HDI on a 0.1% significance level. Gold & Silver remains unchanged from model 3 to 4, which means a zero effect on HDI (not significant). The coefficient for Minerals on HDI is positive, and significant (5% level). The effect is nevertheless very small and close to "no effect".

Governance is introduced to the regression in model 5. A bit surprising, but uplifting in light of robustness, is that the coefficients for production of Petroleum, Natural Gas, Gold & Silver

[86] See Sachs and Warner (1995, 1999, 2001), Auty (1990) or Gelb (1988)
[87] Minerals is the sum of production of 8 different minerals, for details see section 3.4.1

and Minerals experience close to no change when Governance is added to the model. Petroleum still has the same effect on HDI as in model 3 and 4; the only change is that the coefficient is significant at a 1% level (no longer 0.1%). Natural Gas has the same coefficient as in model 3 and 4, and it is still significant at a 0.1% level. When it comes to Gold & Silver, the effect has a trivial increase from model 4 (from zero to 0.001). As in previous models, the coefficient for Gold & Silver on HDI is not significant. Governance has a negative effect on HDI (-0.064), an effect that is not significant. This is quite similar to the findings of the resource rents models, when no other control variables were included.

When GDP is added to the regression in model 6, the majority of the coefficients loose significance (compared to previous models). The exception is the coefficient for Governance on HDI (-0.075) that is significant at a 10% level. GDP is the variable with the strongest effect on HDI (0.189) in model 6, an effect that is significant at a 0.01% level. The déjà vu from the resource rents model (4), when GDP was introduced there, is apparent. GDP affect the resource rents model and the production model very similarly, with a strong significant effect when it is introduced, an effect that disappears when Time is introduced in the last model.

When Time is included to the regression, there is only one other variable that has significant effect on HDI, namely Mineral production. The regression coefficient for Minerals is as close to zero as it gets (0.001). Time on the other hand, has a small positive effect on HDI (0.005) that is significant at a 1% level. In short, the complete model indicates that production of natural resources, regardless of the type, does not affect HDI. The differences in coefficients for Petroleum, Gas, Gold & Silver and Minerals are trivial the way they are specified in these models. The general picture is a very, very small positive effect, where Minerals is the only one that has a significant effect.

Based on previous research, I expected to find more signs of dissimilarity in how production of different types of natural resources affects social development (HDI). Admittedly, the findings in the production model are not directly comparable to the resource rents model, as the production model variables do not say anything about the riches from natural resource production. Any differences that my model could not detect might have been manifested if I had used Mineral Rents, Petroleum Rents and so forth. On the other hand, the production model confirms the first impression from the finding in the resource rents model, namely if there exist any relationship between natural resources and social development, it's microscopic. However, the lack of findings might be related to slow changes in the dependent variable. I therefore find it relevant to use lagged variables in the extensive models. A preliminary conclusion is that the zero findings indicate that we cannot accept the hypothesis that the production of different types of natural resources has a different effect on social development.

4.3.4 Extensive models

The results presented have illustrated that the inclusion of different control variables influences the estimated effect of the independent variables on social development. Although some researchers advice against including too many variables in a model (Hosmer and Lemenshow 2000) , I still want to run extensive models with considered control variables in order to see if that changes the regression estimates of the independent variables. The results of the extensive resource rents model can be seen in table 4.9 on the next page. The results of the two other models are presented in appendix 4.

None of the "additional" variables have significant effect in the models, and to a little extent improve the models explanatory power. Most important, the coefficients for the independent variables of interest do not change when additional control variables are included. This can in other words be regarded as a first "confirmation" of the results stability. In addition, the estimation of extensive models confirmed that it was a good idea to exclude several control variables from the models[88].

[88] This is not very surprising in light of the model comparison carried out in section 4.2.2.

Table 4.9 extensive Resource rents regression models

	FE (1)	*FE(2)*	*FE (3)*	*FE(4)*	*FE (5)*	*FE(6)*
	HDI	*HDI*	*HDI*	*HDI*	*HDI*	*HDI*
Natural resource rents (% GDP)	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(-0.84)	(-0.84)	(-0.84)	(-0.87)	(-0.83)	(-0.91)
Governance	0.003	0.003	0.004	0.004	0.003	0.004
	(0.17)	(0.17)	(0.20)	(0.20)	(0.17)	(0.19)
Democracy	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006
	(-1.21)	(-1.21)	(-1.04)	(-0.96)	(-1.26)	(-1.02)
GDP (log)	0.014	0.013	0.012	0.013	0.012	0.010
	(0.90)	(0.92)	(0.86)	(0.84)	(0.71)	(0.63)
Time	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
	(10.23)	(5.56)	(8.22)	(11.13)	(9.61)	(5.90)
Population (log)		-0.010				-0.009
		(-0.18)				(-0.14)
Access Sanitation			-0.000			-0.000
			(-0.19)			(-0.06)
Access Water				-0.000		-0.000
				(-0.12)		(-0.02)
Unemployment					-0.000	-0.000
					(-0.11)	(-0.15)
Authority taxes						
Constant	0.499**	0.665	0.516**	0.511**	0.511**	0.676
	(3.85)	(0.75)	(3.91)	(3.43)	(3.26)	(0.68)
r2	0.944	0.944	0.944	0.944	0.944	0.944
N	330	330	330	330	330	330
r2_a	0.943	0.943	0.943	0.943	0.943	0.942

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with cluster robust standard errors in parenthesese.

The number of obs is 330 in all models due to missing GDP values for Argentina and missing Governance values for Belize

4.4 Robustness

There might be characteristic of the data, the choice of statistical method or specifications in the model that influence the results. Relying on just one model specification or method might misguide us when drawing conclusions. In light of these aspects, I evaluate the stability of my results with the aim of drawing final conclusions based on a solid foundation.

4.4.1 Initial checks

To evaluate the stability of my results, I will first investigate changes in the estimated effects when including lagged dependent variable and independent variables. Then I will replace "time" with five-year / ten-year periods to see if the estimated effects change. As a final part of these initial assessments, I will run comparisons of the models with and without influential observations, before moving on to alternative ways of estimating the models.

First, I check if the estimated effects change when using a one- and two-year lagged dependent variable[89]. The rationale behind this is that the estimated effects are very small or equal to zero and by lagging the dependent variable we might be able to detect “slow” effects that were not visible in the original models. A downside by doing this is that it reduces the number of observations, since each lag excludes first year observations on the variable that is lagged. In the resource rents model[90] the use of lagged dependent variable does not change the estimated effects of natural resource rents on HDI, indicating that the negative “zero effect” holds. The estimated effect of Governance on HDI only experiences a slight increase, from 0.003 in the original model, to 0.004 in the models with lagged dependent variable. The biggest changes in estimated effects are in the control variables, especially a reduction of the estimated effect of GDP, from a positive effect (0.014) to a negative effect on HDI (-0.02). The GDP coefficient is not significant in any of the models. A similar change is visible when evaluating the use of lagged HDI in the production data model[91]. The estimated effects of different natural production data only experiences trivial changes, but some reduction in Governance and change in GDP (from positive to negative) is visible. GDP and HDI correlate with each other and follow the same trend, so it is not surprising that the effect of GDP on HDI decreases when HDI is lagged with one and two years. What is surprising is that the

[89] As there is little difference in the resource rents model with and without interaction effects, not all robustness tests are run on the latter, with the intention of streamlining the work. The exception is when deviance occurs in the variables that are included in the interaction term.
[90] Table 8.3 in appendix 4 present the results of the resource rents model using lagged HDI
[91] Table 8.5 in appendix 4 present the results of the production data model using lagged HDI

effect of GDP on HDI turns, from a positive to a negative. In spite of this, the use of lagged dependent variable verifies the initial estimated effects of natural resources and governance on HDI. In short, my main results still holds.

As a second, precautionary measure, I investigate if the estimated effects change when the independent variables are lagged. In the resource rents model, the biggest change from the original model is changes in GDP[92], but the changes are trivial compared to the changes in GDP when using lagged dependent variable (HDI). The estimated effect of Natural resource rents on HDI remains unchanged when using lagged Natural resource rents and lagged Governance. The estimated effect of Governance on HDI changes from 0.003 to 0.006 in the model where Governance is lagged with one year. The effect is still not significant and confirms the initial null results between Governance and HDI. In the production model with lagged independent variables, there is no change in the estimated effects of the natural resource production variables[93]. The significant effect of Mineral production on HDI still holds when Mineral production is lagged with one year. The estimated effects of Governance and GDP experience some minor changes, but neither of these variables have significant effect in any models.

The estimations with lagged dependent and independent variables indicate that my models capture the relationship (or no relationship) between the independent and dependent variables quite good and that "slow variation" is excluded as a possible reason for null findings.

The evaluation of my results continues with replacing "Time" with two time periods (as a dummy variable). This does not change the estimated "zero effect" natural resource rents have on HDI, but the negative effect of Governance is stronger when the time dummy is introduced (however, not significant)[94]. I do the same for the production data, where only microscopic changes in the estimated effects of production data on HDI appear[95]. The estimated effect of Governance on HDI changes from an insignificant positive effect to an insignificant negative

[92] Table 8.4 in appendix 4 present the results of the resource rents model using a one year lag on resource rents and governance.
[93] Table 8.5 in appendix 4 present the results of the production data model using a one year lag for production variables and governance.
[94] Table 8.6 in the appendix present the results when two time periods replace the time dummy in the resource rents model
[95] Table 8.7 in the appendix present the results when two time periods replace the time dummy is introduced in the production data model.

effect when the time dummy is introduced; indicating that increased value on Governance has a stronger negative effect in the 1990s than after 2001.

Lastly, I estimated the models without influential values, identified in section 4.2.1, in order to evaluate if influential values for a country could be leading the results. In the resource rents model, four additional regressions were conducted[96]. The proses led to additional confirmation of no relationship between resource rents and development. The estimated effect of Governance on HDI increased (from 0.003 to 0.018) in the model where Bolivia was excluded and in the model where Venezuela was excluded, but remained more stable when Guatemala and Paraguay were excluded. Still, none of the Governance coefficients turned out significant, approving the original findings. In the Production model[97], only slight changes occur in the coefficients for Petroleum, Gas, Gold & Silver and Minerals when countries are excluded. The only noteworthy change is that the estimated effect of Governance is significant (5% level) and positive (0.039) when Venezuela is excluded from the analysis. Since Venezuela also is excluded from the resource rents regressions and the Governance indicator is part of both models, I find it a bit strange that the impact of removing Venezuela resulted differently in the two models. However, it is the dramatic "ups and downs" in the Governance indicator from 2002 to 2012 that are the cause of Venezuela's influential values. This is not surprising, in light of the general strike, lockouts and attempted political coup against Hugo Chávez in 2002 and the political instability that followed (Bull 2016: 110).

The initial checks indicate that my results are quite stable, especially concerning the lack of evidence of a causal relationship between natural resource rents and natural resource production on HDI. The only sign of sensitivity in my results concerns the significant positive effect of Governance on HDI, when Venezuela was excluded from the production model.

4.4.2 Model estimation alternatives

My results have so far proved to be robust, especially the findings of a non-existing relationship between natural resources and social development in Latin America. All the additional estimations have used the same framework as the main model, namely the fixed effects model with cluster robust standard errors. A famous example of changes in effects

[96] Resource rents: Bolivia, Guatemala, Paraguay and Venezuela were excluded separately. See table 9.0 in appendix 4.

[97] Production model: Bolivia, Brazil, Mexico, Peru and Venezuela were excluded separately. See table 9.1 in appendix 4.

when using different frameworks is found in Acemoglu and Robinson (2008). When revisiting the previously proven positive relationship between income level and democracy Acemoglu, Johnson, Robinson, and Yared (2008) discovered that the effect disappeared when using fixed effects. In my case, it might be the opposite. Could my null findings between natural resources and social development, as well as governance and social development, be caused by the framework I choose?

Well, the answer is either yes or no. When comparing the results of the resource rents model using alternative models, all the models gave the same result concerning the relationship between natural resource rents and social development: 0.000[98]. The estimated effect of Governance on HDI however, changed and to some extent led to contradictory results. The biggest change is in the OLS PSCE model, where Governance has a positive effect on HDI (0.117), significant at a 0,001 level[99]. In the production data, all estimated effects in the OLS PCSE model are significant. If OLS PCSE was the only model, I would most likely have accepted hypothesis 3, that the production of different types of natural resources affects social development differently. However, OLS PCSE model does not control for country specific effects, which I argue is highly relevant since Latin American countries are diverse, both in sosio-economic factors, as well as in demography and geography. The Hausman test[100] supported the use of fixed effects, based on the nature of the data. In short, both theoretical and statistical arguments led to the conclusion of controlling for country specific effects.

4.4.3 Alternative dependent variable

This thesis presents both similarities and contrasts to previous research. The most obvious difference is my choice of dependent variable. Several contributions to the resource literature focus on natural resource’s effect on economic development, whereas I use social development as my indicator. The rationale behind the deduction of hypotheses and choice of independent variables, were also to a great extent based on what previous research had concluded on when investigating the relationship between natural resource wealth and economic development. In order to investigate if my results are conditioned by my choice of dependent variable, I run my models using GDP per Capita in constant USD 2011 as the

[98] The proceding sign changes in the models, but they all give zero results. See table 9.2 in the appendix.
[99] The autoregresive FE = -0.02 and bootstrapped FE = 0.003
[100] See section 3.8 and appendix 2.

dependent variable, switching place with the HDI as control variable. The variable is log transformed (see section 3.5).

In the resource rents model, with GDP as dependent variable, my zero findings regarding natural resource rents are confirmed. Interestingly, Governance has a strong positive effect on GDP, an effect that is significant at 5% level. In the production model, with GDP as dependent variable, Governance gets almost the same estimated effect on GDP, significant at a 1% level (0.391)[101]. The model indicates that the finding of Mehlum et al. (2006)[102] is true when a GDP based measure is used, but not when using HDI. These findings will be discussed in the next section.

Table 4.4.1: Alt. dependent variable, Resource rents with interaction

	Original HDI	Resource rents GDP (log)	Incl. Interaction GDP (log)
Natural resource rents (% GDP)	-0.000	0.003	-0.008
	(0.000)	(0.91)	(-1.76)
Governance	0.003	0.390*	0.184
	(0.020)	(2.93)	(1.40)
Democracy	-0.006	-0.009	-0.031
	(0.005)	(-0.28)	(-0.89)
GDP (log)	0.014	1.009	1.225
	(0.015)	(1.10)	(1.18)
Time	0.005***	0.015*	0.014*
	(0.000)	(2.67)	(2.34)
Rents*Governance			0.023**
			(3.41)
Constant	0.499**	8.061***	8.038***
	(0.130)	(13.51)	(12.30)
Number of obs.	330	330	330
r2	0.944	0.779	0.802
r2 adjusted	0.943	0.776	0.798

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table report regression coefficients with cluster robust standard errors in parentheses.

[101] The results are presented in appendix 4.
[102] a positive effect of natural resources on economic growth is condition by high institutional quality

4.5 Discussion

The goal of this thesis was to compliment an ongoing academic debate that has many contributions, more than the ones mentioned here, and diverging conclusions on the relationship between natural resources, development and intermediate variables. It is previous researcher's conclusions on these relationships that makes my "zero findings" interesting.

In light of the studies by Auty (1990, 2001), Sachs and Warner (1995, 1999, 2001) and Gelb et al. (1988) it would not be "far-fetched" to expect some similarity in the relationship between Natural resource rents (% GDP) and HDI, and the negative relationship they discovered between natural resources exports (as % total export) and economic development (GDP growth-based measures). On the other hand, there are those who claim that riches from natural resources have a positive effect on a country's development[103], indicating that positive findings could be likely. Lastly, there are also scholars who argue that a positive effect of natural resource abundance on economic development is conditioned either by political factors or institutional factors[104]. In short, the resource literature displays a diversity of potential positive or negative relationships, conditional or not. My main findings are on the contrary, above all characterized by a large number of zeros.

The results of the analysis and robustness tests give a quite clear answer to the first hypothesis. Based on my findings, there is little evidence to confirm that riches in natural resources have a positive effect on social development (HDI). Nor can I discard the hypothesis, and argue that it is the other way around. There is simply no measurable effect between Natural resource rents (% of GDP) and HDI in the Latin American countries included in my analysis, when I use Governance, Democracy, GDP and Time as controls. In light of previous research, I had expected a clearer indication of either a positive relationship, or negative one, between riches in natural resources and social development.

My second hypothesis was deducted in order to investigate if a positive relationship between natural resource rents depended on high governance scores. By using an interaction term in my model, with resemblance to the one used by Mehlum et al (2006), I expected to get similar results. My findings on the other hand, tell a different story. What might be evident when using economic development as the dependent variable is, not necessarily relevant when social development is used as the dependent variable. I was surprised when finding a close to

[103] See Mikesell (1997) and Brunnchweiler and Bulte (2008)
[104] See Mehlum et al. (2006), van der Ploeg (2011), Lederman and Maloney(2007)

non-existing relationship between the quality of governance and social development (see 4.2.3). As a result, my findings could not confirm the assumption that high scores on governance lead to a greater positive effect of riches from natural resources on social development. Still, it should be mentioned that the governance estimates were a bit more sensitive than the estimated effect of natural resources on social development[105].

The third hypothesis assumed that production of different types of natural resources would result in different outcomes on social development. The rationale behind investigating this was partly because the majority of scholars had focused on either fuels or minerals, or all natural resources, and secondly because claims had been raised of mineral producing countries being a more visible victims of the resource curse (Mikesell 1997). The production of minerals was the only variable with a significant effect on social development (among natural resource production measures). However, I would not claim that *one* significant effect of 0.001 reinforces Mikesell's argument. When investigating if the choice of framework influenced my results, the OLS PCSE model indicated that different production data had different effect on social development. However, both theoretical and statistical arguments led me to rely on the estimates from the fixed effects model, controlling for country specific effects. The OLS PCSE model could have led me to accept the third hypothesis. The stable results in other estimations and robustness test have, nevertheless, led me to conclude that there is no convincing evidence of the third hypothesis. In short, my results indicate that production of different types of natural resources effect social development the same way (no effect) in Latin American countries.

4.5.1 Absence of evidence, or evidence of absence?

"absence of evidence is not evidence of absence (…) all that has been shown is an absence of evidence of a difference" (Altman and Bland 1995). Altman and Bland (1995) wrote this in light of misinterpretation of differences that did not reach significance in health research. One of their arguments concerns the strict rules of drawing inference based on significance level, that can be almost impossible in health trials were the samples are small. The number of observations that are included in the analysis is relevant when discussing my results. Other contributions to the resource literature studied a wide range of countries and usually have a

[105] For instance, the estimates changed when Venezuela was excluded from the regression. In the OLS PSCE model, Governance had a significant effect on natural resources

longer time span, giving them thousands of observations[106]. In this thesis, only countries in Latin America were included and the time limit was to some extent reduced by the data availability. In addition, missing data reduced the number of observed values[107].If the argument of Altman and Bland (1995) was valid in my case; we would not see any significant effects at all. There is especially one estimated model I find interesting in light of efficiency, the model where GDP is used as dependent variable. If the lack of significant effects in the original models was caused by a small number of observations, we would most likely not find significant effects in regressions with the exact same indicators and number of observations. The strong positive effect of governance on GDP in that model put Altman and Bland's argument to a test, as well as the model represent some interesting findings.

4.5.2 Difference that matters

I can still answer the main research questions of the thesis, even if the null findings made confirming or discarding the hypotheses complicated. Based on my findings, riches from natural resources do not affect social development Latin American countries. This conclusion holds, even when I examine if the same applies to economic development. In other words, riches from natural resources do not affect either social development (HDI) or economic development (GDP per capita) in Latin America (1991-2012). In light of my findings, I argue that production of different types of resources do not affect social development differently.

The results in the original resource rents model, using HDI as dependent variable, resulted in a slight positive effect of governance on social development, however not significant. The interaction term created to capture the conditioned effect of natural resource rents on social development, depending on the score of governance, also indicated no relationship. The simple answer to the second research question is therefore "no". However, the quality of governance does affect economic development (GDP per capita) directly and by interaction. This implies that the conclusion of Mehlum et al. (2006), that the effect of natural resource abundance on economic development is condition by high institutional quality, is confirmed in my results, but proven to be less relevant in light of social development.

[106] Mehlum et al (2006) study 87 countries from 1965-1990 and Sachs and Warner (1999) study 69 countries from 1970-1990.
[107] See section 3.6 and 3.7

In short, the relationship between natural resources, governance and social development in Latin America, is characterized by no relationship. For economic development on the other hand, governance quality matter in a greater extent.

5. Concluding remarks

The introduction of this thesis begins with the highly debated question; Are riches in natural resources a "curse" for a country´s development? My aim was to compliment the resource literature, by investigating the relationship between riches from natural resource and social development in Latin America, with a specific focus on how quality of governance influences this relationship. I was also curious to see if the production of different types of natural resources could affect social development differently.

With basis in existing literature, I developed the research design and framework for empirical analysis focusing on two main models I refer to as a) the resource rents model and b) the production data model. The former in order to investigate the relationship between riches in natural resources (Natural resource rents as % of GDP), and how governance influences this relationship. The latter consists of different types of natural resource production data and developed in order to investigate if those have different effect on social development.

My findings and evaluation of robustness conclude that there is little evidence to confirm that riches in natural resources have a positive effect on social development (HDI) in Latin American countries. Nor do my results point in the opposite direction. In continuation, the quality of governance does not seem to influence this relationship and lastly, the production of different types of natural resources does not seem to affect social development differently in Latin America. In short, my results are null findings.

In evaluating my findings, another interesting result presented itself: governance has a strong positive effect on economic development (GDP) in Latin America. In light of the almost non-existing estimated effect of governance on social development, I found this a bit surprising. It nevertheless tells us that the difference between social development and economic development is not at all trivial, and that my choice of dependent variable indeed is a contrast and complimentary to the resource literature. Why governance effects economic development and not social development in Latin America, is a question that I leave hanging in the air, hoping that someone will find it interesting enough to pursue further.

My last comment is simple. This thesis provides no evidence
of a resource curse in Latin America.

Bibliography

Acemoglu, D., S. Johnson, J.A Robinson, and P. Yared. 2008. “Income and democracy.” *American Economic Review* 98(3): 808–842.
Acemoglu, D., J.A Robinson, and S. Johnson. 2001. “The colonial origins of comparative development: an empirical investigation.” *American Economic Review* 91: 1369–1401.
Adcock, Robert, and David Collier. 2001. “Measurement validity: A shared standard for qualitative and quantitative research.” In *American Political Science Association*, Cambridge Univ Press, p. 529–546. http://journals.cambridge.org/abstract_S0003055401003100 (Accessed December 10, 2015).
Agresti, Alan, and Barbara Finlay. 2008. *Statistical Methods for the Social Sciences*. 4 edition. Upper Saddle River, N.J: Prentice Hall.
Alexeev, Michael, and Robert Conrad. 2009. “The Elusive Curse of Oil.” *The Review of Economics and Statistics* 91(3): 586–598.
Allison, Paul D. 2015. “Imputation by Predictive Mean Matching: Promise & Peril.” *Statistical Horizons*. https://statisticalhorizons.com/predictive-mean-matching (Accessed November 11, 2017).
Allison, Paul D. 2001. *Missing Data*. SAGE Publications.
Altman, Douglas G., and Martin J Bland. 1995. “Statistics Notes: Absence of evidence is not evidence of absence.” *BMJ* 311.
Ardanaz, Martín, Carlos Scartascini, and Mariano Tommasi. 2011. “Political Institutions, Policymaking, and Economic Policy in Latin America.” *The Oxford Handbook of Latin American Economies*. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199571048.001.0001/oxfordhb-9780199571048-e-3 (Accessed February 19, 2016).
Aslaksen, Silje. 2010. “Oil and democracy: More than a cross-country correlation?” *Journal of Peace Research* 47(4): 421–431.
Auty, Richard M. 1990. *Resource-Based Industrialization: Sowing the Oil in Eight Developing Countries*. New York: Oxford University Press.
Auty, Richard M. 2001. “The political economy of resource-driven growth.” *European Economic Review* 45(4–6): 839–846.
Baer, Werner. 1972. “Import Substitution and Industrialization in Latin America: Experiences and Interpretations.” *Latin American Research Review* 7(1): 95–122.
Beck, N., and J. N. Katz. 1996. “Nuisance vs. Substance: Specifying and Estimating Time-Series-Cross-Section Models.” *Political Analysis* 6(1): 1–36.
Beck, Nathaniel, and Jonathan N. Katz. 1995. “What to do (and Not to Do) with Time-Series Cross-Section Data.” *American Political Science Review* 89(03): 634–647.
Bevir, Mark. 2008. *Key Concepts in Governance*. SAGE.
Blomström, Magnus, and Patricioeditores Magnus Blomström Meller. 1991. *Diverging paths: comparing a century of Scandinavian and Latin American economic development*. Inter-American Development Bank.
Boix, Carles, Michael Miller, and Sebastian Rosato. 2012. “A complete data set of political regimes, 1800–2007.” *Comparative Political Studies*: 0010414012463905.
Brahmbhatt, Milan, Otaviano Canuto, and Ekaterina Vostroknutova. 2010. 7 *Dealing with Dutch Disease*. The World Bank. http://siteresources.worldbank.org/INTPREMNET/Resources/EP16.pdf (Accessed January 17, 2016).

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14(1): 63–82.
Brunnschweiler, Christa N., and Erwin H. Bulte. 2008. "The resource curse revisited and revised: A tale of paradoxes and red herrings." *Journal of Environmental Economics and Management* 55(3): 248–264.
Bull, Benedicte. 2016. "Governance in the Aftermath of NeoLiberalism: Aid, Elites and State Capacity in Central America." *Forum for Development Studies*: 1–23.
Bull, Benedicte. 2015. *Latinamerikanske utfordringer*. 2. utgave. Cappelen Damm.
Cameron, Adrian Colin, and Pravin K. Trivedi. 2010. 2 *Microeconometrics using stata*. Stata press College Station, TX.
Cecchini, Simone, Fernando Filgueira, and Claudia Robles. 2014. 49 *Social Protection systems in Latin America and the Caribbean, a comparative view*. Santiago: Economic Commission for Latin America and the Caribbean.
Chen, Xiao, Philip B. Ender, Michael Mitchell, and Christine Wells. 2003. *Regression with Stata*. https://stats.idre.ucla.edu/stat/stata/webbooks/reg/default.htm . (Accessed November 23, 2017).
Christophersen, Knut Andreas. 2013. *Introduksjon til statistisk analyse, Regresjonsbaserte metoder og anvendelser*. 1.utgave. Oslo: Gyldendal.
Corden, W. Max, and J. Peter Neary. 1982. "Booming Sector and De-Industrialisation in a Small Open Economy." *The Economic Journal* 92(368): 825.
Dahl, Robert A. 1991. *Democracy and Its Critics*. New edition edition. Yale University Press.
ECLAC. 2014a. *Compacts for Equality, Towards a Sustainable Future*. Lima, Peru: Economic Commission for Latin America and the Caribbean.
ECLAC. 2014b. *Social Panorama of Latin America 2014*. Santiago, Chile: Economic Commission for Latin America and the Caribbean. Text. http://www.cepal.org/en/publications/37626-social-panorama-latin-america-2014 (Accessed February 5, 2016).
ECLAC. 2017. 272 *Social Panorama of Latin America 2016*. Santiago, Chile: Economic Commission for Latin America and the Caribbean.
ECLAC. 2010. *Time for Equality, Closing gaps, opening trails*. Brasilia, Brazil: Economic Commission for Latin America and the Caribbean.
ECLAC/UNASUR. 2013. 108 *Natural resources within the Union of South American Nations, Status and trends for a region development agenda*. Santiago: Economic Commission for Latin America and the Caribbean/Unión de Naciones Suramericanas.
Franko, Patrice. 2007. *The Puzzle of Latin American Economic Development*. Third Edition. USA: Rowman & Littlefield Publishers, Inc.
Fukuyama, Francis. 2013. "What Is Governance?" *Governance: An international Journal of Policy, Administration and Institutions* 26(3): 347–368.
Gasparini, Leonardo, and Nora Lustig. 2011. "The Rise and Fall of Income Inequality in Latin America" eds. Jose Antonio Ocampo and Ros Jaime. *The Oxford Handbook of Latin American Economics*. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199571048.001.0001/oxfordhb-9780199571048-e-27 (Accessed March 9, 2016).
Gelb et al. 1988. *Oil Windfalls: Blessing or Curse?* Washington, DC: World Bank.
Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. 1 edition. Cambridge ; New York: Cambridge University Press.

Grynspan, Rebeca, and Luis F. López-Calva. 2011. “Multidimensional Poverty in Latin America: Concept, Measurement, and Policy.” *The Oxford Handbook of Latin American Economies*. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199571048.001.0001/oxfordhb-9780199571048-e-28 (Accessed February 19, 2016).
Haber, Stephen, and Victor Menaldo. 2011. “Do natural resources fuel authoritarianism? A reappraisal of the resource curse.” *American political science Review* 105(1): 1–26.
Hanson, Jonathan, and Rachel Sigman. 2013. *Leviathan’s Latent Dimensions: Measuring State Capacity for Comparative Political Research*. Rochester, NY: Social Science Research Network. SSRN Scholarly Paper. https://papers.ssrn.com/abstract=1899933 (Accessed December 16, 2017).
Honaker, James, and Gary King. 2010. “What to Do about Missing Values in Time-Series Cross-Section Data.” *American Journal of Political Science* 54(2): 561–581.
Honaker, James, Gary King, and Matthew Blackwell. 2011. “Amelia II: A program for missing data.” *Journal of statistical software* 45(7): 1–47.
Hosmer, David W., and Stanley Lemenshow. 2000. *Applied Logistic Regression*. Second. New York: John Wiley & Sons.
Høyland, Bjørn, and H. M. Nygaard. 2011. “Missing uncertainty and uncertain missing in the study of civil war onset.” *Oslo: Department of Political Science, University of Oslo, Working Paper*. http://www.uio.no/english/research/interfaculty-research-areas/democracy/news-and-events/events/conferences/2012/papers-2012/Hoyland-Nygaard-wshop7.pdf (Accessed May 6, 2016).
Jaccard, James, and Robert Turrisi. 2003. *Interaction Effects in Multiple Regression*. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc. http://methods.sagepub.com/book/interaction-effects-in-multiple-regression (Accessed December 10, 2017).
Johnson, Chalmers. 1982. *Miti and the Japanese Miracle: The Growth of Industrial Policy : 1925-1975*. Stanford University Press.
Johnson, Chalmers. 1999. “The Developmental State: Adyssey of a Concept.” In ed. Meredith Woo-Cumings. Cornell, CA: Cornell University Press.
Kam, Cindy D., and Robert J. Franzese. 2007. *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor, Michigan: University of Michigan Press. https://www.press.umich.edu/206871/modeling_and_interpreting_interactive_hypotheses_in_regression_analysis (Accessed December 10, 2017).
Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2011. “The Worldwide Governance Indicators: Methodology and Analytical Issues.” *Hague Journal on the Rule of Law* 3(02): 220–246.
Kennedy, Peter. 2003. *A Guide to Econometrics*. 6th Edition. Wiley. http://eu.wiley.com/WileyCDA/WileyTitle/productCd-EHEP001025.html (Accessed March 29, 2016).
King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, N.J: Princeton University Press.
Knack, S., and P. Keefer. 1995. “Institutions and economic performance: cross- country tests using alternative institutional measures.” *Economics and Politics* 7: 207–227.
Krugman, Paul. 1987. “The narrow moving band, the Dutch disease, and the competitive consequences of Mrs. Thatcher.” *Journal of Development Economics* 27(1–2): 41–55.
Lederman, Daniel, and William Francis Maloney. 2007. *Natural Resources, Neither Curse Nor Destiny*. World Bank Publications.

Lee, C.A., and R.C. Hill. 2008. *Using Stata for Principles of Econometrics*. Third Edition. Hoboken, New Jersey: John Wiley & Sons.
Leftwich, Adrian. 1993. "Governance, democracy and development in the Third World." *Third World Quarterly* 14(3): 605–624.
Little, Roderick J. A., and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data*. 2 edition. Hoboken, N.J: Wiley-Interscience.
Long, J. Scott, and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata, Second Edition*. Stata Press.
Maloney, William F. 2002. "Missed Opportunities: Innovation and Resource-Based Growth in Latin America." *Economía* 3(1): 111–167.
McNeish, John-Andrew, and Axel Borchgrevink. 2015. "Chapter 1, Introduction: Recovering power from energy - Reconsidering the linkages between energy and development." In *Contested Powers: The Politics of Energy and Development in Latin America*, eds. John-Andrew McNeish, Axel Borchgrevnik, and Owen Logan. Zed Books, Limited.
Mehlum, Halvor, Karl Moene, and Ragnar Torvik. 2006. "Institutions and the resource curse." *The Economic Journal* 116(508): 1–20.
Menard, Scott. 2010. *Logistic Regression: From Introductory to Advanced Concepts and Applications*. SAGE.
Meng, Xiao-Li. 1994. "Multiple-Imputation Inferences with Uncongenial Sources of Input." *Statistical Science* 9(4): 538–558.
Midtbø, Tor. 2012. *Stata: En entusiastisk innføring*. Oslo: Universitetsforlaget. http://www.universitetsforlaget.no/nettbutikk/stata-uf.html (Accessed November 8, 2017).
Mikesell, Raymon F. 1997. "Explaining the resource curse, with special reference to mineral-exporting countries." *Resource Policy* 23(4): 191–199.
Morris, Tim P., Ian R. White, and Patrick Royston. 2014. "Tuning multiple imputation by predictive mean matching and local residual draws." *BMC Medical Research Methodology* 14: 75.
Nguyen, Cattram D., John B. Carlin, and Katherine J. Lee. 2017. "Model checking in multiple imputation: an overview and case study." *Emerging Themes in Epidemiology* 14. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5569512/ (Accessed November 11, 2017).
North, D.C. 1963. "Aspects of economic growth in the United States, 1815-1860." In *Experience of Economic Growth, ed. B.E. Supple*, New York: Random House.
Pedersen, Jørn. 2013. "Democracy and Famine Revisited : Investigating the effects of contestation, participation and civil liberties." https://www.duo.uio.no/handle/10852/36965 (Accessed March 21, 2016).
Pennings, Paul, Hans Keman, and Jan Kleinnijenhuis. 2017. "Multivariate Analysis and Causal Inference." In *Doing Research in Political Science*, London: SAGE Publications Ltd, p. 132–182. http://methods.sagepub.com/book/doing-research-in-political-science/n8.xml (Accessed November 26, 2017).
Peugh, James L., and Craig K. Enders. 2004. "Missing data in educational research: A review of reporting practices and suggestions for improvement." *Review of educational research* 74(4): 525–556.
Pierre, Jon, and B. Guy Peters. 2000. *Governance, Politics, and the State*. St. Martin's Press.
van der Ploeg, Frederick. 2011. "Natural Resources: Curse or Blessing?" *Journal of Economic Literature* 49(2): 366–420.
Rhodes, R. A. W. 1996. "The New Governance: Governing without Government." *Political Studies* 44(4): 652–667.
Ross, Michael L. 2001. "Does oil hinder democracy?" *World Politics* 53(3): 325–361.

Rothstein, Bo, and Jan Teorell. 2008. “What Is Quality of Government? A Theory of Impartial Government Institutions.” *Governance* 21(2): 165–190.
Sachs, Jeffrey D., and Andrew M. Warner. 1995. *Natural Resource Abundance and Economic Growth*. National Bureau of Economic Research. http://www.nber.org/papers/w5398 (Accessed September 22, 2015).
Sachs, Jeffrey D., and Andrew M. Warner. 1999. “The big push, natural resource booms and growth.” *Journal of Development Economics* 59: 43–76.
Sachs, Jeffrey D., and Andrew M. Warner. 2001. “The curse of natural resources.” *European Economic Review* (45): 827–838.
Schafer, Joseph L., and John W. Graham. 2002. “Missing data: Our view of the state of the art.” *Psychological Methods* 7(2): 147–177.
Sen, Amartya. 2001. *Development as Freedom*. OUP Oxford.
Sen, Amartya. 1983. “Development: Which Way Now?” *The Economic Journal* 93(372): 745.
Sinnott, Emily, John Nash, and Augusto de la Torre. 2010. 100 *Natural resources in Latin America and the Caribbean, beyond booms and busts?* Washington, DC: The World Bank.
Skog, Ole-Jørgen. 2010. *Å forklare sosiale fenomener*. Second editon. Gyldendal Norsk Forlag. http://www.gyldendal.no/Faglitteratur/Samfunnsfag/Statistikk-metode/AA-forklare-sosiale-fenomener (Accessed May 6, 2016).
StataCorp, L. P. 2013. “STATA Multiple-imputation, Reference manual.” *Release 13*.
Stewart, Frances. 1985. *Planning to Meet Basic Needs*. London: Palgrave Macmillan.
Stijns, Jean-Philippe C. 2005. “Natural resource abundance and economic growth revisited.” *Resources Policy* 30(2): 107–130.
Stock, James H. W., and Mark W. Watson. 2012. *Introduction to Econometrics*. 3 edition. Pearson.
Stock, James H., and Mark W. Watson. 2015. *Introduction to Econometrics*. Updated third edition. Pearson Education.
Strand, Havard, Havard Hegre, Scott Gates, and Marianne Dahl. 2012. “Why Waves? Global Patterns of Democratization, 1820–2008.” *Typescript, PRIO*. http://www.cidcm.umd.edu/workshop/papers/Dahl_CIDCM2013.pdf (Accessed May 5, 2016).
Sunkel, Osvaldo. 1969. “National development policy and external dependence in Latin America.” *The Journal of Development Studies* 6(1): 23–48.
Teorell, Jan et al. 2016. “QoG Standard Dataset 2016.” http://dx.doi.org/10.18157/QoGStdJan16 (Accessed March 24, 2016).
Teorell, Jan, Stefan Dahlberg, Sören Holmberg, Bo Rothstein, and Richard Svensson. 2016. *The Quality of Government Standard Dataset*. The Quality of Government Institute: University og Gothenburg. http://www.qog.pol.gu.se doi:10.18157/QoGStdJan16.
“The Brazilian Oil & Gas Industry.” 2013. *PricewaterhouseCoopers*. https://www.pwc.com.br/pt/publicacoes/setores-atividade/assets/oil-gas/oeg-tsp-13.pdf.
The World Bank. 2016. “World Databank.” http://data.worldbank.org/ (Accessed January 21, 2016).
UNDP. 1990. 141 *Human Development Report 1990*. Oxford University Press, New York: United Nations Development Programme. http://hdr.undp.org/sites/default/files/reports/219/hdr_1990_en_complete_nostats.pdf (Accessed February 18, 2016).

UNDP. 2010. 238 *The Real Wealth of Nations: Pathways to Human Development*. New York: United Nations Development Programme. Human Development Report 2010. http://hdr.undp.org/sites/default/files/reports/270/hdr_2010_en_complete_reprint.pdf (Accessed February 18, 2016).

UNDP. 2015. *Work for Human Development*. New York: United Nations Development Programme. Human Development Report. http://hdr.undp.org/sites/default/files/2015_human_development_report_1.pdf.

UNESCO. 2014. *Regional Report about Education for All in Latin America and the Caribbean*. Regional Bureau of Education for Latin American and the Caribeean: UNESCO. Global Education for All Meeting, Muskat Oman. http://www.unesco.org/fileadmin/MULTIMEDIA/HQ/ED/ED_new/pdf/LAC-GEM-2014-ENG.pdf.

"Uneven Recovery in Latin America: A Regional Overview." 2015. *HKTDC Research*. http://economists-pick-research.hktdc.com/business-news/article/Research-Articles/Uneven-Recovery-in-Latin-America-A-Regional-Overview/rp/en/1/1X000000/1X0A2FSZ.htm (Accessed September 27, 2015).

Vink, Gerko, Laurence E. Frank, Jeroen Pannekoek, and Stef Buuren. 2014. "Predictive mean matching imputation of semicontinuous variables." *Statistica Neerlandica* 68(1): 61–90.

Watson, Mark W., and James H. Stock. 2007. "Chapter 10: Regression with Panel Data." In *Introduction to Econometrics,* Pearson Education, p. 349–382.

White, Ian R., Patrick Royston, and Angela M. Wood. 2011. "Multiple imputation using chained equations: issues and guidance for practice." *Statistics in medicine* 30(4): 377–399.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Wooldridge, Jeffrey M. 2009. *Introductory Econometrics: A Modern Approach, 4th Edition*. 4th edition. Mason, OH: South-Western.

Wooldridge, Michael. 2009. *An Introduction to MultiAgent Systems*. John Wiley & Sons.

Wright, Gavin, and Jesse Czelusta. 2004. "Why Economies Slow, The Myth of the Resource Curse." *Challange* 47(2): 6–38.

Appendices

Table 1a: List of country-year observations in the dataset.

Country	**Time**
Argentina	1991-2012
Bolivia	1991-2012
Brazil	1991-2012
Belize	1991-2012
Chile	1991-2012
Colombia	1991-2012
Costa Rica	1991-2012
Ecuador	1991-2012
El Salvador	1991-2012
Guatemala	1991-2012
Honduras	1991-2012
Mexico	1991-2012
Panama	1991-2012
Paraguay	1991-2012
Peru	1991-2012
Uruguay	1991-2012
Venezuela	1991-2012

Appendix 1: Multiple imputation with diagnostics

Table 6.1: Preliminary regression using listwise deletion, production data model

Fixed-effects (within) regression — Number of obs = 132
Group variable: ccode — Number groups = 13

R-sq: — Obs per group:

within = 0.9685 — min = 2
between = 0.0197 — avg = 10.2
overall = 0.1503 — max = 22

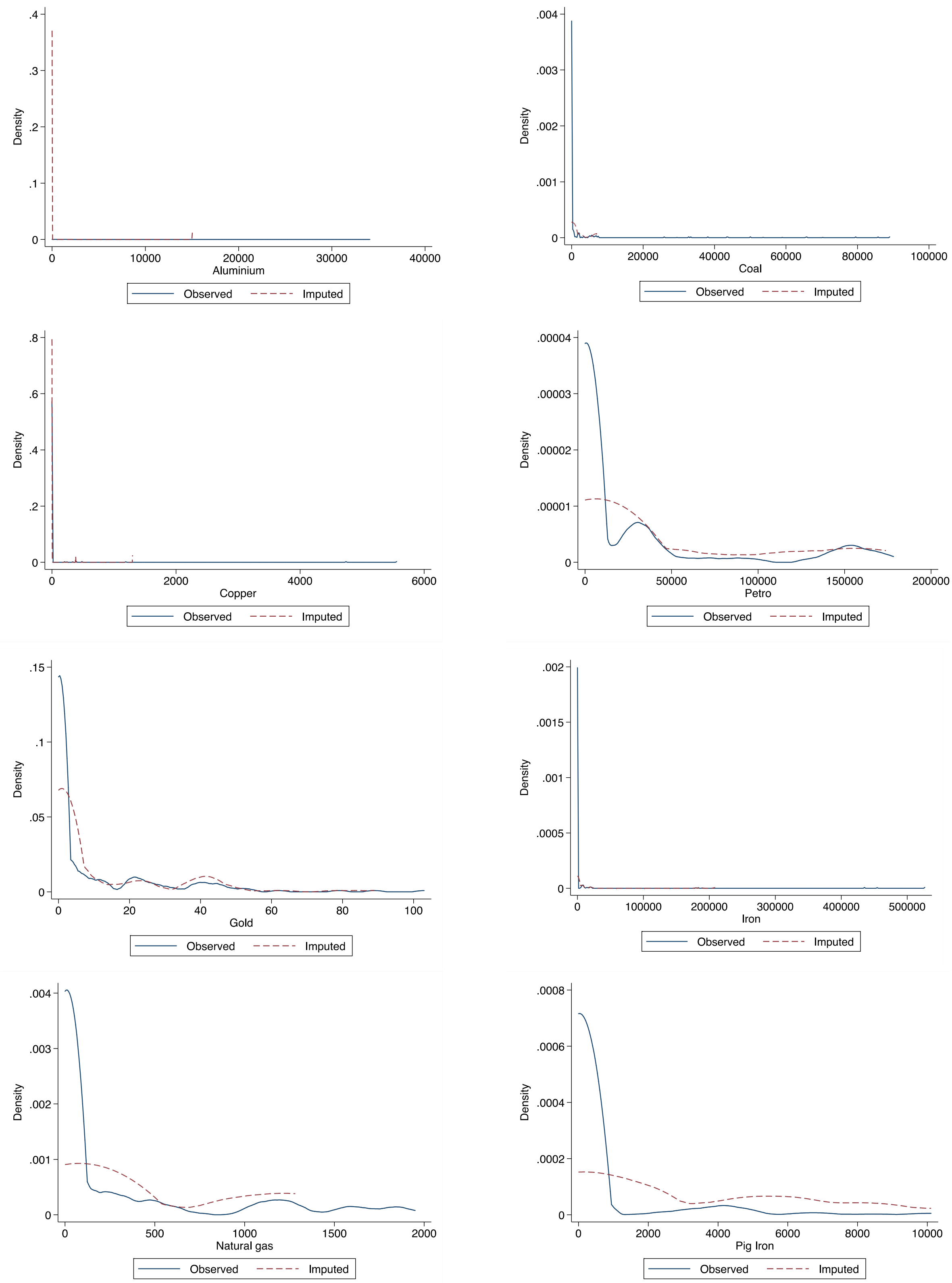
F(7,112) = 492.63
corr(u_i, Xb) = -0.0944 — Prob > F = 0.0000

HDI	Coef.	Std. Err.	t	P>t	[95% Conf.Interval]	
Petroleum	-1.29e-07	1.86e-07	-0.70	0.487	-4.97e-07	2.39e-07
Natural Gas	3.56e-06	.0000104	0.34	0.734	-.0000171	.0000242
Gold & silver	-.0000134	6.83e-06	-1.96	0.052	-.0000269	1.43e-07
Minerals	8.59e-07	5.19e-07	1.65	0.101	-1.70e-07	1.89e-06
Governance	.046486	.0062193	7.47	0.000	.0341631	.0588088
GDP	-1.36e-06	5.09e-07	-2.67	0.009	-2.37e-06	-3.52e-07
time	.0056003	.0001502	37.27	0.000	.0053026	.005898
_cons	.6112644	.0055991	109.17	0.000	.6001705	.6223584

Table 6.2: Variables in the imputation model

To be imputed	Part of analysis	Additional
Aluminum	GDP	GNI
Copper	Population	Access Sanit
Coal	Governance	Access water
Petroleum	HDI index	Life Expectancy
Gold	Time	Unemployment
Iron		Taxes
Natural gas		World Prices Oil
Pig iron		World Prices Min.
Silver		World Prices Gas
Tin		
Wire rod		
Zink		
Democracy		

Figure 6.1. Kernel density plot of the observed and imputed values for one imputed dataset.



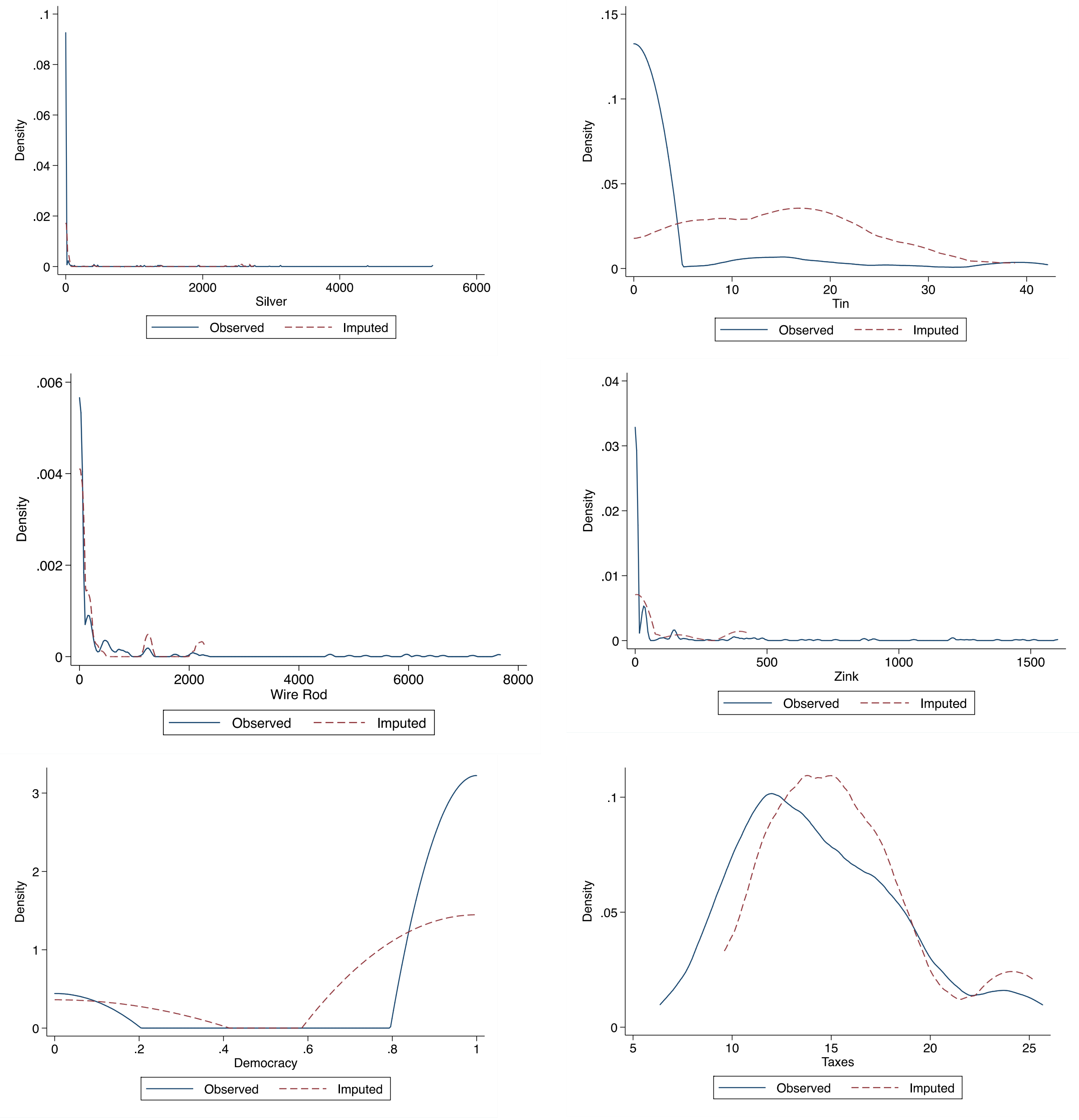
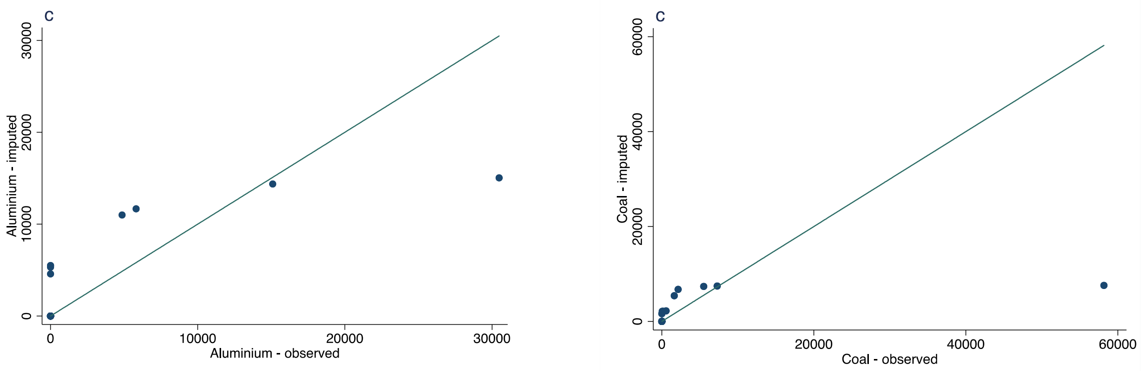
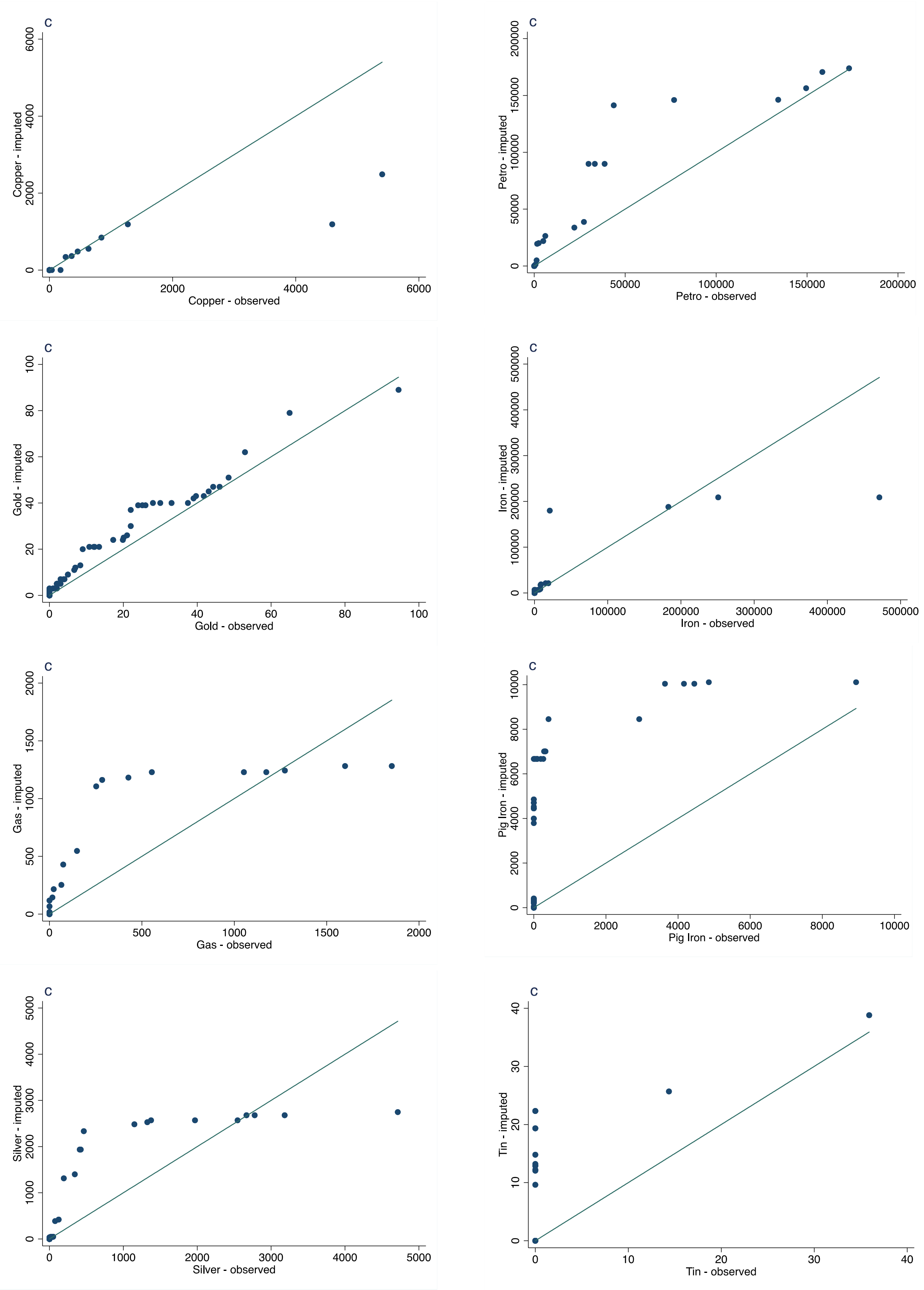
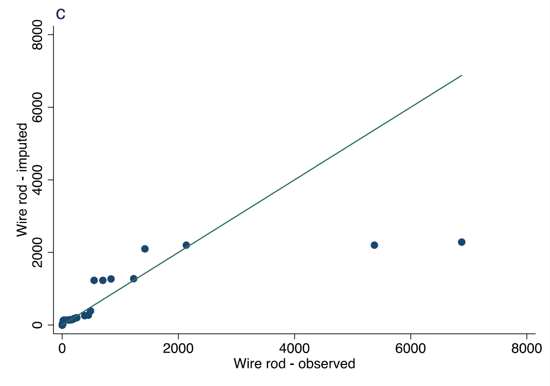
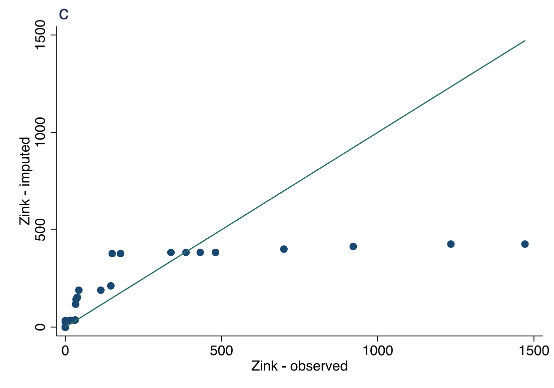


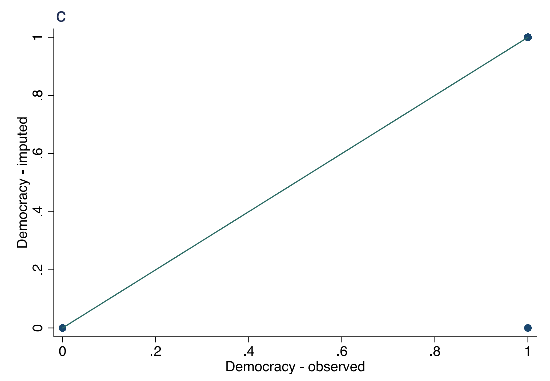
Figure 6.2 Plot of the quantiles of the imputed data against quantiles of the observed data (quantile-quantile plot)

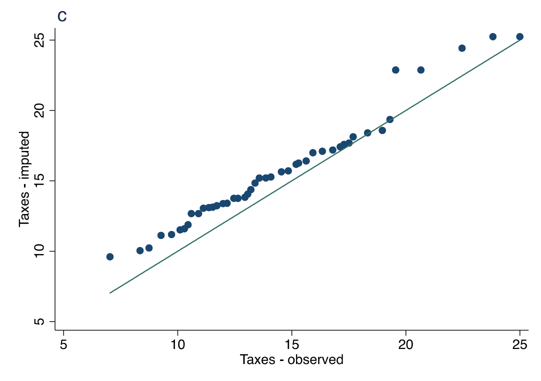



c
Copper - imputed
Copper - observed
c
Petro - imputed
Petro - observed
c
Gold - imputed
Gold - observed
c
Iron - imputed
Iron - observed
c
Gas - imputed
Gas - observed
c
Pig Iron - imputed
Pig Iron - observed
c
Silver - imputed
Silver - observed
c
Tin - imputed
Tin - observed


C
Wire rod - imputed
Wire rod - observed
0
2000
4000
6000
8000


C
Zink - imputed
Zink - observed
0
500
1000
1500


C
Democracy - imputed
Democracy - observed
0
.2
.4
.6
.8
1


C
Taxes - imputed
Taxes - observed
5
10
15
20
25

Appendix 2: Pre estimation diagnostics

Table 6.3. Hausman test comparing FE vs RE Resource Rents model

	Coefficients			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	Fixed Effects	Random Effects	Difference	S.E.
Governance	.0033705	.0019398	.0014307	.
Natural resource rents (% GDP)	-.0002614	-.0002546	-6.85e-06	.
Democracy	-.0062819	-.0062392	-.0000427	.
Population (log)	-.00963	-.0083818	-.0012482	.0118648
GDP (log)	.0125468	.0189365	-.0063897	.0017331
Time	.0050821	.0049345	.0001476	.0002194

b = consistent under Ho and Ha; obtained from xtreg

B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(6) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 127.77

Prob>chi2 = 0.0000

(V_b-V_B is not positive definite)

Table 6.4. Hausman test comparing FE vs RE Production model

	Coefficients			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	Fixed Effects	Random Effects	Difference	S.E.
Petroleum	.000373	-.0000758	-.0002973	.
Natural gas	.0006552	.0010721	-.0004169	.
Gold_Silver	.0001185	.0000384	-.0001569	.
Minerals	.000493	.0004478	.0000452	.
Governance	.0139258	.010629	.0032968	.
GDP (log)	.0077216	.0298632	-.0221417	.
Time	.0053804	.0046932	.0006872	.0001865

b = consistent under Ho and Ha; obtained from xtreg

B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(8) (b-B)'[(V_b-V_B)^(-1)](b-B)

= 5659.98

Prob>chi2 = 0.0000

(V_b-V_B is not positive definite)

Figure 6.3: DFBeta Plot of influential values in the natural resource rents model (years)

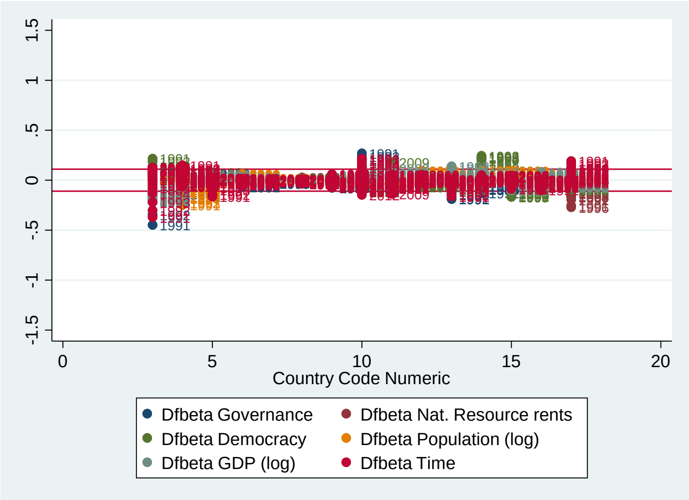


Figure 6.4. Cook's D versus studentized residuals Resource rents model

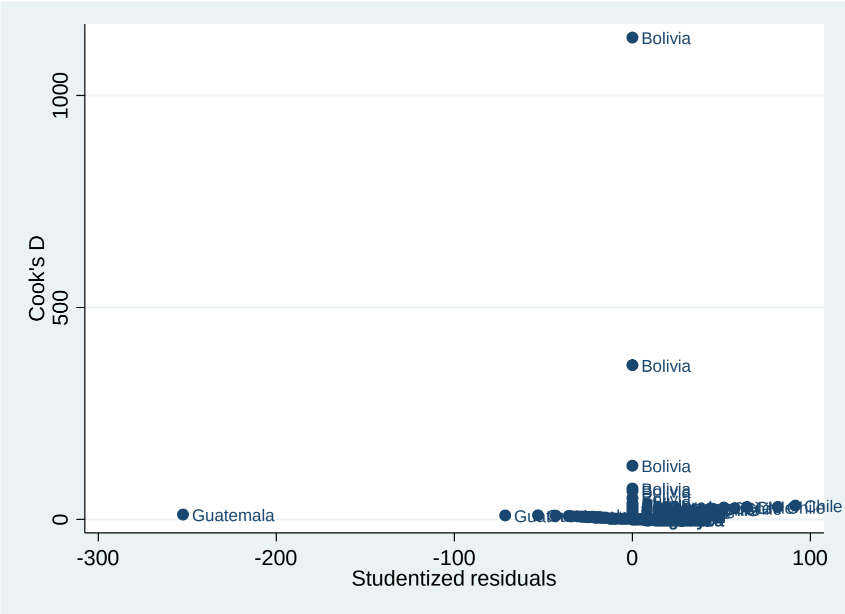


Figure 6.5: Leverage versus residual squared plots for production data model

(first row before deletion of influential imputed values and second row after deletion)

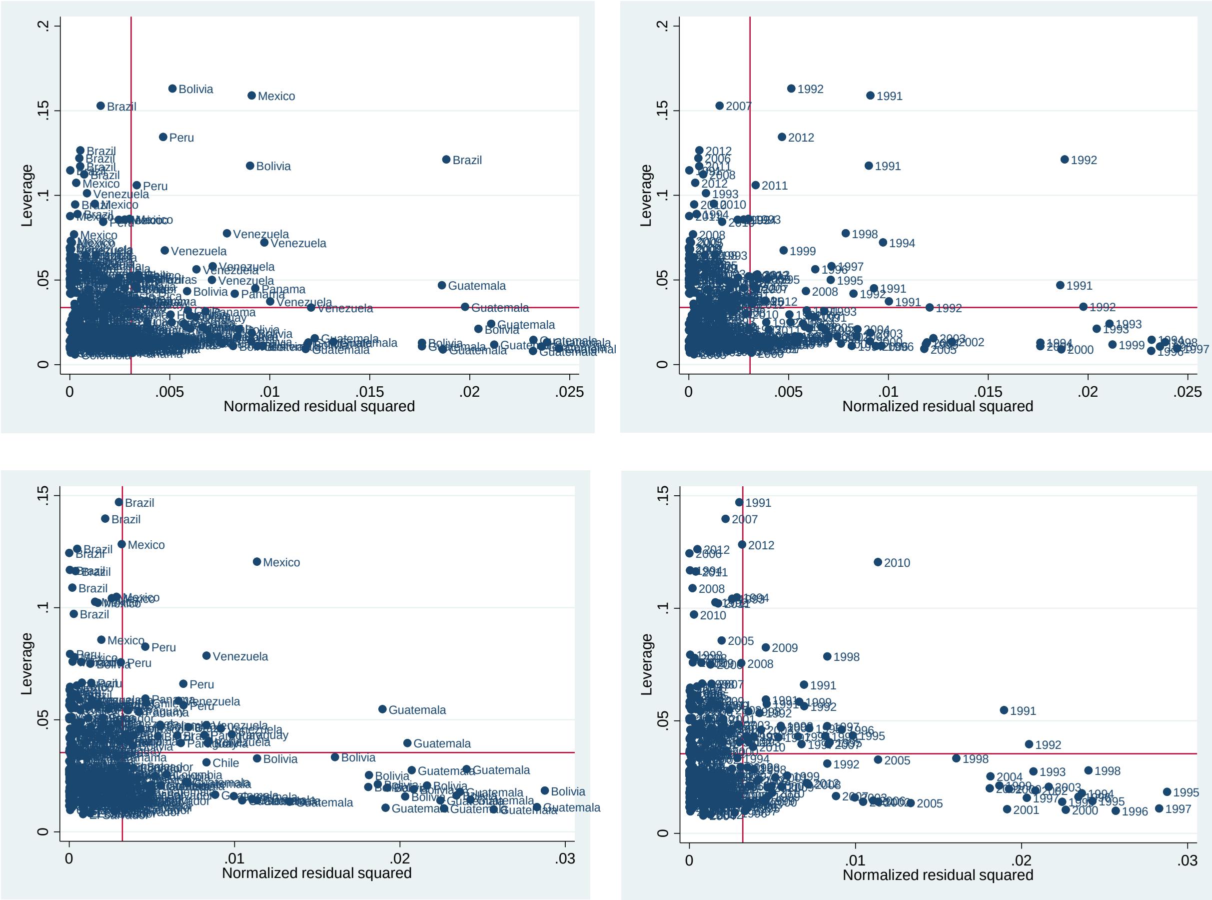
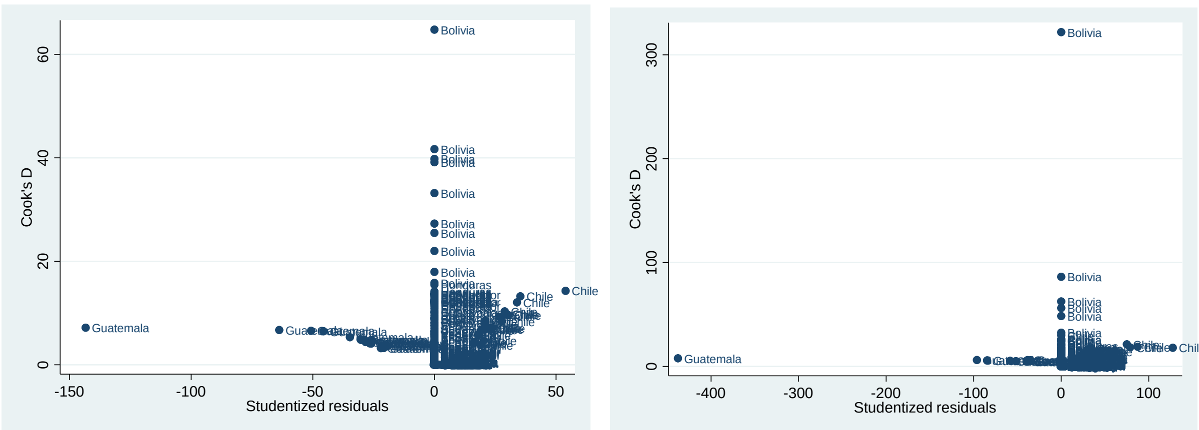
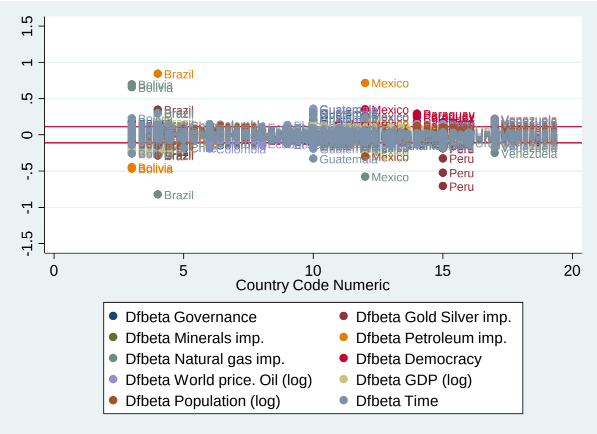


Figure 6.6. Cook's D versus studentized residuals production data model

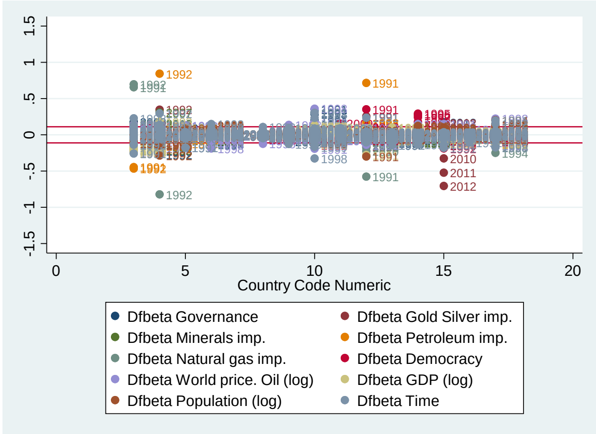


a) Production model before removing influential imp.val b)Production model after removing influential imp.val

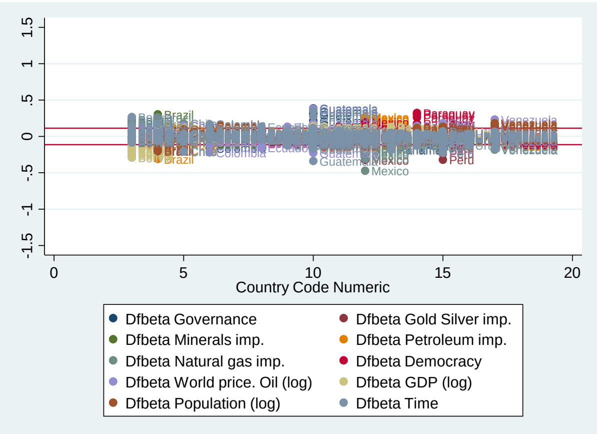
Figure 6.7: DF Beta Plot of influential values in the production data model (A and B)



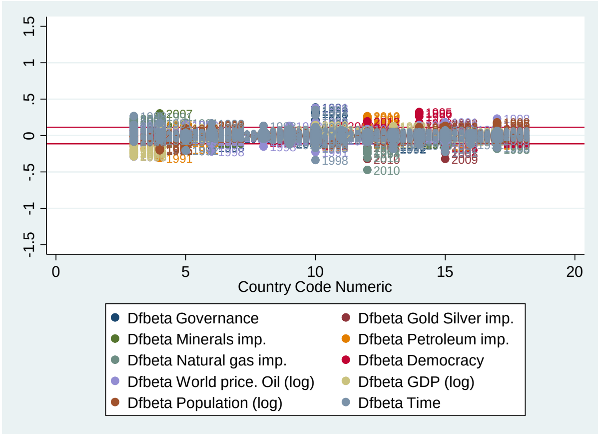
DF BETA Plot A: Model with imputed values (country)



DF BETA Plot A: Model with imputed values (year)



DF BETA Plot B: With influential imputed values deleted



DF BETA Plot B: With influential imputed values deleted (also presented in section 5.2)

Appendix 3: Regression assumptions

Figure 6.8: Linearity assumption,natural resource rents model using acrprplot

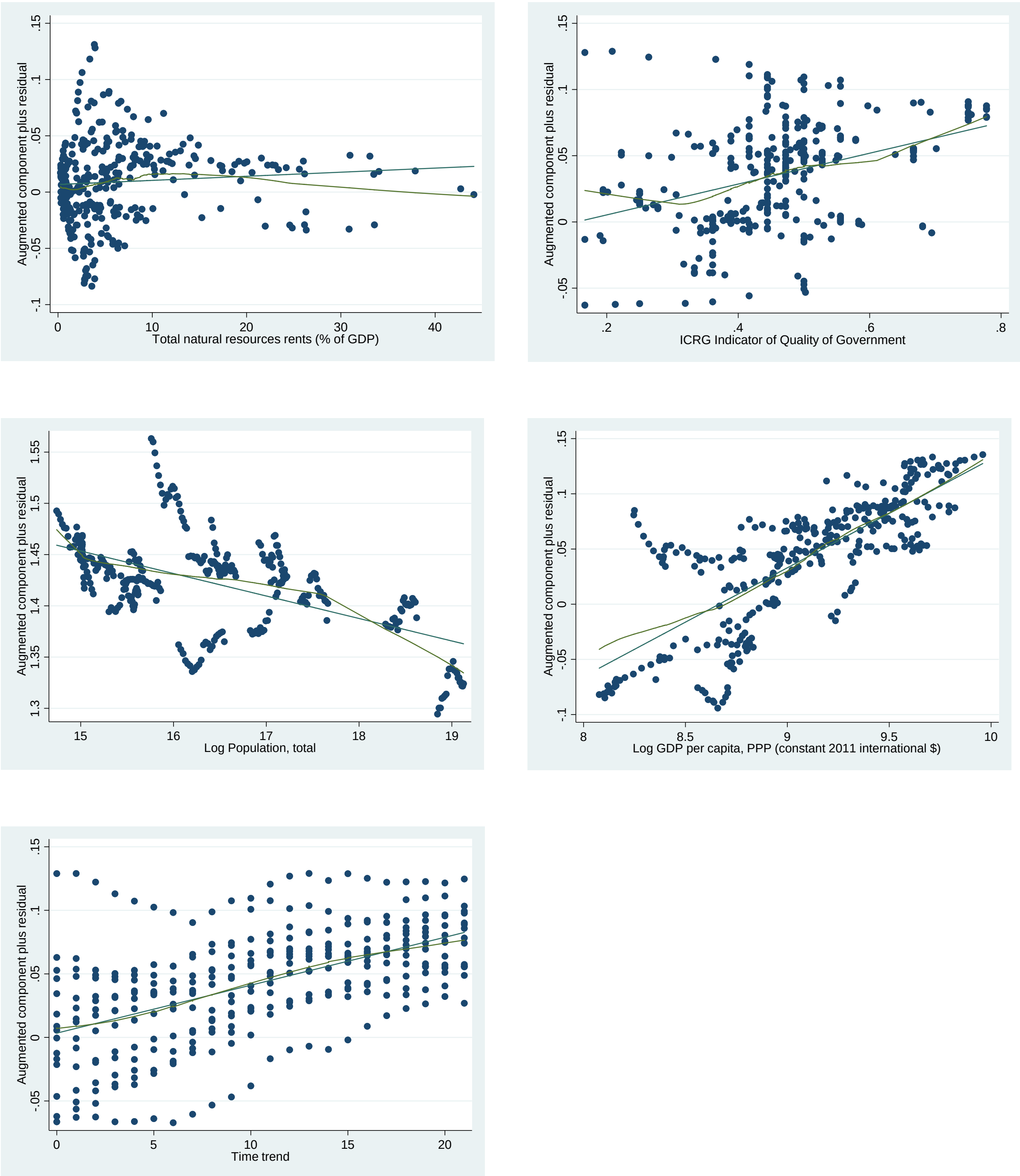
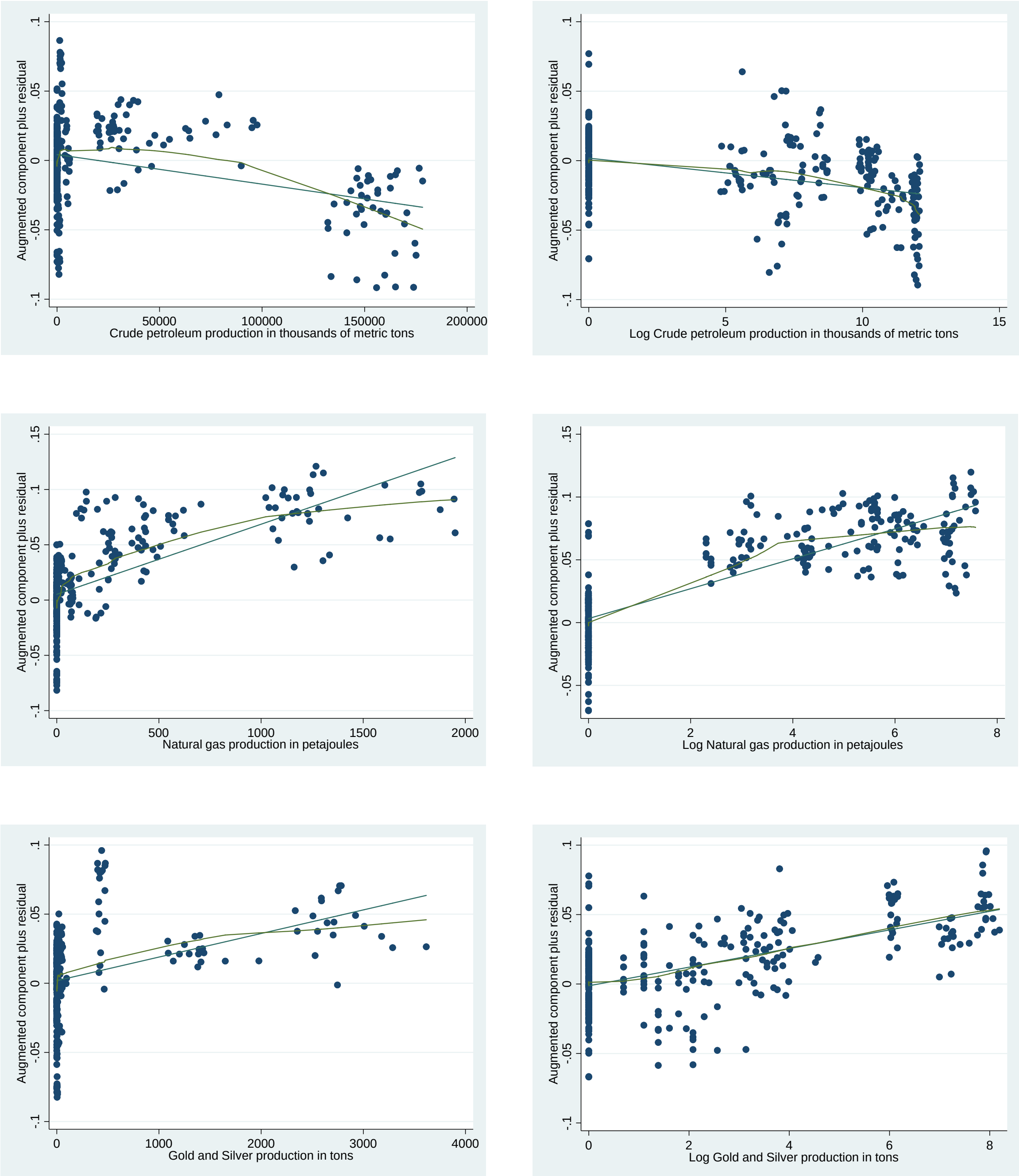
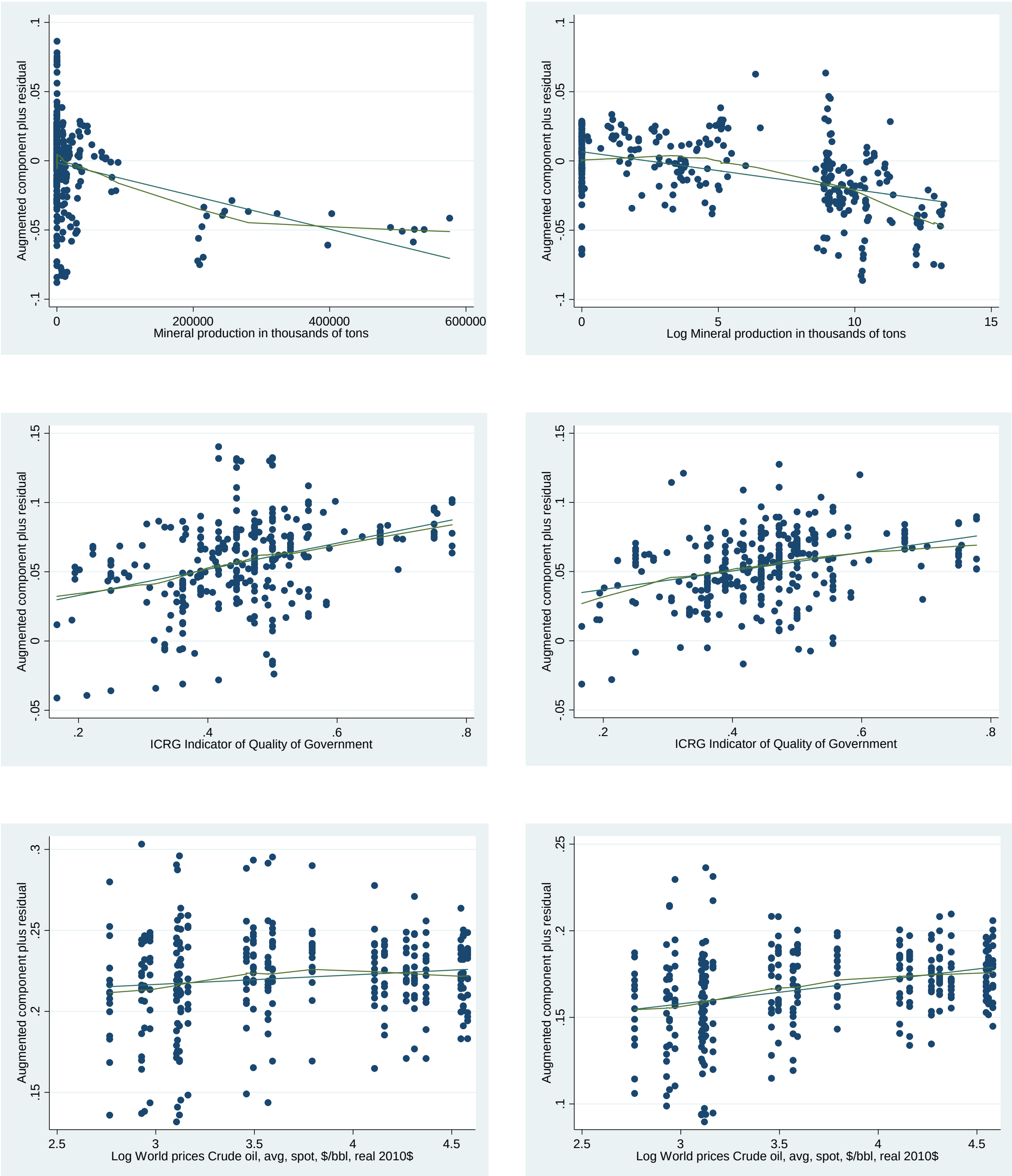
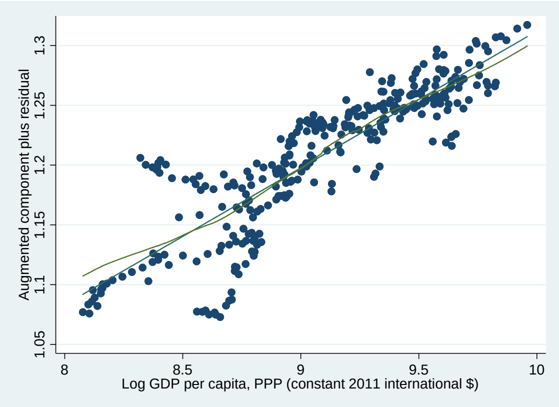
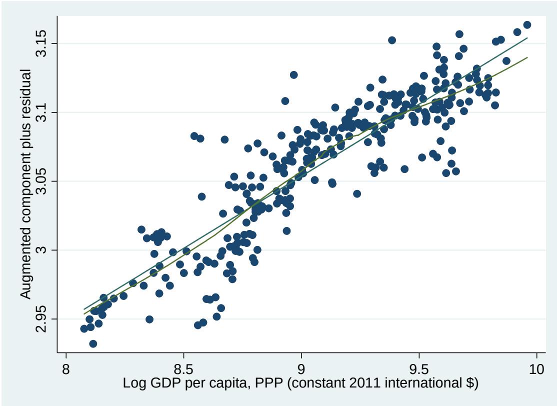


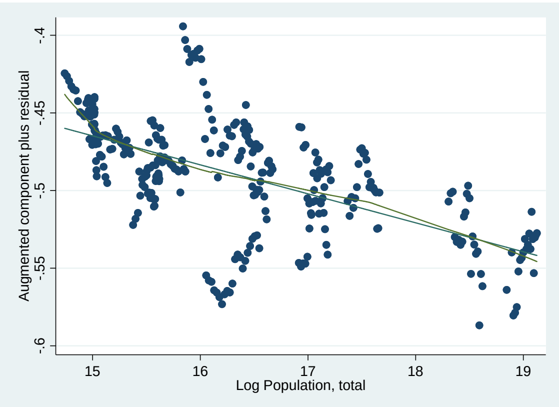
Figure 6.9: Linearity assumption, production data model using acrprplot
The right column includes log(+) transformation of production data

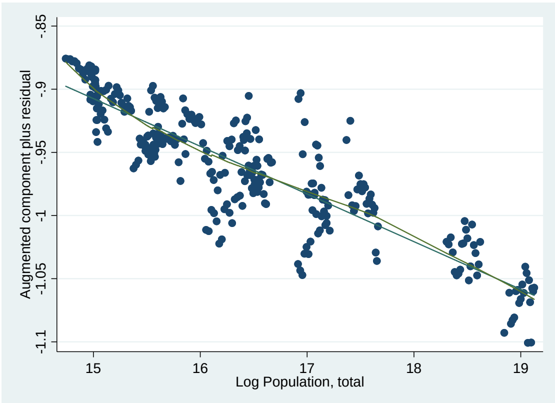


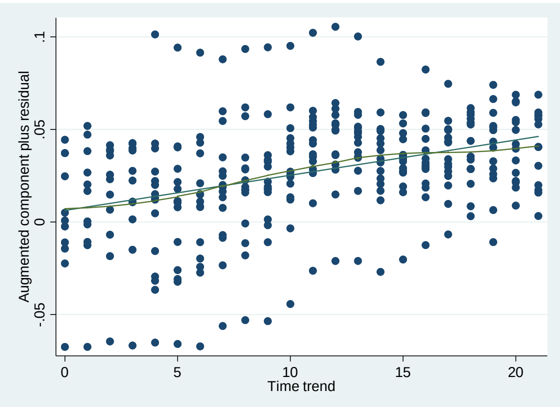

Augmented component plus residual
Mineral production in thousands of tons
Log Mineral production in thousands of tons
ICRG Indicator of Quality of Government
Log World prices Crude oil, avg, spot, $/bbl, real 2010$


Augmented component plus residual
1.05
1.1
1.15
1.2
1.25
1.3
8
8.5
9
9.5
10
Log GDP per capita, PPP (constant 2011 international $)


Augmented component plus residual
2.95
3
3.05
3.1
3.15
8
8.5
9
9.5
10
Log GDP per capita, PPP (constant 2011 international $)


Augmented component plus residual
-.6
-.55
-.5
-.45
-.4
15
16
17
18
19
Log Population, total


Augmented component plus residual
-1.1
-1.05
-1
-.95
-.9
-.85
15
16
17
18
19
Log Population, total


Augmented component plus residual
-.05
0
.05
.1
0
5
10
15
20
Time trend

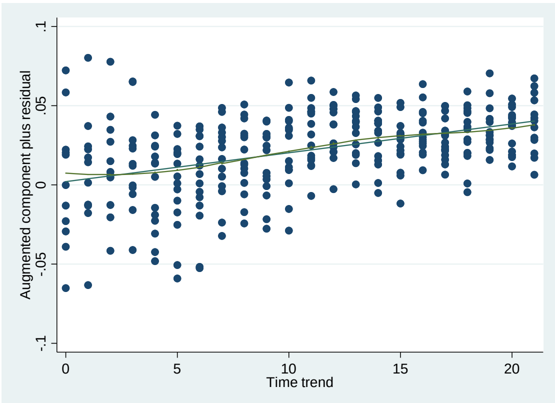

Augmented component plus residual
-.1
-.05
0
.05
.1
0
5
10
15
20
Time trend

Figure 7.0. Residual distribution using Kernel density estimates, natural resource rents model

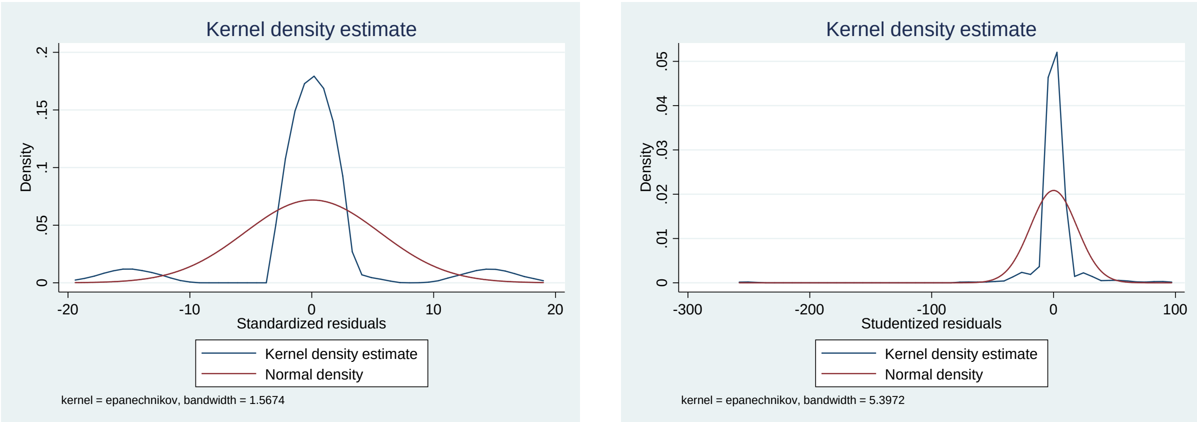


Figure 7.1. Residual distribution using Kernel density estimates, production model

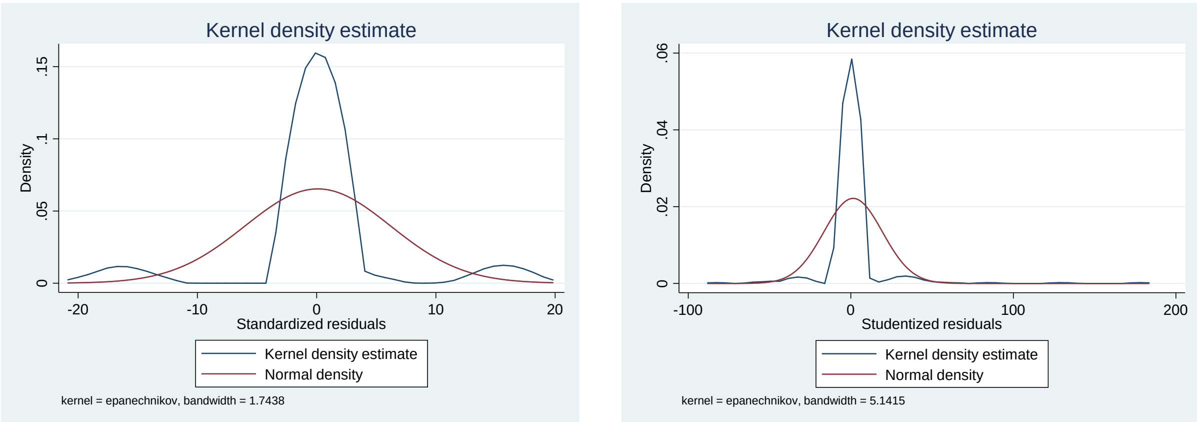


Table 7.1: Investigating heteroscedasticity, production data

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
	Prob > chi2
HDI	0.0000
Petroleum	0.1805
Natural gas	0.3034
Gold_Silver	0.1836
Minerals	0.0397
Governance	0.0312
Democracy	0.2879
W. Prices oil	0.0000
GDP (log)	0.0126
Population (log)	0.0847
Time	0.0000

Table 7.0: Investigating heteroskedaticity, resource rents

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
	Prob > chi2
HDI	0.0000
Governance	0.0007
Natural resource rents (% GDP)	0.0048
Democracy	0.0925
Population (log)	0.9194
GDP (log)	0.0000
Time	0.0000

Table 7.2: Wooldridge test for autocorrelation in nat. rents model

H0: no first-order autocorrelation		
F(1, 14)	=	117.476
Prob > F	=	0.0000

Table 7.3: Wooldridge test for autocorrelation in production model

H0: no first-order autocorrelation		
F(1, 14)	=	114.504
Prob > F	=	0.0000

Table 7.3: Model specification diagnostics for resource rents model, RESET (OVTEST)

Ramsey RESET test using powers of the fitted values of HDI
Ho: model has no omitted variables
F(3, 320) = 3.27
Prob > F = 0.0215

Ramsey RESET test using powers of the independent variables
Ho: model has no omitted variables
F(15, 308) = 3.84
Prob > F = 0.0000

Table 7.4: Model specification diagnostics for resource rents model, using linktest.

Source	SS	df	MS			
				Number of obs	=	330
				F(2, 327)	=	433.56
Model	1.15071635	2	.575358173	Prob > F	=	0.0000
Residual	.43394615	327	.001327052	R-squared	=	0.7262
				Adj R-squared	=	0.7245
Total	1.5846625	329	.004816603	Root MSE	=	.03643

hdi_index	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_hat	.4385457	.5762877	0.76	0.447	-.6951535	1.572245
_hatsq	.4164795	.426739	0.98	0.330	-.4230208	1.25598
_cons	.1877585	.1937291	0.97	0.333	-.1933541	.5688711

Table 7.5: Model specification diagnostics for production data model, RESET (OVTEST) (population and w.prices exluded.)

Ramsey RESET test using powers of the fitted values of HDI

Ho: model has no omitted variables

F(3, 299) = 2.50

Prob > F = 0.0597

Ramsey RESET test using powers of the independent variables

Ho: model has no omitted variables

F(24, 278) = 14.70

Prob > F = 0.000

Table 7.6: Model specification diagnostics for production data model, using linktest production data model (population and democracy excluded.)

Source	SS	df	MS			
				Number of obs	=	311
				F(2, 308)	=	659.83
Model	1.23274625	2	.616373127	Prob > F	=	0.0000
Residual	.287715932	308	.000934143	R-squared	=	0.8108
				Adj R-squared	=	0.8095
Total	1.52046219	310	.004904717	Root MSE	=	.03056

hdi_index	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_hat	1.831247	.4543383	4.03	0.000	.9372473	2.725247
_hatsq	-.6190087	.3377111	-1.83	0.068	-1.283522	.0455042
_cons	-.2766144	.1520463	-1.82	0.070	-.5757953	.0225665

Table 7.7: Comparing r2 and r2 adjusted resource rents model (one control added at a time)

Variable	Base	Democracy	GDP	Population	Time	Access
Governance	-.0447961*	-.0464622*	-.08296832***	.01380657	.00720601	-.01406331
Natural resource rents (% GDP)	.00348856***	.00344732***	.00022116	.00070395**	-.00023135	.00100508***
Democracy		.01218566*				
Population (log)			.19591565***			
GDP (log)				.2907278***		
Time					.00513393***	
Access sanitation						.00509985***
Constant	.67190428***	.6614305***	-1.0735939***	-4.1266373***	.61874945***	.29795989***
N	352	350	330	352	352	352
r2	.14586078	.15766422	.75164956	.84022285	.94250994	.75540475
r2_a	**.10238663**	**.11185744**	**.73811764**	**.83158624**	**.93940237**	**.74218338**

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 7.8: Comparing r2 and r2 adjusted in different resource rents models

Variable	No control var.	Dem, GDP, Time	+ Population	+ access sanit	- Democracy
Governance	-.0447961*	.0034158	.00337046	.00371084	-.00001037
Natural resource rents (% GDP)	.00348856***	-.00024851	-.00026143	-.00026294	-.00025515
Democracy		-.00624492***	-.00628188***	-.00610016***	
GDP (log)		.01361901*	.01254682	.01196788	.01430331*
Time		.00490923***	.00508207***	.00512694***	.00496774***
Population (log)			-.00963	-.00806647	-.00787923
Access sanitation				-.00007379	
Constant	.67190428***	.49854438***	.66464318**	.64882048*	.6172418*
N	352	330	330	330	330
r2	.14586078	.94358552	.94367234	.94369131	.94098991
r2_a	**.10238663**	**.94012786**	**.94002653**	**.93985208**	**.93737316**

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 7.9: Comparing r2 and adjusted r2 in different production data models

Var.	Only prod.	+ Gov	+ Dem	+ GDP	+ Time	+ W.Prices	+ Pop
Petro	.00553301***	.00570355***	.00571904***	.00215465***	-.00043344	-.00049226	-.00052964
Gas	.01092491***	.01066653***	.01055458***	.00152136	.00117499	.00120613	.0012192
G_Si	.00004749	.00068179	.00062498	.0003839	.00031018	.00022488	.00023543
Min.	.0030613**	.00328348**	.00328664**	.00077917	.00048362	.00047267	.00045372
Gov.		-.06368776**	-.06400178**	-.07522421***	.02136384**	.01929168**	.01976702**
Dem			.00148083	.00051494	-.00837792***	-.00812533***	-.00812785***
GDP				.18850881***	.00139312	.00237401	.0029656
Time					.00519016***	.00539914***	.0052577***
W. Pr						.00244631	-.00236508
Pop.							.00835242
Const.	.59660084***	.62284999***	.62202583***	-1.0281548***	.59428029***	.59323811***	.45214432
N	311	311	311	311	311	311	311
r2	.19932473	.21933137	.21947298	.78088528	.94893192	.94915449	.94922251
r2_a	.14996803	.16835988	.16564353	.76496345	.94503089	.94507976	.94496146

legend: * p<0.05; ** p<0.01; *** p<0.001

Appendix 4: Additional analysis results

Table 8.1: Extensive resource rents model with interaction term, including all considered control variables

Variable	mod1 b/t	mod2 b/t	mod3 b/t	mod4 b/t	mod5 b/t	mod6 b/t
Natural resource rents (% GDP)	0.000	0.000	0.000	0.000	0.000	0.000
	(0.38)	(0.37)	(0.34)	(0.33)	(0.49)	(0.42)
Governance	0.011	0.011	0.011	0.011	0.011	0.012
	(0.55)	(0.56)	(0.57)	(0.58)	(0.56)	(0.58)
Democracy	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005
	(-1.22)	(-1.22)	(-1.04)	(-0.97)	(-1.20)	(-0.99)
GDP (log)	0.018	0.017	0.017	0.018	0.019	0.018
	(0.98)	(1.00)	(0.92)	(0.89)	(1.14)	(1.11)
Time	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
	(9.23)	(5.43)	(6.94)	(8.25)	(9.35)	(6.15)
nat_qog	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(-0.68)	(-0.69)	(-0.66)	(-0.65)	(-0.80)	(-0.73)
Population (log)		-0.012				-0.012
		(-0.22)				(-0.18)
Access to sanitation			-0.000			-0.000
			(-0.13)			(-0.03)
Access to water				-0.000		0.000
				(-0.03)		(0.04)
Unemployment					0.000	0.000
					(0.11)	(0.09)
Constant	0.454*	0.663	0.466*	0.458*	0.443*	0.640
	(2.81)	(0.73)	(2.62)	(2.26)	(2.98)	(0.65)
Number of obs.	330	330	330	330	330	330
r2	0.944	0.945	0.944	0.944	0.944	0.945
r2 adjusted	0.943	0.943	0.943	0.943	0.943	0.943

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with clustered robust standard errors in parentheses.

Table 8.2: Extensive production data model including all considered control variables

	mod1	mod2	mod3	mod4	mod5
	b/t	b/t	b/t	b/t	b/t
Petroleum (log)	-0.000	-0.000	-0.000	-0.000	-0.001
	(-0.36)	(-0.59)	(-0.48)	(-0.47)	(-0.78)
Gas (log)	0.001	0.001	0.001	0.001	0.001
	(0.52)	(1.05)	(0.59)	(0.52)	(1.06)
Gold and silver (log)	0.000	0.000	-0.000	0.000	0.000
	(0.00)	(1.10)	(-0.34)	(0.03)	(0.93)
Minerals (log)	0.001*	0.000	0.000	0.000*	0.000
	(2.16)	(1.98)	(2.02)	(2.28)	(1.95)
icrg_qog	0.017	0.021	0.014	0.017	0.020
	(0.70)	(1.12)	(0.58)	(0.74)	(1.08)
GDP (log)	0.006	0.001	0.008	0.007	0.003
	(0.43)	(0.11)	(0.51)	(0.48)	(0.22)
Time	0.005***	0.005***	0.005***	0.005***	0.005***
	(12.74)	(12.79)	(14.53)	(5.56)	(5.41)
Democracy		-0.008			-0.008
		(-1.61)			(-1.51)
log_wprice~l			-0.004		-0.002
			(-1.66)		(-0.86)
Population (log)				0.010	0.008
				(0.20)	(0.17)
Constant	0.546***	0.594***	0.547***	0.384	0.452
	(4.38)	(5.39)	(4.36)	(0.48)	(0.56)
Number of obs.	311	311	311	311	311
r2	0.945	0.949	0.945	0.945	0.949
r2 adjusted	0.943	0.948	0.944	0.943	0.948

Notes:* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with clustered robust standard errors in parentheses.

Table 8.3: Using one year lagged HDI as dependent variable, resource rents model

Variable	mod1 HDI	mod2 HDI	mod3 HDI
Natural resource rents (% GDP)	-0.000	-0.000	-0.000
	(-0.84)	(-1.52)	(-1.62)
Governance	0.003	0.004	0.004
	(0.17)	(0.21)	(0.23)
Democracy	-0.006	-0.005	-0.004
	(-1.21)	(-1.06)	(-0.92)
GDP (log)	0.014	0.006	-0.002
	(0.90)	(0.37)	(-0.12)
Time	0.005***	0.005***	0.005***
	(10.23)	(10.28)	(10.78)
Constant	0.499**	0.564***	0.622***
	(3.85)	(4.40)	(5.34)
Number of obs.	330	330	330
r2	0.944	0.940	0.938
r2 adjusted	0.943	0.939	0.937

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with clustered robust standard errors in parentheses.

Table 8.4: Using lagged independent variables, resource rents model

	mod1	mod2	mod3
	HDI	HDI	HDI
Natural resource rents (% GDP)	-0.000		-0.000
	(-0.84)		(-0.71)
Governance	0.003	0.003	
	(0.17)	(0.14)	
Democracy	-0.006	-0.007	-0.007
	(-1.21)	(-1.31)	(-1.25)
GDP (log)	0.014	0.010	0.009
	(0.90)	(0.66)	(0.65)
Time	0.005***	0.005***	0.005***
	(10.23)	(9.99)	(11.00)
Natural resource rents (% GDP, Lag1)		-0.000	
		(-0.30)	
Governance (Lag1)			0.006
			(0.30)
Constant	0.499**	0.529**	0.539***
	(3.85)	(4.00)	(4.54)
Number of obs.	330	315	315
r2	0.944	0.942	0.942
r2 adjusted	0.943	0.941	0.941

Notes: * p < 0.05, ** p < 0.01, *** p < 0.001

The table report regression coefficients with clustered robust standard errors in parentheses.

Table 8.5: Using lagged HDI as dependent variable, production data model*

	mod1	mod2	mod3
	HDI	HDI	HDI
Petroleum (log)	-0.000	-0.000	-0.000
	(-0.36)	(-0.45)	(-0.47)
Gas (log)	0.001	-0.000	-0.000
	(0.52)	(-0.00)	(-0.09)
Gold and silver (log)	0.000	0.000	0.000
	(0.00)	(0.50)	(0.63)
Minerals (log)	0.001*	0.000	0.000
	(2.16)	(1.35)	(0.11)
Governance (log)	0.017	0.014	0.008
	(0.70)	(0.60)	(0.38)
GDP (log)	0.006	-0.001	-0.006
	(0.43)	(-0.04)	(-0.37)
Time	0.005***	0.005***	0.005***
	(12.74)	(12.10)	(11.44)
Constant	0.546***	0.607***	0.650***
	(4.38)	(4.74)	(5.20)
Number of obs.	311	300	289
r2	0.945	0.940	0.937
r2 adjusted	0.943	0.939	0.935

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with clustered robust standard errors in parentheses.

Mod 1 is the original model with HDI.

Mod2 is using one-year lagged HDI
Mod3 uses a two-year lagged HDI

Table 8.5: Using lagged independent variables, production data model

	mod1 HDI	mod2 HDI	mod3 HDI	mod4 HDI	mod5 HDI	mod6 HDI
Petroleum (log)	-0.000		-0.000	-0.000	-0.000	-0.000
	(-0.36)		(-0.06)	(-0.08)	(-0.23)	(-0.19)
Gas (log)	0.001	-0.000		-0.000	-0.001	-0.000
	(0.52)	(-0.22)		(-0.27)	(-0.35)	(-0.30)
Gold and silver (log)	0.000	0.000	0.000		0.000	0.000
	(0.00)	(0.31)	(0.21)		(0.74)	(0.35)
Minerals (log)	0.001*	0.001*	0.001*	0.001*		0.001*
	(2.16)	(2.40)	(2.41)	(2.53)		(2.33)
Governance (log)	0.017	0.013	0.014	0.012	0.011	
	(0.70)	(0.52)	(0.56)	(0.50)	(0.42)	
GDP (log)	0.006	0.006	0.008	0.007	0.006	0.005
	(0.43)	(0.42)	(0.51)	(0.47)	(0.42)	(0.36)
Time	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
	(12.74)	(13.83)	(12.20)	(12.76)	(12.51)	(13.75)
Petroleum (lag)		-0.000				
		(-0.04)				
Gas (lag)			-0.000			
			(-0.16)			
Gold and silver (lag)				0.000		
				(0.12)		
Minerals (lag)					0.001**	
					(3.97)	
Governance (lag)						0.018
						(0.77)
Constant	0.546***	0.550***	0.535**	0.544***	0.552***	0.560***
	(4.38)	(4.38)	(4.09)	(4.32)	(4.39)	(4.74)
Number of obs.	311.000	298.000	297.000	301.000	300.000	300.000
r2	0.945	0.942	0.941	0.942	0.940	0.943
r2 adjusted	0.943	0.941	0.939	0.941	0.939	0.941

Notes: * p < 0.05, ** p < 0.01, *** p < 0.001

The table report regression coefficients with clustered robust standard errors in parentheses.
Mod 1 is the original model. Mod 2 uses lagged petroleum, Mod3 lagged natural gas,

Mod4 lagged Gold_Silver (gs_lag), Mod5 lagged minerals and Mod6 lagged governance.

Table 8.6: Estimation with two time periods, resource rents model

	mod1	mod2
	HDI	HDI
Natural resource rents (% GDP)	-0.000	-0.000
	(-0.84)	(-1.23)
Governance	0.003	-0.020
	(0.17)	(-0.58)
Democracy	-0.006	-0.007
	(-1.21)	(-0.84)
GDP (log)	0.014	0.124***
	(0.90)	(7.26)
Time	0.005***	
	(10.23)	
Time (periods)		-0.032***
		(-6.77)
Constant	0.499**	-0.428*
	(3.85)	(-2.85)
Number of obs.	330	330
r2	0.944	0.840
r2 adjusted	0.943	0.837

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with clustered robust standard errors in parentheses.

Table 8.9: Estimation with two time periods, production model

	mod1	mod2
	HDI	HDI
Petroleum (log)	-0.000	0.002
	(-0.36)	(1.93)
Gas (log)	0.001	0.001
	(0.52)	(0.32)
Gold and silver (log)	0.000	0.000
	(0.00)	(0.46)
Minerals (log)	0.001*	0.001
	(2.16)	(1.00)
Governance (log)	0.017	-0.013
	(0.70)	(-0.34)
GDP (log)	0.006	0.124***
	(0.43)	(6.65)
Time	0.005***	
	(12.74)	
Time (periods)		-0.027***
		(-6.69)
Constant	0.546***	-0.447*
	(4.38)	(-2.83)
Number of obs.	311	311
r2	0.945	0.852
r2 adjusted	0.943	0.849

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with clustered robust standard errors in parentheses.

Table 9.0: Estimation excluding influential units, Resource rents model					
	Original HDI	Bolivia HDI	Venezuela HDI	Guatemala HDI	Paraguay HDI
Natural resource rents (% GDP)	-0.000	-0.000	-0.001	-0.000	-0.000
	(-0.84)	(-0.86)	(-1.18)	(-0.66)	(-0.86)
Governance	0.003	0.018	0.018	-0.004	0.003
	(0.17)	(1.08)	(0.99)	(-0.16)	(0.15)
Democracy	-0.006	-0.007	0.000	-0.006	-0.006
	(-1.21)	(-1.52)	(0.15)	(-1.04)	(-0.97)
GDP (log)	0.014	0.007	0.015	0.021	0.013
	(0.90)	(0.49)	(0.90)	(1.47)	(0.75)
Time	0.005***	0.005***	0.005***	0.005***	0.005***
	(10.23)	(13.29)	(9.22)	(10.00)	(9.48)
Constant	0.499**	0.548***	0.479**	0.441**	0.508**
	(3.85)	(4.37)	(3.42)	(3.56)	(3.51)
Number of obs.	330	308	308	308	308
r2	0.944	0.958	0.950	0.944	0.941
r2 adjusted	0.943	0.958	0.949	0.943	0.940

Notes: * p < 0.05, ** p < 0.01, *** p < 0.001
The table report regression coefficients with clustered robust standard errors in parentheses.
Bolivia, Venezuela, Guatemala and Paraguay are excluded from the regression separately in each model

Table 9.1: Estimation excluding influential units, Production model						
	Original HDI	Brazil HDI	Bolivia HDI	Mexico HDI	Peru HDI	Venezuela HDI
Petroleum (log)	-0.000	-0.000	-0.000	-0.000	-0.001	-0.000
	(-0.36)	(-0.29)	(-0.50)	(-0.45)	(-0.65)	(-0.50)
Gas (log)	0.001	0.001	0.001	0.001	0.000	0.001
	(0.52)	(0.62)	(1.34)	(0.55)	(0.18)	(0.94)
Gold and silver (log)	0.000	-0.000	-0.000	0.000	-0.000	0.000
	(0.00)	(-0.40)	(-0.16)	(0.17)	(-0.27)	(0.50)
Minerals (log)	0.001*	0.001*	0.000	0.000	0.001	0.001*
	(2.16)	(2.19)	(1.67)	(2.01)	(1.96)	(2.20)
Governance (log)	0.017	0.021	0.018	0.018	0.015	0.039*
	(0.70)	(0.81)	(0.74)	(0.73)	(0.61)	(2.77)
GDP (log)	0.006	0.007	0.005	0.004	0.009	-0.000
	(0.43)	(0.50)	(0.34)	(0.24)	(0.55)	(-0.03)
Time	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
	(12.74)	(12.77)	(12.83)	(11.36)	(12.38)	(13.01)
Constant	0.546***	0.542***	0.554**	0.566***	0.528**	0.595***
	(4.38)	(4.45)	(4.17)	(4.30)	(3.97)	(5.64)
Number of obs.	311.000	291.000	294.000	290.000	292.000	293.000
r2	0.945	0.943	0.951	0.942	0.944	0.958
r2 adjusted	0.943	0.942	0.950	0.940	0.943	0.957

Notes: * p < 0.05, ** p < 0.01, *** p < 0.001
The table report regression coefficients with clustered robus standard errors in parentheses.
Brazil, Bolivia, Mexico, Peru and Venezuela are excluded from the regression separately in each model

Table 9.2: Robustness tests with alternative statistical models (Resource Rents)

	FE CL	FE AR	PCSE	BS
	HDI	HDI	HDI	HDI
Natural resource rents (% GDP)	-0.000	0.000	-0.000	-0.000
	(-0.84)	(0.13)	(-0.42)	(-0.63)
Governance	0.003	-0.002	0.117***	0.003
	(0.17)	(-0.53)	(4.22)	(0.15)
Democracy	-0.006	-0.001	-0.011***	-0.006
	(-1.21)	(-1.12)	(-3.61)	(-1.12)
GDP (log)	0.014	0.041***	0.080***	0.014
	(0.90)	(7.78)	(22.93)	(0.72)
Time	0.005***	0.003***	0.004***	0.005***
	(10.23)	(6.20)	(19.56)	(9.63)
Constant	0.499**	0.292***	-0.143***	0.499**
	(3.85)	(103.30)	(-6.75)	(3.05)
Number of obs.	330.000	315.000	330.000	330.000
r2	0.944		0.603	0.944
r2 adjusted	0.943	0.313		0.940

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

FE CL=Fixed effects clustred standard errors,
FE AR= Fixed effects Autoregresive
PCSE=Panel corrected standard errors,
BS= Bootstrap st.errors

Table 9.3: Robustness tests with alternative statistical models (Production model)

	FE CL	FE AR	PCSE	BS
	HDI	HDI	HDI	HDI
Petroleum (log)	-0.000	0.000	-0.005***	-0.000
	(-0.36)	(0.15)	(-9.44)	(-0.26)
Gas (log)	0.001	0.000	0.008***	0.001
	(0.52)	(0.18)	(7.35)	(0.23)
Gold and silver (log)	0.000	-0.000	0.009***	0.000
	(0.00)	(-0.62)	(7.68)	(0.00)
Minerals (log)	0.001*	-0.000	-0.006***	0.001*
	(2.16)	(-1.04)	(-7.30)	(2.05)
Governance (log)	0.017	-0.005	0.031*	0.017
	(0.70)	(-0.92)	(2.45)	(0.62)
GDP (log)	0.006	0.051***	0.114***	0.006
	(0.43)	(12.22)	(35.23)	(0.45)
Time	0.005***	0.002***	0.003***	0.005***
	(12.74)	(5.37)	(15.41)	(14.66)
Constant	0.546***	0.204***	-0.395***	0.546***
	(4.38)	(79.34)	(-16.18)	(4.65)
Number of obs.	311.000	296.000	311.000	311.000
r2	0.945		0.809	0.945
r2 adjusted	0.943	0.439		0.941

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

FE CL=Fixed effects clustred standard errors,
FE AR= Fixed effects Autoregresive
PCSE=Panel corrected standard errors,
BS= Bootstrap st.errors

Table 9.4: Estimation with alt. dependent variable

	HDI	GDP (log)
Petroleum (log)	-0.000	-0.006
	(-0.36)	(-1.33)
Gas (log)	0.001	0.007
	(0.52)	(0.69)
Gold and silver (log)	0.000	-0.001
	(0.00)	(-0.30)
Minerals (log)	0.001*	0.002
	(2.16)	(0.42)
Governance (log)	0.017	0.391**
	(0.70)	(3.36)
GDP (log)	0.006	
	(0.43)	
Time	0.005***	0.019**
	(12.74)	(3.42)
HDI Index		0.434
		(0.47)
Constant	0.546***	8.410***
	(4.38)	(14.10)
Number of obs.	311	311
r2	0.945	0.803
r2 adjusted	0.943	0.799

Notes: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table report regression coefficients with clustered standard errors in parentheses.

Appendix 5: Data and do-files

The relevant data and do-files are on the accompanying memory stick, or can be provided by request (contact: marianne.andenaes@gmail.com).