Preface

After spending six fantastic years in Trondheim, I have finally achieved what I initially came here to do. A master thesis in Political Science has somehow materialized during the past months, and the process is now at an end. First and foremost, I wish to thank my supervisor, Jonathon Moses. His patience has been beyond expectations, and this thesis would have been much less than what it is if it were not for his encouragement and constructive feedbacks. I am forever grateful. I also wish to thank Tor Georg Jakobsen for much-needed statistical help in times of need. However, it must be emphasized that all errors occurring in this thesis are entirely my own.

Further I wish to thank my best and dearest friend, Thomas. Thank you for your always kind and encouraging words and for always believing in me – and for making me believe in myself. I would also thank my mother and father for everything they have given me through the years. Many an examination period would have been much darker and longer without the prospect of coming home to you. I must also thank my two sisters, Guro and Sigrid, for always answering their phones. I cannot imagine how this past six months would have been without our frequent phone calls.

Finally, a very special thanks is directed to Studentersamfundet in Trondheim and the student festivals UKA-09, UKA-11 and UKA-13 for making my time in Trondheim somewhat longer, but far more memorable than I had ever dreamt. All the long nights, early mornings and the best of friendships, I owe it all to you.

Index

List of tables

List of figures

1 Introduction

Despite its formidable economic growth over the past decades, India still struggles with divergent economic growth across states. Some states are still knee deep in problems related to poverty and slow economic development, while others have surged ahead. As India has large geographical and natural habitat diversity, the occurrences of natural resources are also unevenly distributed across the states (Sachs, Bajpai and Ramiah 2002). Considering the variation in natural resource occurrence together with divergent economic growth, it would be interesting to see whether the divergent economic growth could be explained by the resource curse literature.

There is a large literature on the resource curse internationally and a growing literature on the national level, consisting mainly of work done on the US states. They find the curse to hold at the regional level as well (Freeman 2005; Gerard 2011; Papyrakis and Gerlagh 2007). A study done on the county level, including more than 3000 US counties, have found the resource curse to explain the economic growth also here (James and Aadland 2011). Through these studies, the resource curse is demonstrated to be applicable across levels. The US and India share some of the features that makes it possible to transfer value from the US to India. They both have rather autonomous states, and are thus good units for analysis. They also extend through several climatic zones and the occurrence of natural resources varies considerably within the countries. Considering that the resource curse has been demonstrated to affect economic growth across countries as well as across the US states, it is possible that it will hold at a regional level in India also. This is what I intend to study with this thesis.

1.1 Political motivation

India has experienced remarkable economic growth in the recent years, with an average annual Gross Domestic Product (GDP) growth of 7.3% for the past ten years (World Bank no date, a). This growth has earned India the position as the fourth largest economy in the world today. Despite this impressive growth, a whooping one third of the world's population living in poverty resides in India (World Bank no date, b), making it the country with the largest concentration of poor people today (Topalova 2008:3). With over 400 million citizens living in poverty, it is evident that not everyone has benefited from India's growing wealth (World Bank no date, b). Several empirical studies have been done on the economic disparities across the Indian states, and most find the disparities to widen further (see for example Bhattacharya and Sakthivel 2004; Dasgupta, Maiti, Mukherjee, Sarkar and Chakrabarti 2000; Dholakia 1994; Kurian 2000; Purfield 2006; Rao, Shand and Kalirajan 1999; Sachs, Bajpai and Ramiah 2002; Topalova 2008). Even though the findings are not consistent on this subject area, it is clear that regional disparity in India is of serious concern.

Most of the poor people in India reside in poor and slow growing regions. It is crucial that these regions also benefit from the economic development to avoid potential social, political and further economic difficulties as a result of widening inequality (Purfield 2006:3). The state's legitimacy is dependent on the people's trust in its ability to safeguard their interests. The economic situation and the welfare of the citizens are acknowledged as the most important elements in the formation of trust towards the government. When the economic and social conditions are unsatisfactory, citizens react with lower political and institutional trust (Listhaug and Ringdal 2008; Miller and Listhaug 1999). It can be assumed that these results can be applied to the developing world as well. When political trust is absent, the state's ability to successfully implement rules and reforms might wither and crumble. However, economic growth is no guarantee for equality or eradication of poverty. Neither equality nor poverty alleviation are questions at hand in this thesis, but these problems are highlighted to stress why the divergent economic growth is of political interest, and what the ultimate consequences could be in states where sustained poor economic growth is prevalent.

1.2 Academic motivation

There has been a growing body of literature on the resource curse for the past decades, spreading from the international to the national level. The gap in the literature on the relationship between the divergent economic growth and the geographical spread of natural resources in India is evident. As the resource curse already has been proved to hold at the regional level in the US, we might ask whether this will be the case in India as well. Thus, in this thesis I seek to explore whether some of the divergent growth rates in India can be explained by the resource curse.

Even though India in some aspects makes a good unit for this kind of study, it proves to be a challenging one in other areas. The lack of extensive and available data is an evident problem, hard to overcome. Studies done on the international level and on the US states have included a range of variables in search of what affects economic growth. These will be difficult, or for the most part, impossible to replicate with the data I have at hand. However, we can still establish whether a resource curse is present among Indian states based on the data at hand.

To conduct this study, I use time-series cross-section regression analyses. Based on the resource curse literature and literature from empirical studies on divergent economic growth in India, a set of control variables are constructed. Seven hypotheses are formulated and tested for, seeking to study the fit of the resource curse to the divergent growth rates in India. The hypotheses are presented in Chapter 2.5.

The main hypothesis frames the central question for the study. Hypothesis 1 assumes that natural resource dependence is expected to have a negative impact on economic growth. As the data material used in this study is somewhat scarce, it was difficult to construct as many control variables as desired. Through the literature of divergent growth rates in India, as well as through the available data material, hypothesis 2 is formulated. It assumes that human capital and infrastructure are expected to have a significant impact on economic growth. Hypothesis 3 is associated with this, and seeks to investigate whether resource dependent states have poorer scores on human capital and infrastructure, as is expected.

In addition to investigating the impact natural resources have on economic growth, the contribution of the secondary and tertiary sector can also be investigated. Just as the primary sector is expected to have a negative effect on economic growth, the secondary and tertiary sectors are expected to demonstrate a positive effect. The literatures on divergent growth rates in India highlight the tertiary sector as the most important in creating growth, a contrast to the notion that the secondary usually is the main engine for growth. Hypothesis 4a argues that the tertiary sector has a bigger explanatory power on economic growth than the secondary sector.

As the secondary and tertiary sectors have a reputation for being long-term growth-promoting industries, it is also useful to see what factors seem to drive growth in these sectors. As the IT industries in India is assumed to have a high impact on the growth in the tertiary sector, hypothesis 4b asserts that states with a higher share of tertiary sector in state gross domestic product (SGDP) will also have higher teledensity. Skilled labour and possibilities for freight are expected to be equally important to the secondary sector. Hypothesis 5a therefore assumes that states with a higher share of secondary sector in SGDP will have better possibilities for freight. Hypothesis 5b assumes that states with higher share of secondary sector in SGDP will have better scores on human capital.

The thesis is organized as follows: Chapter 2 constitutes of the applied literature. The chapter is divided into two parts, the first dealing with the resource curse literature, the second with growth promoting factors specific for India. Chapter 3 is the method chapter. The first part accounts for the construction of the variables. The second part deals with time-series cross-section analyses and the statistical assumptions this method tends to breach. The chapter also accounts for the statistical tests that are run. Chapter 4 presents the results from the regression analyses, including tables and figures. Chapter 5 analyzes the results, and interprets the findings. Chapter 6 discuss the results from the regression analyses together with the literature presented in Chapter 2. Chapter 7 summarizes the main findings from Chapter 5 and 6, and concludes. References, Appendix A and Appendix B follow thereafter. Appendix A consists of descriptive statistics and a detailed list of the variables, their sources and construction. Appendix B consists of statistical tests, regression analyses not included in the text and robust testing of the regression results.

2 Theory and empirical research

The extent of the literature on the subject of the resource curse is both too large and too wide for the scope of this thesis. The body of literature comprises of studies done on a variety of aspects of the resource curse and of factors that might impact it. In what follows, I will firstly account for the studies done primarily on the main features of the resource curse. In Chapter 2.2 on transmission channels, I will account for some of the literature done on more specific areas of interest. As previously mentioned, the lack of data poses a problem in recreating the variables used in earlier studies on the subject. This is why I seek to identify factors detrimental to economic growth on the Indian case in the convergence literature in Chapter 2.3. Even though I will not be able to include measures from all the aspects of the literature, it is important to account for the main features of the theory. This is done to give the thesis a sound foundation for the discussion later in the analysis.

2.1 The resource curse

According to basic intuition and conventional economic reasoning, natural resources should be a desired asset to any economy as they can be extracted, exploited and ultimately converted into economic growth (Bulte, Damania and Deacon 2005:1029; Gerard 2011:1; James and Aadland 2010:440). However, studies done on natural resources and their effect on economic growth have found that countries abundant in natural resources tend to have a lower economic growth than countries that are scarce in resources (see Auty 2001; Sachs and Warner 1995; 2001, among others). Through history, several resource scarce countries have had an economic growth outperforming that of contemporary resource rich countries. The resource poor Netherlands outperformed resource abundant Spain in the seventeenth century, and Switzerland and Japan outperformed resource abundant Russia in the nineteenth and twentieth century. In more recent decades, the resource scarce “tiger economies” in East Asia, comprising of South Korea, Taiwan, Hong Kong and Singapore, have surged ahead of several resource abundant countries such as Nigeria and Venezuela (Auty 2001:315; Sachs and Warner 1995:2).

Although there have been cases of resource scarce countries outperforming resource rich countries for a long time, it seems that the scope of the resource curse as we know it today has emerged only the past few decades. Auty finds that the resource curse is sensitive to the time period of investigation, with it holding better in the recent decades (2001:3-4). The explanation proposed is that the previous high transportation costs made the physical location of the resources detrimental to whom could exploit them, and thus create economic growth (Sachs and Warner 1995:3). However, with the revolutions in global freight, transportation has become both easier and cheaper (Sachs and Warner 2001:833), enabling resource scarce countries to become world leading exporters of processed goods after importing cheap natural resources (Sachs and Warner 1995:3).

Even though the pattern of resource abundance combined with poor economic growth had already been observed for some time, the first quantitative analysis on the subject was conducted by Sachs and Warner in 1995. In a cross-country regression using data from 97 countries spanning from 1970-1989, they found a negative relationship between the size of the natural resource sector and economic growth, and that these results held up after controlling for a range of other variables known to be important to economic growth (Sachs and Warner 1995). Since then, there have been done several cross-country studies on the subject of the resource curse. Although the curse poses as a “paradox” or “conceptual puzzle” that runs contrary to conventional economical thinking, the phenomenon has now become a well established truth among most economists (James and Aadland 2011:440; Papyrakis and Gerlagh 2004:182; Sachs and Warner 1995:3). Even though the empirical support for the resource curse hardly is bulletproof, it poses as a quite strong demonstrable empirical fact (Sachs and Warner 2001).

There is no obvious reason why the mere presence of natural resources should frustrate economic growth, and there are resource rich countries that seem to have escaped the curse entirely. Two examples frequently given are Norway and Botswana, who have managed to turn respectively oil- and diamond riches into lasting sustained economic growth (Gylfason 2001:851; Papyrakis and Gerlagh 2007:1012). As Gylfason points out: “it needs to be emphasized that it is not the existence of natural wealth as such that seems to be the problem, but rather the failure of public authorities to avert the dangers that accompany the gifts of nature.” (2001:851). Concluding their study from 1995, Sachs and Warner propose a set of hypothesis about the relationship between resource wealth and economic growth. They further elaborate on these hypotheses in the study from 2001, accounting for the “crowding-out logic”: “Natural resources crowd-out activity *x*. Activity *x* drives growth. Therefore, natural resources harm growth.” (Sachs and Warner 2001:833). However, as there are a diversity of views of what really drives growth, there are also a diversity of opinions about how the resource curse works: “In other words, a complete answer to what is behind the curse of natural resources therefore awaits a better answer to the question about what ultimately drives growth.” (Sachs and Warner 2001:833). Even though there are a range of views on what drives growth and therefore also on how the resource curse works, there are some leading explanations for this phenomenon in the literature. These leading explanations are summed up and presented by Gylfason (2001), Papyrakis and Gerlagh (2003; 2004; 2007) and Sachs and Warner (1995; 2001), all whom have done comprehensive analyses on the subject. Based on these works, four transmission channels are presented.

2.2 Transmission channels

Transmission channels can be defined as the effect natural resources have on other explanatory variables. In their study from 2003, Papyrakis and Gerlagh sum up the transmission channels presented throughout the subject literature. Through cross-country regressions they display the effect natural resources have on other factors detrimental to growth. In doing so, they go a long way in explaining how resource abundance can work through indirect channels, and thus how they indirectly can affect economic growth. Further, they find that natural resources contribute positively to economic growth when all of the transmission channels are accounted for (Papyrakis and Gerlagh 2003). The rest of Chapter 2.2 is based on the leading literature of Gylfason, Papyrakis and Gerlagh, and Sachs and Warner. In what follows, the four transmission channels accounted for by these authors are presented, and their importance in determining economic growth is pointed put. In doing so, I will also draw upon some of the specialized literature on the resource curse. The four channels frequently given in the more narrow literature and summarized and presented in the comprehensive literature are: a) corruption and rent-seeking, b) openness and terms of trade, c) investment and d) education.

All four transmission channels are accounted for, despite the lack of sufficient data to map out all of them in the analysis to come. This is done to provide adequate background information on the hypotheses on how the resource curse affects economic growth. This is essential to understand the discussion in the subsequent analysis. Each chapter is summed up by accounting for how the transmission channel is expected to affect the economic growth in the regression analysis, and whether a variable from the channel are included or not.

2.2.1 Corruption, rent-seeking and institutional quality

Economies relying on natural resources are prone to corruption and rent-seeking. When the international prices for a commodity are high or there are new discoveries of natural resource deposits, the state will have an opportunity to gain vast riches in a short period of time. This could lead to a feeding frenzy as several interest groups will be interested in a slice of the readymade pie (Sachs and Warner 1995:4). Natural resources are often associated with the emergence of politically powerful interest groups that attempt to gain access to the rents through influencing and bribing politicians. Their aim is to get the politicians make decisions favouring them, usually over other, growth promoting activities (Papyrakis and Gerlagh 2003:11; 2007:1022-1023). One of the concerns in the corruption- and rent-seeking literature is that the reallocation of talent and knowledge from growth promoting and productive activities in the economy will be allocated to rent-seeking activities (Leite and Weidmann 1999:5). In their study on the relationship between natural resources and corruption, Leite and Weidmann found that corruption indeed is an important factor in explaining the slow growth for resource rich economies.

They also found that strong institutions could counteract the negative effects of corruption on economic growth (Leite and Weidmann 1999:31). Even though corruption by some is viewed as a desired lubricant making it easier to bypass inefficient bureaucracy and regulations, the most common view is that it impedes growth and hurts innovative activities, and thus is undesired (Mo 2001:66). Unestablished entrepreneurs usually do not have connections or the funds necessary to work their way around a dishonest bureaucracy. Therefore, chances are that their talent and effort will be allocated into rent-seeking activities if they consider it to be easier and more lucrative (Mo 2001:67). Mo also finds that corruption seems to be more prevalent where other forms of institutional weakness can be found as well. This compliments the findings of Leite and Weidmann about strong institutions’ ability to prevent corruption. Mo also considers the possibility that corruption could be interpreted as an indication of a wider set of institutional problems (Mo 2001:76).

Mehlum, Moene and Torvik (2006) confirm the findings of Leite and Weidmann and Mo. In a study from 2006 they show that whether or not a country falls victim for the resource curse is determined by the quality of its political and legal institutions (Mehlum, Moene and Torvik 2006:16). Even though the resource abundance obviously cannot affect the institutions, the institutions can affect how the natural resources are utilized. Thus, the institutions can also have an effect on the economic growth. Mehlum et al. distinguish between grabber friendly and producer friendly institutions, and they propose the explanation that the natural resources are testing their quality. Countries with producer friendly institutions pass the test, while countries with grabber friendly institutions fail (Mehlum et al. 2006:3). By dividing a sample of 42 countries into two groups based on the quality of their institution, they find that the resource curse only occurs among the group with grabber friendly institutions (Mehlum et al. 2006:1).

Robinson, Torvik and Verdier (2006) further explain *how* poor institutions can cause a resource curse. They create a political model of resource extraction, and through this they show that low quality institutions make it easier and more attractive for politicians to redistribute rents from the natural resources to influence the outcome of elections. Thus, low quality institutions invite politicians to make bad policy choices, while high quality institutions make such policy choices infeasible. As politicians seek to redistribute rents during resource booms towards influencing elections and away from productivity, resource booms thus tend to make other parts of the economy inefficient as well (Robinson et al. 2006:465-466). In a selection of resource abundant countries, Robinson et al. show that the countries that have escaped the resource curse score high on an index of institutional quality, and that the countries who have suffered under the resource curse score lower on institutional quality (Robinson et al. 2006:465). This compliments the findings of Mehlum et al. (2006) that high quality institutions can counteract corruption and rent-seeking behaviour, and that the countries with producer friendly institutions seem to have escaped the resource curse.

Corruption and institutional quality can both have an effect on the state level. It is reasonable to imagine the emergence of powerful interest groups on the state level, where it might be easier to influence state officials. State officials are also dependent on re-election to keep their jobs, and there are no obvious reasons to assume that the corruption level or the impacts of interest groups are lower at the state level than the federal level. In some cases, the impact might even be stronger. There might be a stronger feeling of ownership to resources discovered within the home state, giving the public a stronger incentive to exert pressure to get their share of the riches.

In their study on the US states, Papyrakis and Gerlagh use the number of prosecuted corrupted officials per 10 000 citizens as a proxy of corruption (2007:1033). Similar data exist on the state level in India, but not for a long enough time period. As there are no good measures available for the institutional quality either, this transmission channel will not be included in the analysis.

2.2.2 Openness and terms of trade

In their seminal work from 1995, Sachs and Warner find the openness of the economy to be strongly and positively associated with economic growth. In their study, openness is defined as integration in the global economy, and that a country has maintained low tariffs and quotas through the years studied, 1970-1989 (1995:9). For the most part, openness remains significant when other variables are included. Natural resources are found to have a negative impact on openness and terms of trade (Papyrakis and Gerlagh 2004:189), indicating a quite strong effect that cannot easily be explained by other factors. Developing states often pursue protectionist and state-led development strategies to protect domestic producers from competition, and to try to prevent the Dutch Disease (Papyrakis and Gerlagh 2003:12; Sachs and Warner 1995:18). Papyrakis and Gerlagh explain the Dutch Disease as follows:

> [N]atural resource booms increase domestic income and the demand for goods, triggering inflation and an overvaluation of the domestic currency. The relative price of all non-traded goods increase, the terms of trade deteriorate, and exports become expensive relative to world market prices and decline (Papyrakis and Gerlagh 2003:12).

The protectionist and state-led development strategies reduce the openness of the economy, and many see these strategies as harmful for the economic growth in the long run. It will also exclude the state from the global economy, which is assumed to have a negative impact on long-term growth (Papyrakis and Gerlagh 2004:189).

The measurement of openness and terms of trade on the international level cannot be directly transferred to the state level. Most of the political measures to improve openness and terms of trade have to be implemented by the federal government. There is little states can do to enhance their openness in the international market. In their study of the US states, Papyrakis and Gerlagh use international immigration relative to the state's population as a proxy to measure openness. They assume that states with a more open economy will receive more foreign immigrants than states with closed economies (2007:1019). Freeman (2009) and Gerard (2011) don't include a measure of openness or terms of trade at all in their studies on the same topic. Papyrakis and Gerlagh acknowledge the difficulties of measuring openness and terms of trade on the state level, and state that

> We notice that the mechanisms that link resource abundance to openness must be different for the state level when compared to the country level. At a state level, resource abundance cannot lead to a raise in trade tariffs or import quotas...” (Papyrakis and Gerlagh 2007:1021).

The state measures on openness and terms of trade can therefore not be directly compared with those at the national level. It is for example hard to imagine how states can fall victim to the Dutch Disease by an overvaluation of the currency, as the currency is the same across the states. Further, it is not particularly interesting to control for openness when conducting a study where the conditions across the units are quite similar. It is more fruitful to discuss the states differing abilities to attract foreign and domestic investment. Thus, both infrastructure and skilled labor might be deemed as factors affecting the states’ own terms of trade, in that they can make themselves more attractive for investors by improving on these areas. These factors are quite dissimilar across the states, some states being more desirable for investments than others. Variables for both infrastructure and education are included in the analysis. I will get back to this in Chapter 2.3 on growth-promoting factors in India.

2.2.3 Investment

Of the four transmission channels, Papyrakis and Gerlagh demonstrated that investment is the most important (2003:11; 2004:189). Sachs and Warner find an inverse relationship between resource abundance and export growth in manufactures (2001:834). They also find that natural resource intensive economies tend to have a higher price level, which can explain why this is the case. The manufacturing industry is often reliant on domestic input of natural resources in their production. As the price level for the resources needed in manufacturing is higher than normal, it costs more to produce the products, and competitiveness on the global market will suffer (Sachs and Warner 2001:834). Hence, one hypothesis is that the presence of natural resources drives up the manufacture prices, rendering exports uncompetitive. Another hypothesis is that the government fail to actively promote export of manufactures (Sachs and Warner 2001:835). This is elaborated on by Papyrakis and Gerlagh (2007). As natural resources assure a continuous income of wealth for the foreseeable future, this might make future wealth and welfare seem less dependent on investments in manufacturing (Papyrakis and Gerlagh 2007:1020).

Investment can be measured in several ways. Sachs and Warner measure it as real gross domestic investment to real gross domestic product averaged over the time period in their study (1995:24). The same measure is used by Papyrakis and Gerlagh (2003:7; 2004:185). The inclusion of this variable is always significant and positive in these studies and it is robust to the inclusion of several other variables into the model. This indicates that government investment has a positive effect on economic growth, and that this is not easily explained by other factors. In the study by Papyrakis and Gerlagh (2007) done on the US states, investment is measured as "the share of industrial machinery production in GSP[1]". Also here, investment is always positive and significant (Papyrakis and Gerlagh 2007:1033).

It is important to bear in mind the differences in development level between the US and India, both in terms of infrastructure and human capital. The Indian federal government identify education and infrastructure as the biggest obstacles to improving the growth rate in manufacturing (Planning Commission of India 2006:31). These factors need to be in place before investments in this industry will occur, in turn creating economic growth. Considering the developmental differences between the US and India, it is reasonable to assume that the investment in infrastructure and human capital are as detrimental to growth in India as direct investments in manufacturing are in the US.

Thus, investment in infrastructure and human capital can be good proxies of growth-promoting investments tailored to the Indian case. However, there are no good measures available for infrastructure, human capital or investment in the manufacturing sector. It can be assumed that states that invest more in infrastructure and human capital also attract more investment, and are therefore expected to have a higher share of secondary sector in SGDP. To test for these assumptions, interaction terms can be included between the secondary sector and the control variables of infrastructure and human capital. These relationships will be investigated in the analyses in Chapter 4.

---

[1] GSP is the Gross State Product. This measures the same as State Gross Domestic Product (SGDP), and they are used interchangeably. As the Indian government use SGDP, this is the abbreviation I will use throughout this thesis.

2.2.4 Education

Education is widely viewed as a prerequisite for economic growth, as education has several positive ramifications in the economy. Education is known to foster democracy, equality and to improve health, making it an end in itself (Gylfason 2001:851). Another important aspect of education is that more educated people means more people involved in the high-skilled work force, which again functions as an engine for economic growth. Higher education can thereby be directly linked to economic growth.

The primary industry, being based on natural resources, is less dependent on high skilled labour than secondary and tertiary industries. When the primary industry is the most important employer, there might be a bigger incentive to get a job than an education. As education is not a prerequisite to get a job, education thus becomes excess (Papyrakis and Gerlagh 2007:1021). In line with this hypothesis, Sachs and Warner present the explanation that "easy riches lead to sloth" (1995:4). Natural resources are the easy riches, and the sloth occurs when the government fails to see the continued need to invest in education, as the need for high skilled labour is low when the demand for natural resources is high. Consequently, fewer people will bother getting an education, as jobs are easy to get even without an education (Papyrakis and Gerlagh 2003:12).

However, education tends to shift the comparative advantage away from the primary sector towards manufacturing and services in the secondary and tertiary sector (Gylfason 2001:856). The unknown *x* that drives growth is often associated with manufacturing. Forces pushing the economy away from manufacturing lower the growth rate of the economy. As education pushes the economy towards manufacturing, education should be strongly desired by the government as a growth-promoting measure (Sachs and Warner 1995:5). The Indian government acknowledge this, and state that "...it is important to recognize that better health and education are the necessary pre-conditions for sustained long-term growth." (Planning Commission of India 2006:2). The Indian Government has an explicit goal of achieving a higher growth rate in the manufacturing sector. However, it also identifies signs that a rapid growth in manufacturing can result in a shortage of the high quality skills needed (Planning Commission of India 2006:7). Pursuing growth in manufacturing alone will not be sufficient to ensure economic growth. As the primary sector comprises of less high skilled labour, workers released from this sector will not have the skill or expertise needed to offer employers in the secondary industry (Gylfason 2001:856). As a continuation of this, future expansion of other sectors that require high skilled labor will also be restricted by a lack of education. A commitment to investing in manufacturing therefore has to go hand in hand with an equal commitment to education.

In his study on the connection between natural resources, economic growth and education from 2001, Gylfason finds a negative and statistically significant relationship between a country's expenditure on education and natural capital. Countries with a higher share of natural capital in the economy tend to have lower expenditure on education. The same negative and statistically significant relationship is found between natural resources and expected years of schooling for girls, and for secondary-school enrolment for both genders. Gylfason suggests that the country's natural wealth might blind them to the need of educating their children (2001:850). This compliments the hypothesis that "easy riches lead to sloth" presented by Sachs and Warner (1995:4). Other studies done on the relationship between education and economic growth find that gender inequality in education has a negative effect on economic growth (Klasen 2002:345). As it can be assumed that men and women have the same abilities to learn and educate, failing to educate women will rob the society of many bright minds that otherwise could have contributed to economic growth. Klasen further points out that gender bias in education might prevent reduction in fertility and child mortality, as educated women tend to have fewer children and a better knowledge of basic hygiene (2002:346). Educating girls may lead them into jobs and thus reducing fertility. High fertility rates have been demonstrated to have a negative impact on growth. Thus, educating women will lower the fertility rates, which will make the output per person higher (Barro 2001:14).

In the general literature on the international level, Papyrakis and Gerlagh measure schooling or education as the logarithm of average secondary schooling for the time period investigated (2003:7; 2004:192). In their study on the US states, the same authors use "the contribution of educational services in GSP in 1986" (Papyrakis and Gerlagh 2007:1033). Freeman use a measure of the percentage of the population with at least a bachelor's degree (2009:8), while James and Aadland use the percent of the population graduated from High School and from College in their study on the US counties (2011:442). None of these measures fit the Indian case. Again, the issue of levels of development arises. India still struggles with low school attendance and low literacy rates. The idea of higher education is still far from most people's minds. Therefore, measures of literacy rates suit the Indian case better than measures of higher education. Two measures of literacy rates are included in this study. One measures the relative literacy rate, the other measure the divergent literacy rates between genders. These will be elaborated on in Chapter 2.3.1 on human capital. Literacy is expected to have a significant effect on economic growth.

2.3 Growth promoting factors in India

Most of the literatures done on the convergence theory in India tend to conclude that the economic growth is diverging across the states. However, there are studies contradicting these finding, fronting a convergent view. Considering the scope of this thesis, I will not get involved with the question of the economic convergence of the Indian states *per se*. I will simply use the existing literature on the subject to better understand what drives growth in India. As previously mentioned, it will lead to obvious measurement errors if the same measurements from the US case are indiscriminately used to describe the Indian case. If education is measured by those with a bachelor’s degree in India, this would measure the elite rather than the general education level in the population. Therefore, it is important to identify growth factors specific for India.

Although the conclusions are not unequivocally as to whether the poorer states are converging to the richer states or not, they do confirm the fact that there are great economical disparities between the Indian states (Bhattacharya and Sakthivel 2004; Dasgupta et al. 2000; Dholakia 1994; Kurian 2000; Rao et al. 1999; Sachs et al. 2002). Several of these studies also go a long way in explaining what causes these disparities. In doing so, they explore a wide range of factors determining economic growth in the Indian states, trying to identify what promotes growth and what hinders it. Students of the convergence theory in India strive to identify factors driving economic growth, seeking to explain the differences in economic growth across the states. Much in the same way, students of the resource curse strive to identify through which channels the resource curse impacts growth, and which measures can be taken to counteract it. Thus, the convergence theory and resource curse literature both seek to identify growth specific factors. Some of the arguments already made in the previous chapters will therefore be repeated in the chapters that follow.

Each chapter will end with a brief description of what variables I have been able to construct and what they are meant to measure. A more detailed description of each variable with sources and calculations will follow in Chapter 3.1. A detailed description of sources and coding of each variable are attached in Appendix A.2.

2.3.1 Human capital

Education and health indicators are pointed out as important factors for economic growth in India (Dasgupta et al. 2000:2422; Kurian 2000:540; Rao et al. 1999:772). Kurian (2000) emphasizes the literacy level as perhaps the most important indicator of development in a society. On the case of female literacy, he says that "[t]he true index of development of a society is the level of female literacy which can be considered as the bottom line as far as literacy is concerned"(Kurian 2000:540). Although the female literacy rate is a good measurement of development, the gender difference between boys and girls gives an even better indication. The literacy level can be low for both genders, and it is therefore important to either include both genders or look at the gender difference (Kurian 2000:540).

In addition to the literacy level, the infant mortality rate (IMR) and life expectancy are also sound human capital indicators of a society's development. In highly developed societies, women are usually expected to outlive men. In India, the opposite was the case for a long time, indicating poor health care services and a high mortality rate linked to child birth. Kurian found great differences among the Indian states on both cases (2000:541), giving reason to further investigate the relationship between economic growth and health indicators. The Indian government also draws parallels between inequality and development. "Another important divide relates to gender. It begins with the declining sex ratio, goes on to literacy differential between girls and boys and culminates in the high rate of maternal mortality." (Planning Commission of India 2006:10).

In a study from 1995, Guio, Murthi and Drèze link fertility and mortality rates to education, and to female education in particular (1995:747). This link can work through several channels. It is a well known fact that education breeds further education in the next generation. Educated women tend to have higher ambitions for their children, wanting them to get an education. It is easier to follow up on a small number of children through school, rather than many. Educated women are more likely to have other sources of prestige and fulfilment than rearing children. In addition, they are less dependent on their sons for social status and future economic security, giving them fewer incentives for creating large families. Further, childbearing takes time away from economically productive work, which might lead them to minimize such time-consuming activities (Guio et al. 1995:748). Educated women are also more likely to have a higher knowledge about basic health care, nutrition and hygiene. This is most important in rural parts of the country where the health care system is poorly developed. With this basic knowledge, the child mortality rate is expected to go down (Guio et al. 1995:748). Thus, education in itself is likely to reduce child mortality rates and fertility rates. A reduction of fertility is an important contribution to stabilizing the population growth, which in turn will raise the per capita income.

Gender inequality is acknowledged to be a problem in India both by empirical analysis as well as by the government. Also, it is being recognized as an important factor when it comes to economic growth. Fortunately, there are quite a lot of data with many time observations available on human indicators, making it possible to construct several variables. Two variables for literacy have been constructed. *LiteracyRate* is the share of the state population literate, out of 100. *LiteracyDisparity* measure the difference in gender literacy rates. *LiteracyRate* is expected to have a positive effect on economic growth, as higher literacy rates should indicate a more developed state. *LiteracyDisparity* is expected to have a negative effect, as higher disparity indicates lower gender equality. Finally, *IMR* is included as a measure of infant mortality rates. Infant mortality rates are measured by the number of deaths of infants under the age of 1 per 1000 infants born. This is expected to have a negative effect on economic growth, as a high value of infant mortality rates indicates poor health care services and poor maternal health, indicating a low developed society.

2.3.2 Urbanization

Sachs et al. (2002) however, found these human capital factors to have a surprisingly negative effect on economic growth. Literacy rates appear to be positive, but lose their significance once urbanization is included. In other words, the initial observed effect of literacy rates and mortality rates, are also demonstrated by the urbanization variable (Sachs et al. 2002:14). This demonstrates most likely that urbanized citizens have higher literacy rates and better health care. The divide between urban and rural India has become an evident problem (Planning Commission of India 2006:10). Rural India struggles with poor economic growth for several reasons. Even though primary schooling is now a fundamental right, there are few schools and lack of teachers in the rural areas (Bajpai and Sachs 2011:6). Some areas in rural India even still lack sound sanitation systems and clean drinking water. Further, the poor infrastructure, both in terms of freight and IT, contributes to lower foreign and domestic investment rates than in the urban areas (Bajpai and Sachs 2011:21-22; Bhattacharya and Sakthivel 2004:2).

I have access to some data on urban and rural population on the state level. *UrbanPopulationGrowth* is a measure of the state-wise percentage growth in urban population for each year from 1980 to 2009. A measure of the share of the state population living in cities would be a better measure as it would control for the state's population size. However, the urbanization growth might be a good proxy for a transition from the primary sector towards the secondary or tertiary sectors. If this variable successfully measures this development, we should expect high urbanization rate to have a significant and positive effect of economic growth. In accordance with previous findings, the human capital variables are expected to lose their significance, as urbanization will account for some of the same effect as these.

2.3.3 Infrastructure

Access to steady infrastructure enables access to markets, and is vital to manufacturing in attracting both foreign and domestic investments (Bhattacharya and Sakthivel 2004:1071). Broadly speaking, infrastructure varies from roads and railways to energy and communication (Kurian 2000:546; Sachs et al. 2002:9-10). As several of the Indian states have a long coastline, closeness to a big harbour is by Sachs et al. (2002) pointed out as a possible factor explaining economic growth. In China, coastal areas have experienced much faster economic growth than the interior areas. It is therefore natural to assume that the same will be the case in India (Sachs et al. 2002:9). The ports need major improvement to meet the current needs and to bring them up to international standards, both in terms of size and modernization (Planning Commission of India 2006:44). Even though the Indian ports are in need of expansion and modernization, it is likely that the existing ports, and thus the states where they are located, benefit from the pressure from today's needs.

As many states are land locked, the general infrastructures in these states are of especially high importance. Both railway and airport traffic need upgrading and expansion to meet the current and growing needs. Railways can transport large amounts of containers from the inland towards the coast, as well as across the country. Satisfactory railway connection can make up for some of the disadvantage states without coastlines experience, in that transportation from manufacture location to shipping location is made easier (Planning Commission of India 2006:43-44). Expansion and improvement of the road system are of great importance to the rural areas, enabling them to access the market economy. Adequate roads will help improve societal problems as well, linking villages to schools and hospitals. However, as the state governments are responsible for roads within the state and districts, many of these roads have not received the required attention and financing (Planning Commission of India 2006:42). Infrastructure related to the IT business, such as Internet and mobile phone coverage, are also of great importance, especially for the tertiary sector.

Data on railroads in India are available, but most of this data are on the national level. Some data on the state level are given, but there are not enough time measurements to be able to make use of them in this analysis. As the measures of railroads are not expected to change much from year to year, a dummy set or a scale could have been constructed from the available data. However, as there are enough data available on roads on the state level, a proxy for infrastructure have been made using this data.

The variable for the state-wise road density, *RoadDensity*, has been constructed. The variable is measured as road length in kilometres per 100 km^2 of area. As the measurement is neutral to the state's size, it is directly comparable across the units. To distinguish states with a coastline from those without, two different dummy variables are made. *DummyCoastline* distinguishes between states with a coastline and those without. The second coastline variable distinguishes between long and short coastline as well as land locked states by *CoastlineA*, *CoastlineB* and *CoastlineC*. Finally, *Teledensity* is included. This measures the teledensity in each state by the number of telephone connections for every hundred individuals living within an area. As it is not calculated as the total number of telephones, but as the average number of phones within a given area, it is directly comparable across the units.

Preferably, a variable on mobile phones and Internet access should have been included. Unfortunately, there are not sufficient data on this, as there are no measurements on Internet and mobile phones until recent years. Existing data would be a poor measure, as the development on Internet and mobile phone coverage are expected to have developed drastically during the past thirty years. However, it can be assumed that the teledensity works as a good proxy for Internet and mobile phones coverage. Those areas with high teledensity are expected to have a well developed coverage on Internet and mobile phones in addition to telephones.

2.3.4 Sector composition

The stylized fact among many economists is that a transition from agricultural to industrial production is significant in economic development, and is viewed as an important structural change (Kurian 2000:541). In general, the share of the primary sector declines and the share of the secondary sector grows as the economy progresses. The tertiary sector overtakes the secondary sector only when the economy has attained a fairly high level of development. However, the transition in most Indian states has gone directly from the primary sector to the tertiary sector (Bhattacharya and Sakthivel 2004:13). The growth of the tertiary sector reflects its low labour costs and India’s growing work force. The revolution in information and communication technology has enabled India to provide services within the knowledge industry globally (Planning Commission of India 2006:37-38). These industries are usually not dependent on freight to deliver their services, but on sound Internet and telecommunications. Therefore, these industries might be successfully established in inland regions.

It is interesting to see whether the presumption that the secondary sector usually is the main engine for growth, is applicable to India. Empirical findings presented here indicate otherwise, namely that the tertiary sector has had a more important role in economic growth in most Indian states. To test for this, the variables *Secondary* and *Tertiary* are created. These show the annual share of respectively secondary and tertiary sector in total SGDP from 1980 through 2009. Both are expected to affect the economic growth positively. As the tertiary sector often is reliant on infrastructure such as Internet and mobile phone coverage, the interaction term *Teledensity***Tertiary* is included in the regression analyses in Chapter 4. The aim is to test hypothesis 4b, that states with a bigger tertiary sector will have well appraised teledensity. Similarly, the interaction term *DummyCoastline***Secondary* is included to see whether having a coastline have a bigger impact on economic growth, and that this positively affects the secondary sector. There is also reason to assume that the interaction term *RoadDensity***Secondary* will provide similar results, as sound infrastructure is regarded a prerequisite of growth in the secondary sector. The latter two interaction terms test hypothesis

5a, that states with a higher share of secondary sector in SGDP will have better possibilities for freight.

2.4 Growth variation in India

According to the resource curse literature, states abundant in natural resources are expected to experience a lower economic growth rate than those who are resource scarce. Based on data from the 15 Indian states included in my analysis, I will in this chapter try and map out the connection between economic growth and dependence on natural resources. There are 28 states and Union Territory's (UT's) in India, but there are several reasons why only 15 are included in this study. The most obvious is the lack of data on the smaller states. Further, the SGDP for the smaller states are so much lower than for the bigger states, that it is not necessarily wise to compare them without using per capita measurements. As noted in Chapter 3, I have not access to such data, and the smaller states are therefore left out. However, the states included in my analysis are the biggest and most populous states in India, accounting for about 95% of the population (Kurian 2000:539; Sachs et al. 2002:3).

Goa is an exception as it is smaller and less populous than the other states in the analysis. Odisha is excluded, as I was unable to obtain the necessary data for this state. This is a shame, as Odisha is one of the poorest and slowest growing states, as well as being highly dependent on agriculture. The state also has some of the most productive mines and quarrying in the country, further making it an interesting state considering the effect minerals are known to have on economic growth (Sachs et al. 2002:17). This gives reason to expect that if the resource curse is proven to hold at the state level, Odisha should be one of the states hardest hit. Goa is included to make up for losing Odisha, and to make the N larger. Even though Goa is a small state, it is experiencing a high economic growth along with a low dependence on the primary sector, making it a prime unit for this study.

I have obtained my data from Indiastat.com, and the data for each state is divided in three time periods. The first period runs from 1980-1981 to 1992-1993, the second from 1993-1994 to 1999-2000 and the third from 1999-2000 to 2009-2010. All the data was obtained in current values, and re-calculated to constant 2010 prices to make them comparable. However, the data is not necessarily comparable across the three time period for which the data is divided into. As growth per year is calculated by the SGDP from previous year and the current year, it could lead to misleading results if the year values are not suited for comparison. It is the growth rates for the first year in each dataset or time period that could lead to these errors. Therefore, I have simply skipped the growth rates for the first year in each dataset, thereby excluding 1993-1994 and 2000-2001. The same have been done in both Table 2.1 and Table 2.2. Further explanation on this data issue is provided in Chapter 3.1.

2.4.1 Growth rates

As the growth rates from one year to the next are very volatile, I have calculated the average growth rates for the different time periods. In this way it is easier to see the growth trends for each state for each time period, as well as for the entire period.

Table 2.1 Average growth in SGDP

State	1980-1992	1993-1999	2000-2010	All years
Andhra Pradesh	5.5	4.3	6.2	5.3
Assam	5.1	2.4	4.3	3.9
Bihar	3.5	4.0	5.9	4.5
Goa	4.4	9.0	9.4	7.6
Gujarat	5.8	4.9	10.5	7.1
Haryana	5.1	4.7	9.1	6.3
Karnataka	5.3	5.4	6.9	5.9
Kerala	4.1	6.0	7.3	5.8
Madhya Pradesh	3.6	3.4	4.1	3.7
Maharashtra	5.6	4.3	8.0	6.0
Punjab	5.3	3.1	5.1	4.5
Rajasthan	7.4	6.1	5.8	6.4
Tamil Nadu	5.3	4.6	5.8	5.2
Uttar Pradesh	4.0	3.5	5.0	4.2
West Bengal	3.2	5.9	6.5	5.2

States with growth rates below average are highlighted in blue

Table 2.1 shows the average growth rates for each of the Indian states in percentage points. The states with the lowest economic growth rates are highlighted in blue. The state with the lowest growth rate for the period 1980-2010 is Madhya Pradesh, with an average growth rate of 3.7 percent for the past 30 years. By looking at the average growth rates for the three different time periods, we see that the low growth rates for Madhya Pradesh have been sustained throughout. The state experiencing the highest growth rate is Goa, with an average yearly growth of 7.6 percent. However, in the case of Goa, the growth rates diverge from 4.4 percent for the first time period to 9.4 percent for the last time period. It seems that the economic growth accelerated sometime around 1993, when average yearly growth jumps from 4.4 to 9.0 percent.

The states with the lowest economic growth, in addition to Madhya Pradesh are Assam, Bihar, Uttar Pradesh and Punjab, none experiencing average growth above 4.5 percent. Bihar seems to be experiencing an increase in economic growth, but the three others have a rather even growth pattern. On the other end of the scale, we see that Gujarat, Haryana and Rajasthan all experienced an average growth rate over 6.0 percent for the entire period, making them the top scorers when it comes to economic growth. Gujarat and Haryana seem to have experienced acceleration in growth sometime after 2000, when the average growth rates almost doubles. Rajasthan, on the other hand, has experienced a more sustained growth throughout the period, although it actually has declining growth rates with every time period.

2.4.2 Resource dependence

According to the resource curse literature, states experiencing low economic growth will also be more dependent on natural resources. Here, natural resource dependence is measured as the share of SGDP accounted for by the primary sector. The primary sector consists of agriculture, mining, forestry and fishing. I have followed the same procedure as in Table 2.1, with the average for each of the time periods, as well as the average for the entire period. The states with highest share of primary sector in SGDP for the entire period are highlighted in blue.

Table 2.2 Average share of primary sector in SGDP

State	1980-1992	1993-1999	2000-2010	All years
Andhra Pradesh	38.9	34.1	29.1	34.0
Assam	49.8	45.9	36.0	43.9
Bihar	46.0	43.9	30.4	40.1
Goa	21.1	17.0	14.9	17.7
Gujarat	33.9	24.6	19.0	25.8
Haryana	45.9	37.3	24.0	35.7
Karnataka	39.0	32.5	21.2	30.9
Kerala	33.3	28.4	17.5	26.4
Madhya Pradesh	44.5	39.5	31.3	38.4
Maharashtra	23.9	18.6	14.7	19.1
Punjab	46.2	43.2	33.2	40.9
Rajasthan	48.0	36.0	28.4	37.5
Tamil Nadu	33.8	33.1	30.0	32.3
Uttar Pradesh	43.7	37.5	32.1	37.8
West Bengal	32.9	34.5	26.1	31.2

States with the highest average share of primary sector in SGDP are highlighted in blue

At first glance, Table 2.2 shows a pattern somewhat similar as what can be seen in Table 2.1. Assam is the state most dependent on natural resources, with an average of 43.9 percent of SGDP accounted for by the primary sector. Even though this number has decreased, it is still the state most dependent on natural resources with 36.0 percent in the third time period. Assam is also one of the states with the lowest growth in SGDP. Bihar and Punjab are also highly dependent on the primary sector, which accounts for an average of over 40 percent for the entire period, and they both experienced an average economic growth of 4.5 percent. As can bee see from the blue highlighting in Table 2.1, this growth rate earns them a place among the poor performing states. Madhya Pradesh, Uttar Pradesh and Rajasthan all have a primary sector accounting for over 35.0 percent of SGDP for all the time periods. All of them are becoming less dependent on the primary sector for each time period. Rajasthan and Haryana are the only ones among the seven states most dependent on natural resources that do not belong in the group of poorly performing states in Table 2.1.

Goa is the state least dependent on natural resources, with the primary sector accounting for only 17.7 percent of SGDP for the entire period. Second is Maharashtra with 19.1 percent. Gujarat and Kerala both have a primary sector accounting for less than 30.0 percent of the SGDP for the entire period. In general, all of the states have become less dependent on natural resources for each time period. Goa is both the state least dependent on natural resources, as well as being the overall growth winner. Among the eight states being least dependent on natural resources, only Tamil Nadu and West Bengal earn a place among the poor performing states. The other six states being less dependent on natural resources, all belong in the higher end of the growth rate scale.

By applying the resource literature to the Indian case, we should expect to see a pattern where the primary sector accounts for a higher share in the SGDP in the states with a lower economic growth, and vice versa. Assam, Uttar Pradesh and Bihar seem to be the most obvious cases of high resource dependence and low economic growth. Goa is the most striking case of the opposite, low resource dependence and high economic growth.

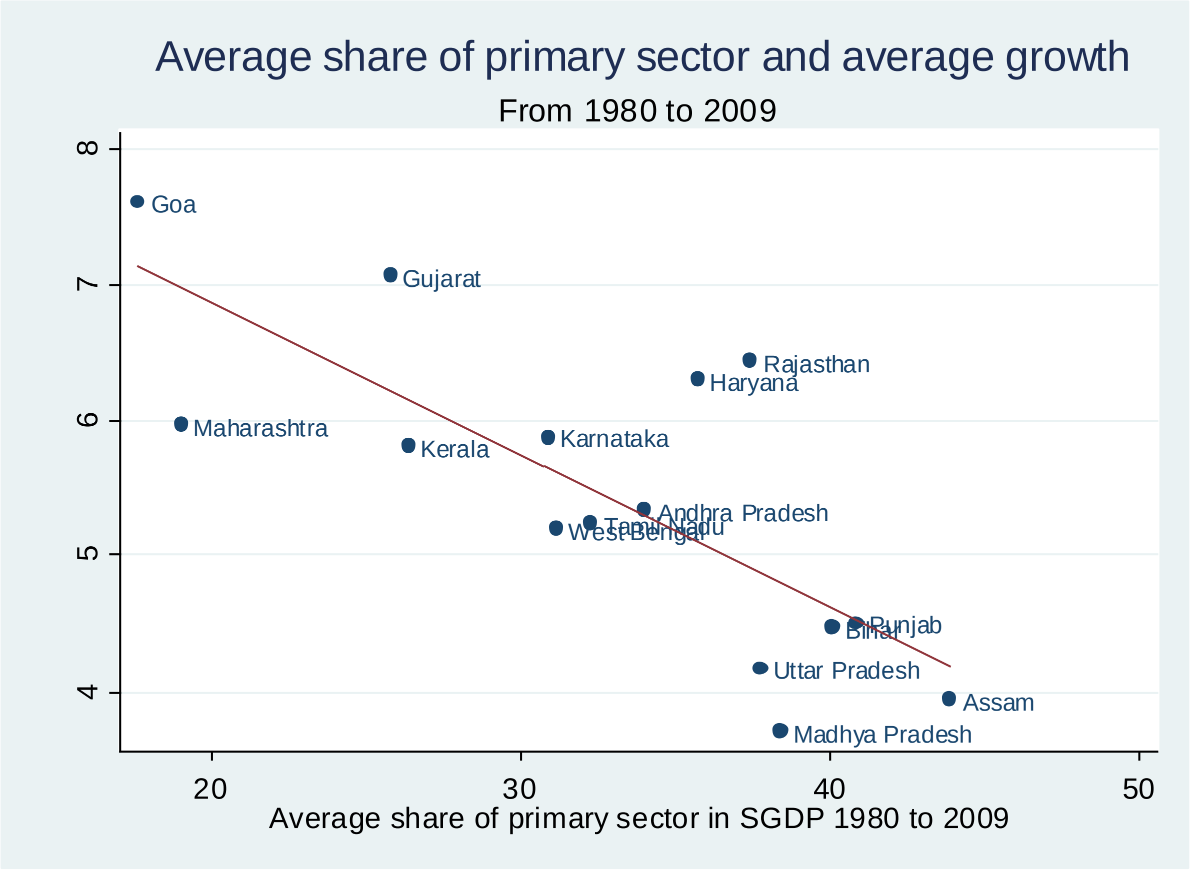

Figure 2.1 Average growth rates of SGDP and average primary sector as share of SGDP for the time period 1980-2009

Figure 2.1 is a scatter plot of average growth in SGDP and average of the share of primary sector in SGDP, both for the entire time period. The line displays a clear negative trend. States with a high share of natural resources in SGDP experience a lower economic growth. These observations are situated at the lower right of the figure. At the upper right, the states with a low share of natural resources in SGDP and high economic growth are situated. Goa clearly demonstrates a high economic growth and a low dependence on the primary sector. This figure confirms the superficial findings in Table 2.1 and Table 2.2.

.

2.5 Hypotheses

In Chapters 2.2 to 2.4, literature and empirical research done on the resource curse and growth determinants in India have been presented. The literature is wide and complex. To summarize the literature, a set of hypotheses is presented. These hypotheses will be the base on which the analyses and discussion will be done.

The paradox of resource abundant countries' slow economic growth has become a well established empirical fact. Based on a growing body of subject literature, this phenomenon has been demonstrated to hold at both the international, regional and municipal level. On the basis of the findings from regional studies conducted on the US states, the resource curse is expected to have explanatory power also on the Indian case, with resource rich states growing slower than those that are resource scarce. The main hypothesis of this thesis is therefore:

> H1: Natural resource dependence is expected to have a negative impact on economic growth. Resource abundant states are expected to have a lower economic growth than those that are resource scarce.

Through the literature on the transmission channels and growth determinants in India, several factors are identified as being harmful or beneficial for economic growth. These factors range from institutional quality to human capital and infrastructure. Considering the available data, only some of these determinants will be directly tested for, but some of them should work well as dummies:

> H2: Human capital and infrastructure are expected to have a significant impact on economic growth.

As human capital and infrastructure have been highlighted as essential factors in creating economic growth, they should be expected to be less prominent in states where economic growth is low. As low economic growth is expected to be caused by the presence of natural resources, we can test for an interaction term between them:

> H3: Resource abundant states are expected to have poorer scores on human capital and infrastructure.

The secondary and tertiary sectors are frequently highlighted in the subject literature as detrimental to creating higher economic growth. The economy is expected to shift from the primary to secondary sector as it develops. A transition to the tertiary sector is expected to take place after the secondary sector is well established within the economy. Evidences indicate that this not necessarily is the case in the Indian economy, as many states have gone directly from being dependent on the primary sector to a high growth in the tertiary sector, bypassing the secondary sector.

> H4a: The tertiary sector has a bigger explanatory power on economic growth than the secondary sector.

The tertiary sector consists of several industries, but the IT business is highlighted as one of the most important industries within it where India has a comparative advantage. The IT business is not dependent on access to freight to deliver its goods. However, it is dependent on Internet and telephone coverage. It is not expected that the tertiary sector is equally important in all states. Based on these assumptions, a test for a hypothesis related to H4a is tested for:

> H4b: States with a higher share of tertiary sector in SGDP will also have higher teledensity.

The secondary sector is associated with economic growth. Even though the tertiary sector is expected to have a greater impact on economic growth in India, the secondary sector should still be expected to affect economic growth positively. Freight and skilled labor have been highlighted as detrimental prerequisites for the secondary sector to thrive. The final hypotheses are therefore:

> H5a: States with a higher share of secondary sector in SGDP will have better possibilities for freight.

And:

> H5b: States with a higher share of secondary sector in SGDP will have better scores on human capital.

3 Data and Method

In this chapter I will account for the variables and the methods used in this thesis.

As sound and comprehensive data on the state level in India is hard to come by, variables necessary for the analyses had to be conducted from scratch. As mentioned earlier, this is also the reason why the dataset used in this study is less detailed than dataset used in international or US studies. Sound data are hard to come by, and if you are lucky enough to find sound data, it is for the most part difficult to get access to. The majority of my data are obtained from Indiastat.com. This is a rather cumbersome webpage, but it contains a lot of data and information. The data obtained from this site are raw material, and all variables used in this study have been processed by me. This chapter gives a detailed description on the processing of data and how the final variables are achieved.

3.1 Variables

As all the data are gathered and processed by me, I wish to account for this process to communicate the choices and considerations that were made in this process. A list of all the variables, their sources and calculations is attached in Appendix A.2.

One issue that arises with this data stretching over the period 1980 to 2009 is the implementation of a new standard of measurement of economic activity by the System of National Accounts (SNA) in 1993 (The United Nations 2014a). This was implemented because of the evolving market economies and economic interactions in the world (The United Nations 2014b).

This new standard set new boundaries for several economic measures, complicating the comparison of data previous to, and after, 1993. However, the Central Statistical Organisation (CSO) in India has extended the 1993-94 series backwards, making them more comparable. Bhattacharya and Sakthivel (2004:3) have discussed this extension of data and found a more sophisticated method of recalculation, taking both price correction and quantum correction into account in a more complicated manner. As I don’t have access to data recalculated using this more sophisticated method used by Bhattacharya and Sakthivel (2004), the original data are used in this thesis.

3.1.1 Growth in State Gross Domestic Product (SGDP)

The dependent variable in this study is economic growth. This is measured as the growth in state GDP for each year from 1980 to 2009. Sachs and Warner use real per capita annual growth rates in their international studies (1995; 2001), and the same measure is used by Papyrakis and Gerlagh (2003; 2004). This measure is also used in studies on the regional level (Freeman 2009; Gerard 2011; Papyrakis and Gerlagh 2007). As previously mentioned, I don't have access to data on growth per capita. Percentage annual growth rates are therefore used instead. All data for the dependent variable have been obtained from Indiastat with the heading *Gross State Domestic Product (GSDP) at Factor Cost by Industry of Origin in* State *{(At Current Prices) (*year *to* year*)}.*

Data have been available for all 15 states included in my analysis. The data on state GDP are divided into three time periods, ranging from 1980-1981 to 1992-1993, from 1993-1994 to 1999-2000 and from 1999-2000 to 2009-2010. This makes up three sheets of data for each state (four for Madhya Pradesh), making a total of 46 sheets which I have combined into one variable. All of the states except for two have a continuous time line of 29 years made up of these three time periods. The first exception is Madhya Pradesh, which, as mentioned, has data drawn from four time periods[2]. The second is Punjab, which is lacking data from 1991 and 1992. These two years are treated as missing in the dataset, causing the observations for Punjab for 1991 and 1992 to fall out of the regressions.

All data are given in current prices when retrieved from Indiastat. I have transformed them into constant 2009 prices using the Indian Consumer Price Index (CPI) for each year from 1980 to 2009 (Indexmundi no date). This enables a comparison of the data across years. As there are no CPI values for each state, the all-Indian CPI have been used. It is reasonable to think that price growth across the states is similar enough to give rather sound estimates of constant prices. To transform prices from current to constant, I have used Equation 1:

---

[2] The data on Madhya Pradesh is from four time periods: 1980-1981 to 1988-1989, 1989-1990 to 1997-1998, 1993-1994 to 1999-2000 and 1999-2000 to 2007-2008. In the data from 1989-1990 to 1997-1998, only the data until 1993-1994 has been used.

$$(SGDP\ State_i\ Year_t * Indian\ CPI\ Year_{2009}) / Indian\ CPI\ Year_t \qquad \text{Eq. (1)}$$

Prices for Rajasthan are the only ones that have been calculated differently, as I was only able to obtain data for the time period 1980-1981 to 1992-1993 in constant 1980 prices. To transform these numbers into constant 2009 prices, Equation 2 was used:

$$(SGDP\ State_i\ Year_{1980} * Indian\ CPI\ Year_{2009}) / Indian\ CPI_{1980} \qquad \text{Eq. (2)}$$

Further, I have calculated the yearly state-wise growth using Equation 3 for each state.

$$(SGDP\ State_i\ Year_t - SGDP\ State_i\ Year_{t-1}) / SGDP\ State_i\ Year_{t-1} \qquad \text{Eq. (3)}$$

However, the growth rates from year to year are very volatile, which the numbers also indicate. Negative growth in one year can be followed by a high positive growth the next year. Thus, it would be wise to look at the growth rate over three to five years, accounting for the volatile growth rates. On the other hand, looking at growth rates per three years would give only one third of the time observations, leaving too few observations to get sound estimates. Another option is to take the average growth for each three years and interpolating them, which might give a more stable time line. However, as the growth rates have a tendency to change from year to year, the interpolation option will not necessarily give us a smooth enough growth rate across time to make it feasible. When we have observations for each year, it is desirable to use them. As the time period is quite long, volatile growth rates are expected to even out in the long run. The average growth rates from Table 2.1 give no reason to suspect otherwise.

Looking at the descriptive statistic of the dependent variable in Table A.1 in Appendix A, it can be seen that the minimum value is -45.38 and the maximum is 42.42. These numbers give reason for a further inspection of the variable, as such growth rates are highly unlikely. Looking at the values in the variable, we see that the first year of the new time periods show very high or very low growth rates. As growth rates are calculated from the last year in the previous time period, these abnormal values indicate that the time periods are not suitable for direct comparison. For example, Bihar has a growth rate of -45 percent in 1993, the first year in the second time period, and 42.4 percent in 2000, the first year in the third time period. Knowing that the growth rates are volatile, the growth rates for the rest of the years in the case of Bihar appear to be normal.

Other states show the same tendency as Bihar, with abnormal growth rates for 1993 and 2000. This indicates that Bihar is not a unique case, and that the variable should be altered in some way before it can be used. To account for the strange behaviour of the growth rates in 1993 and 2000, a dummy variable with the value of 1 is given each state for 1993 and 2000, and 0 for the remaining years are constructed. This dummy must be included in every analysis to account for the odd behaviour of these years. Besides from this, the variable is normally distributed, except for heavy tails caused by the extreme growth rates of Bihar in 1993 and 2000. A simple normality test shows that we have a problem with kurtosis in this variable (Hamilton 2006:129-130). This kurtosis is most likely caused by the abnormal growth values for Bihar in 1993 and 2000. However, as can be seen in Figure B.1 and B.2 in Appendix B, both the values and residuals on the variable are within the normal distribution, and the kurtosis will likely not be a problem in the results.

3.1.2 Natural resources and industrial sectors

The most important independent variable is the measure of natural resource dependence. It is this variable that enables application of the theory on the resource curse to the Indian case. Previous studies have used several different measures of resource abundance. On the international level, Sachs and Warner measure natural resources as the share of primary exports in GDP for the base year in their study, 1971 (1995:8; 2001:830). Measuring natural resources as share of GDP give an indication of how dependent a state is on natural resources and their primary sector. Papyrakis and Gerlagh, on the other hand, use the share of mineral production in GDP for the base year 1971 as a proxy of resource dependency (2003:6; 2004:184). Using mineral production as a proxy excludes important primary resources such as agriculture, which is an important industry in India. Gerard argues that Sachs and Warner’s estimates of natural resources are less than optimal, as it does not take the problem of potential endogeneity into account. The argument made by Gerard is that in countries with low economic growth, the primary sector will be the last one standing, as it requires less skilled labour than the other sectors. This opens up the possibility that the causality runs the other way, and that countries with low economic growth will be more dependent on natural resources due to low economic growth (Gerard 2011:2-3). By distinguishing between dependence of resources and resource abundance, a more precise measure of a country's natural resource wealth can be constructed, and the endogeneity can be accounted for.

However, Gerard (2011) also argues that the abundance, that is, a state's total deposits of natural resources, are much harder to measure and quantify. This makes sound and accurate data hard to come by (Gerard 2011:3). This is probably the reason why most studies on the subject area use resource dependence rather than resource abundance as a measure for natural resources. In his study on the US states, Gerard (2011) is in fact using the same dependency measure as Sachs and Warner (2001), namely the export of natural resources as share of state GDP. This measure is also used by Papyrakis and Gerlagh, with base year 1986 (2007:1033). Freeman, on the other hand, uses the share of total employment accounted for by agriculture and mining, arguing that this is the best available measure for the resource's contribution to economic development (2009:9). Finally, James and Aadland use a measure similar to Gerard (2011) and Papyrakis and Gerlagh (2001), using a measure of the percent of earnings in agriculture, forestry, fishing and mining in 1980 (2011:442).

As Gerard (2011) points put, accurate measures of resource abundance are hard to obtain. Therefore, I make use of a measure of resource dependence frequently used in the subject literature. Based on the available data, I use the state-wise share of primary sector in SGDP for each year from 1980 to 2009. As the primary sector comprises of agriculture, forestry, fishing and mining, my measure resembles the measure used by James and Aadland in their study from 2011. The difference is that I have observations for each of the years, while James and Aadland only used data from 1980 as the base year. To compute the variable *Primary*, I draw upon the same data from Indiastat used to compute the dependent variable, economic growth. As the title *Gross State Domestic Product (GSDP) at Factor Cost by Industry of Origin in* State *{(At Current Prices) (*year *to* year*)}* indicates, the data also contains growth rates per sector and per industry. To calculate the share of agriculture in SGDP as percent, I used Equation 4:

$$\textit{Total Agriculture State}_i \textit{ Year}_t \,/\, \textit{Total SGDP State}_i \textit{ Year}_t \qquad \text{Eq. (4)}$$

Here, I calculated both total agriculture and total SGDP in current prices, as the percentage share would be the same if I had calculated either from current or from constant prices. I created the variables *Secondary* and *Tertiary* using the same approach. The secondary sector, and thus also the variable *Secondary*, consists of manufacturing, electricity, gas and water supply and construction. The tertiary sector, and the variable *Tertiary*, consist of trade, hotel and restaurants, railways, transport by other means, communication, storage, banking and insurance, real estate, ownership, business and legal, public administration and other services (Indiastat no date).

Further, I constructed the variables *Agriculture*, *Mining*, *Forestry* and *Fishing* using the same approach and the same data as for *Primary*, *Secondary* and *Tertiary*. Each is measured as a percentage share of SGDP. These four variables are the industries making up the primary sector. The objective with constructing these variables is to measure the impact each industry within the primary sector has on growth. *Forestry* and *Fishing* make up such a small share of SGDP that they are not expected to have any significant effect. On the other hand, *Mining*, which also includes quarrying, and *Agriculture*, make up a much bigger share of the SGDP in each state, and are thus expected to have some effect. As minerals are notoriously known for their negative impact on growth, it will be interesting to see whether this is the case in India as well. A high share of agriculture in SGDP is viewed as a symptom of a less developed society. Both these effects are possible to investigate further when we have variables isolating their effects on economic growth.

3.1.3 Human capital

Chapter 2.3.1 address the topic of human capital and its effect on economic growth. As mentioned, Indiastat contains a good deal of data on this subject area. I have included three different variables to measure the human capital among the Indian states. *LiteracyRate* measures the state-wise literacy rate, by the number of literates per hundred citizens. Simply put, it measures the percent of the state population that can read and write. Indiastat has census data on state-wise literacy rates from as far back as 1951, up until 2011. The data for all the seven censuses are interpolated, and numbers from 1980 to 2009 are included in the variable. As the variable contains literacy rates from the least developed states in 1980 and the most developed states from 2009, the spread from minimum value of 27.79 to maximum of 93.00 makes sense. The mean of 59.45 is placed right between the min and max, and as the standard deviation is not especially high, this indicates that this variable has a satisfactory normal distribution. A simple histogram confirms this[3].

Indiastat also has data on state-wise literacy rates based on gender. This allows testing for the presumption from Chapter 2.3.1, that disparity in literacy between genders is an important factor in development, and thus also in economic growth. To create *LiteracyDisparity*, the difference between the literacy rates of men and women for each state was calculated using Equation 5:

$$\textit{Literacy rate}_{men}\,\textit{State}_i - \textit{Literacy rate}_{women}\,\textit{State}_i = \textit{Literacy disparity State}_i \quad \text{Eq. (5)}$$

The numbers of literacy disparity were then interpolated, creating *LiteracyDisparity*. Numbers from 1980 to 2009 are included in the variable. A high value of disparity indicates high differences in literacy rates, and thus a low developed society with high gender inequality. A low number indicates a higher developed society with less gender inequality, indicated by equal literacy rates for men and women. The descriptive statistics in Table A.1 in Appendix A indicates a rather normal distributed variable, which a simple histogram also confirms. We should expect both *LiteracyRate* and *LiteracyDisparity* to initially be significant with *LiteracyRate* being positive, while *LiteracyDisparity* being negative. High scores on *LiteracyRate* should indicate a higher developed society, but a high score on *LiteracyDisparity* indicates a high disparity between men and women, and thus a lower developed society. We also should expect the effect of *LiteracyRate* to be accounted for by *LiteracyDisparity*, and that the inclusion of the latter will render the first insignificant.

The final indicator for human capital is state-wise infant mortality rates (IMR). IMR measures the number of deaths of infants under the age of one for each 1000 infants born. Data from 1961, 1971, 1981, 1995, 1999 and 2010 are interpolated, and numbers from 1980 to 2009 are included in the variable. Low values on *IMR* indicates a lower number of infants dead per 1000 infants born, while high values indicate many infants dead per 1000 infants born. *IMR* values are expected to rise with poor infrastructure and low access to health care institutions. *IMR* is therefore a sound measure of several aspects of a state’s development, as will be discussed later. There is a wide range between minimum and maximum value of respectively 13.18 and 133.20. Considering that the developments of health care services and infrastructure have come a long way the past thirty years, these values should not be

---

[3] The histograms of the variables are run in STATA, but not included in the thesis unless there is a problem with the normal distribution.

considered unnatural. The mean is placed between the minimum and maximum values, and indicate a normal distributed variable. This is confirmed by a simple histogram.

3.1.4 Infrastructure

In Chapter 2.3.3 on infrastructure and its impact on economic growth, the background for the variables included in the analysis and what effect they are expected to have is discussed. *RoadDensity* is a measure of the state-wise road density, measured as road length in kms per 100 km^2 of area. Data from 1971-1971, 1981-1982, 1991-1992, 1996-1997 and 2011 are obtained from Indiastat and interpolated in STATA, to get a continuous variable with enough time observations. The data from 1980 to 2009 have been used to construct the variable. By looking at the descriptive statistic in Table A.1 in Appendix A, it can be seen that the minimum value is 74.94 and the maximum value is 31 166.87. As this variable is a measure of the development of building of roads for the past thirty years, this big difference in road density it is not surprising. However, the mean is only 1 215.77, which is pretty far from the maximum value of 31 166.87, indicating that the variable is not normally distributed. The very high standard deviation of 2 186.067 further underlines this assumption.

A simple histogram of the variable displays a highly positive skew. Skewed variables can become more normally distributed through a log transformation, or a power transformation (Hamilton 1992:17-19). Although the independent variables are not required to be normally distributed, skewed variables are associated with statistical problems such as influence and heteroscedasticity (Hamilton 1992:55). Even though the interpretation of a log variable is somewhat different and more difficult, Figure 3.1 and Figure 3.2 demonstrate that the variable is much more normally distributed after the power transformation. The power transformed variable is therefore the better choice. The descriptive statistics for the transformed variable is also included in Table A.1 in Appendix A.

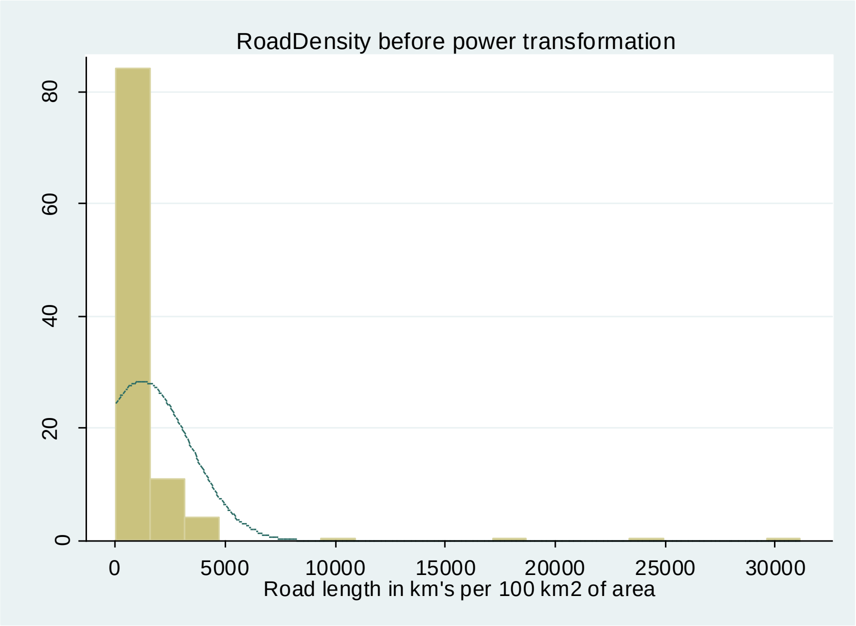

Figure 3.1 RoadDensity before power transformation

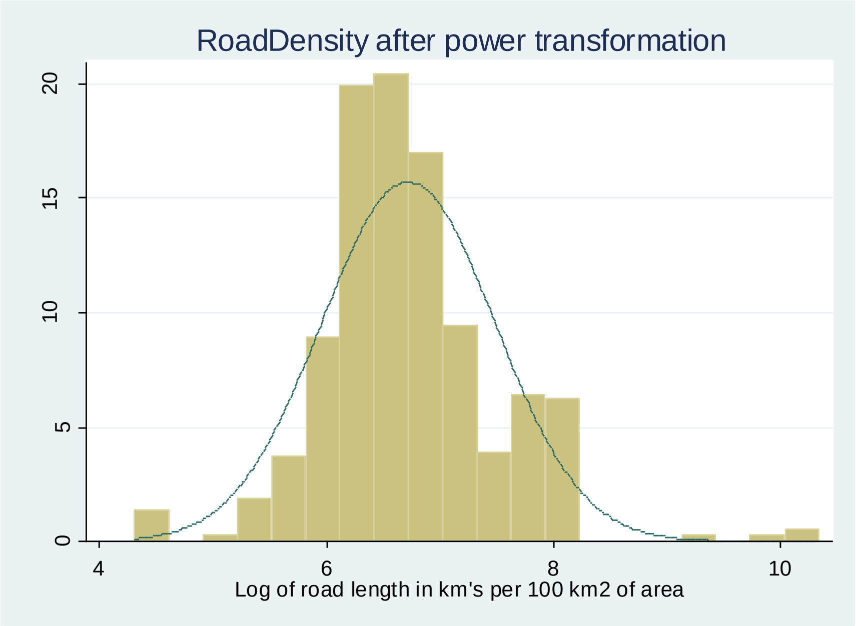

Figure 3.2 RoadDensity after power transformation

As a measure of infrastructure related to the IT sector, I have made a variable measuring the teledensity. Ideally, a variable measuring both Internet and mobile phone coverage should have been included as well. However, as the development and expansion of Internet and mobile phone was only begun in recent years, no such data exist on the state level until the turn of the century. As previously discussed, the development of Internet and mobile phones have sky rocketed the past thirty years, and the recent data cannot be made use of. *Teledensity* is expected to serve as a proxy for the Internet and mobile phone coverage. It is reasonable to assume that the areas with highest teledensity also have the best developed Internet and mobile phone coverage.

Data for the variable are obtained from Indiastat for 1992, 1999-2000, 2002, 2003 and 2012. These have been interpolated, creating a continuous time line. As no data exist on state-wise teledensity previous to 1992, the twelve years prior to this are labelled as missing. However, we still have 17 years of data, which is enough to get rather accurate estimates. From the descriptive statistics we see the low number of observations compared to the rest of the variables, with N = 241. As the minimum value is 0.15 and the maximum value is 71.18, there is a wide spread in the variable. However, as the mean is only 10.97, much lower than the maximum value, and the standard deviation is somewhat high at 15.041, this variable should be further investigated as well. By looking at a simple histogram, we see that also *Teledensity* has a positive skew, and that it might benefit from a power transformation. The histograms in Figure 3.3 and Figure 3.4 display that even though the power transformation has not been as effective as it was for *RoadDensity*, the transformation has given a more normally distributed variable. We choose to use the transformed variable in our analysis. Descriptive statistics for the transformed variable is also included in Table A.1 in Appendix A. It is notable that this variable only has 227 observations, leaving out nearly half of the units of the regression. This variable should therefore be included with some care, as the inclusion of this will also exclude the rest of the units prior to 1992.

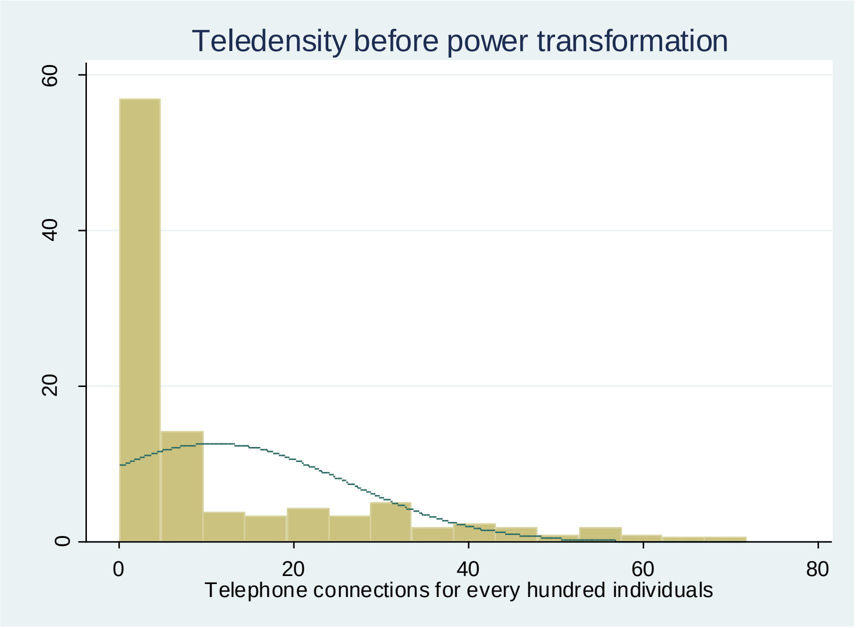

Figure 3.3 Teledensity before power transformation

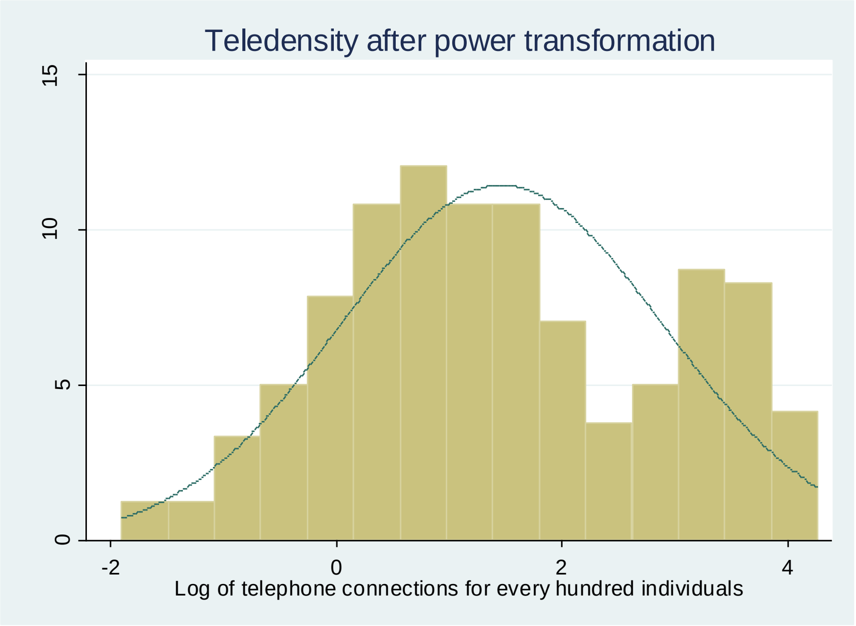

Figure 3.4 Teledensity after power transformation

As discussed in Chapter 2.3.3, closeness to harbour and coastline are pointed out as important for freight of goods, and thus for economic growth. This is measured by two different dummy variables. Data for both are retrieved from Agriculture Research Data Book (2002). The first dummy variable is *DummyCoastline*, simply distinguishing between states with and without a coastline. States with coastline=1, states without coastline=0. As it is possible the length of the coastline has an impact on growth as well, I have made a dummy set with *Coastline A*, *Coastline B* and *Coastline C*. *Coastline A* constitutes of states with a coastline longer than 900km., *Coastline B* constitutes of states with a coastline shorter than 900km., and *Coastline C* constitutes of the states without coastline. *Coastline A* constitutes of Gujarat, Kerala, Andhra Pradesh and West Bengal, all have the value 1. The rest have the value 0. *Coastline B* constitutes of Maharashtra, Tamil Nadu, Karnataka and Goa, all with the value 1. The rest have the value 0. The length of coastlines in this group ranges from 300km. to 840km. The states in *Coastline C*, those without coastline, are Assam, Bihar, Haryana, Madhya Pradesh, Punjab and Uttar Pradesh. All have the value 1, and the states with a coastline have the value 0. However, as *Coastline C* is the reference category and is never included in the model, this is not of much importance. It does however have an importance in interpreting the variables. Both *Coastline A* and *Coastline B* are expected to have a positive and significant effect, as they both are compared to *Coastline C*. *Coastline A* are expected to have a higher value than *Coastline B*, thus indicating that states with longer coastline have higher economic growth than states with shorter coastline. Both *Coastline A* and *Coastline B* are expected to have significantly higher growth than states without coastline.

UrbanPopulationGrowth measures the growth in urban population, measured as the state-wise growth in urban population for each year for the time period 1980 to 2009. Data on the state-wise urban population from 1971, 1981, 1991, 2002, 2003, 2004, 2008, 2009 and 2010 are obtained from Indiastat and interpolated. Then, I used equation 5 for each state to get the growth rate for each year.

$$(Urban\ Population\ State_i\ Year_t - Urban\ Population\ State_i\ Year_{t-1}) / Urban\ Population\ State_i\ Year_{t-1} \qquad \text{Eq. (6)}$$

The variable is not very far from being normally distributed, but it overshoots on the mean. In other words, the mean is a lot bigger than it should be if it had a normal distribution. A power transformation makes its distribution look prettier, but the R^2 in the regression analyses are reduced, making it a poorer fit than the original variable. *UrbanPopulationGrowth* is therefore kept as it is.

3.2 Time-Series Cross-Section analysis

The method used in this thesis is a Time-Series Cross-Section analysis (TSCS). This is a method often used by students of comparative political economy, as it enables studies of effects across time and space at the same time. The name Time-Series Cross-Section refers to the construction of the data, with multiple observations over time for a fixed set of units of interest. The set of units (N) typically ranges from 10 to 100, but there are studies with much higher N. Studies done on the 14 OECD nations is a typical example, but studies are also frequently done on the 50 US states (Beck and Katz 1995:64). There is no strict lower or upper limit on the number of time observations (T) either, but there have to be enough observations to make sense of averaging operations. 20 to 50 time observations are normal. A T below 10 is not desirable, while a bigger T is more helpful than harmful (Beck 2001:272; Beck and Katz 1995:64). As the data in this study has 15 units with 29 time observations, it fit sthe requirements of a TSCS analysis rather well. The regression equation for a TSCS model is:

$$y_{i,t} = x_{i,t}\beta + \varepsilon_{i,t}; \quad i = 1,\dots,N; \quad t =,\dots,T, \qquad \text{Eq. (7)}$$

All units are characterized by the same regression equation at all points in time (Beck and Katz 1995:636). All observations in the dataset have observations for unit (i) and time (t), allowing to investigations of changes over both time and space.

Although there are several beneficial aspects of conducting a TSCS study, the nature of the data breaches many of the Gauss-Markov assumptions. When these assumptions are breached, standard ordinary least squares (OLS) estimations will render inaccurate estimates (Beck 2001:275). TSCS data tend to have correlated errors, known as autocorrelation. As TSCS is measured both across time and space, the data might be autocorrelated across time and space. In other words, the unit’s values are affected by each other, and this year’s values are affected by last year’s values (Beck and Katz 1995:634). Temporal and spatially error correlation bias the estimation of the standard errors in OLS (Hamilton 1992:118). As the standard errors are used to estimate the significance level through the value of *t*, using Equation 7, a wrong estimation of standard errors will also give us imprecise significance levels.

$$t = \frac{\beta}{SE_{\beta}} \quad \text{Eq. (8)}$$

Another assumption usually breached, is that of homoscedasticity, or constant variance in the errors (Hamilton 1992:111). Ideally the variance in the predicted values should be the same for all values within a variable. With heteroscedasticity, the variance in predicted values will not be equal for all values. Heteroscedasticity will also lead to inefficient measure of the standard errors, again giving imprecise significance levels through imprecise values of t.

Influential cases, or outliers, are also a problem that can cause biased estimations. This is not especially prominent in TSCS data, but is a problem in general that should be tested for. The problem with influential cases is that they can substantially affect the regression results, and should therefore be accounted for. If we have measures that have a substantial impact on the regression results, they are influential cases, and should be dealt with in some way. As we already have seen, we are likely to have this problem with the extreme values of Bihar for 1993 and 2000.

Multicollinearity is another problem that can arise. Multicollinearity is caused by too high correlation between two or more variables included in the analysis. Put differently, we get a problem when variables included in the analysis measure too much of the same. Abnormally high standard deviations might be a symptom of multicollinearity. The same could be the case if a variable expected to come out positive, is negative (Hamilton 1992:133-135). Once again, the standard deviation is used to calculate the significance level, and biased standard deviations can therefore become a serious problem. All of these assumptions will be tested for in the next chapter.

3.3 Testing for statistical assumptions

As presented in the previous chapter, there are several bumps in the road that needs to manoeuvre around before the regressions can be embarked on, if we want the answers to be reliable. However, there are several tests and alterations that can be done to prep the data for analyses. In what follows, I will account for the tests I have done and the alterations I have made.

3.3.1 Multicollinearity

We should expect to find multicollinearity for several of the variables, as I some of the variables are intentionally constructed to measure the same thing. For example, *Primary*, *Secondary* and *Tertiary* together make up the SGDP for each state, and measure sector composition and contribution to growth. They are all a measure of the same, and can therefore be expected to have a high intercorrelation among them. The same is likely to be the case for *LiteracyRate* and *LiteracyDisparity*. *LiteracyDisparity* and *LiteracyRate* are both calculated from data expected to be very similar.

To test for multicollinearity, we use the tolerance test. The tolerance test tests the intercorrelation among the X variables included in the analysis, and how much of the variation in the X variables are dependent on the other variables. A tolerance level below .1 or .2 is problematic. If the tolerance level is at .1, only 10 percent of the variation in the variable is independent of the other variables (Hamilton 1992:134). The tolerance tests are included in Appendix B. Table B.1 is a tolerance test for the entire model, and we can see that out of 16 variables, only six are above the critical value .2. This is expected, but still problematic, as a model including all of these variables will lead to substantially higher standard errors, and correspondingly lower *t* statistics and significance levels (Hamilton 2006:202).

In Table B.2, I have tested the intercorrelation between the sectors. The very low tolerance values confirms that *Primary*, *Secondary* and *Tertiary* indeed measure the same, and that they should not be included in the same analysis. Further, Table B.3 demonstrates that the four sub industries of *Primary*, *Agriculture*, *Mining*, *Forestry* and *Fishing* don’t measure much of the same. Their variance is highly independent of the three others. In Table B.4 I have tested the tolerance of the control variables. None are initially problematic, but *LiteracyRate*, *LiteracyDisparity* and *Primary* fall below .2 when *Primary* is included in the regression. As the main purpose is to test the effect *Primary* has on economic growth, it must be keep it in the analyses. However, both *LiteracyDisparity* and *Primary* rise above the .2 limit when *LiteracyRate* is excluded from the regression. I choose to keep *LiteracyDisparity* instead of *LiteracyRate,* as it is considered a better measure for education.

3.3.2 Influential units

The reason to test for influential units is to detect units with values that differ substantially from the other observations. Influential units change the regression output substantially if deleted, and should thus be controlled for (Hamilton 1992:125). Here, I will only test for Cooks D, which measures each unit's influence on the model as a whole (Eikemo and Clausen 2007:165). As have already been seen, some of the values in the dataset give reason for concern, such as Bihar's unusually high and low growth rates in 1993 and 2000. However, this should be accounted for by the dummy variable *DummyGrowth*. Regression analyses controlled for cook's D can be found in Chapter B.4 in Appendix B.

3.3.3 Autocorrelation

Ideally, the data should be free of autocorrelation. That is, the observations should be independent of each other. However, as mentioned, TSCS data tend to correlate with each other both contemporaneously and temporally (Beck 2001: 275). Contemporaneous correlated errors occur e.g. when the errors of one Indian state correlates with the error of another state. This is usually seen between units that are neighbouring countries or close trading partners (Jakobsen and Jakobsen 2007:304). Temporal correlated errors occur when there are correlations between consecutive values. An example is that the economic growth of State_i for Year_t is affected by the economic growth of State_i for Year_{t-1}. The problem with autocorrelation causes biased standard errors, and thus renders the *t* values and significance level invalid (Hamilton 1992:113).

A test for unit roots have been conducted to see whether there are problems with temporally correlated units in this data Unit roots indicate that out data is not stationary, and that the mean and variance change over time (Hamilton 2006:376). To account for this, the variables can be lagged. When an X variable is lagged with -1 year, it indicates that this year's value on Y is determined by last year's value of X. An augmented Dickey-Fuller test, have been conducted on the dependent variable, growth. It does not have a unit root, and should therefore not be lagged. All of the independent variables have non-stationary values, and should thus be lagged. All the independent variables are lagged by one year, except for the coastline dummies, whom are time constant. Robust tests with two and three year lags are attached in Chapter B.4 in Appendix B.

The lagging of the dependent variables is also supported by theory. The building of roads and expansion of the telephone connections are not expected to have an immediate effect on the SGDP. The same is the case for literacy levels and IMR. A positive or negative trend in literacy rates and IMR should not be expected to affect the economic growth until some time has passed. These variables change rather slowly over time, and are, generally speaking, not very volatile. Sector compositions should not be expected to have a big impact on the state's economic growth in the current fiscal year. Thus, a one year lag of all the independent variable can be argued both theoretically and statistically. Through a Durbin-Watson test, we can test for contemporaneous autocorrelation. This test demonstrates that there are no contemporaneous autocorrelation in this data. A Breusch-Pagan test indicates that there are no problems with heteroscedasticity either (Wooldridge 2006:281).

3.3.4 Heterogeneity and fixed or random effects model

It would be naïve and contradictory to the purpose of this thesis to assume homogeneity across the Indian states. The economic disparities and the disparity in the distribution of natural resources are at the core of this study. Earlier, heterogeneity across units was viewed as nuisance, and ignored (Beck 2001:282-283). Heterogeneity can be modelled by allowing the units to have their own intercept on Y. Beck argues that fixed effects are appropriate for TSCS data analysis, as it is the units chosen for the analyses that are of interest, not the units as a sample from a larger population. The random coefficients model, on the other hand, assumes that the sample is drawn from some distribution (Beck 2001:284). However, the fixed effects model does not allow inclusion of time constant variables, such as the dummies for coastline. As *DummyGrowth* has to be included in all analysis to control for the abnormal growth rates of 1993 and 2000, the fixed effects is a poor fit for our model. A Hausman test confirms that random effects is the best fit for our analysis.

4 Results

In this section the results from the regression analyses are presented. I will interpret the output for each model shortly. The in depth analysis will follow in Chapter 5, along with the discussions. In Chapter 4.1 the initial findings and the work towards findings the best model to explain economic growth are presented. These regression analyses are based on the hypotheses H1, H2 and H4a. The main hypothesis, H1, assumes that states dependent on natural resources experience a lower economic growth than states that depend on other sectors. According to hypothesis 2, the human capital and infrastructure should have a significant impact on growth. Hypothesis 4a expects the tertiary sector to have a greater impact on economic growth than the secondary sector. In Chapter 4.2, interaction terms are included to investigate the remaining hypotheses.

The regression analyses are carried out by including one variable at a time to see the impact each variable has on growth. *DummyGrowth* controls for the abnormal growth rates in 1993 and 2000, and are therefore always included.

4.1 Natural resources

By including one variable at a time, the impact each of the variables have on economic growth and on the effect of *Primary* can be seen. This is presented in Table B.6 in Appendix B. *Primary* is significant on the 0.01 level in all models. None of the other variables have a significant effect on growth when included alone with *Primary*. However, this does not mean that they don’t have an impact in determining economic growth in other models.

Table 4.1 Primary resource's impact in economic growth

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	12.87***	12.31***	9.527***	9.623***	19.24***	14.45***
	(-2.637)	(-2.658)	(-0.955)	(-0.959)	(-4.070)	(-2.579)
Primary L1	-0.160***	-0.154***	-0.103***	-0.105***	-0.141***	-0.104***
	(-0.035)	(-0.037)	(-0.029)	(-0.030)	(-0.039)	(-0.032)
DummyCoastA	-0.351	-0.216			-0.693	
	(-0.583)	(-0.601)			(-0.505)	
DummyCoastB	-1.33	-1.093			-1.208	
	(-0.928)	(-0.951)			(-0.871)	
RoadDensity ln L1	-0.263	-0.216			-0.905**	-0.601*
	(-0.248)	(-0.243)			(-0.399)	(-0.313)
UrbanPopGrowth L1		0.075*		0.074	0.100**	0.110**
		(-0.045)		(-0.052)	(-0.046)	(-0.054)
LiteracyDisparity L1			-0.012	-0.012	-0.065	
			(-0.045)	(-0.045)	(-0.047)	
IMR L1			-0.008	-0.007	-0.019	-0.024**
			(-0.013)	(-0.012)	(-0.015)	(-0.012)
DummyGrowth	5.314***	5.290***	5.313***	5.282***	5.380***	5.327***
	(-1.763)	(-1.778)	(-1.743)	(-1.763)	(-1.769)	(-1.772)
R^2	**0.078**	**0.077**	**0.074**	**0.075**	**0.081**	**0.077**
N	419	403	419	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

All variables lagged with one year are labelled with L1 at the end. Power transformed (*RoadDensity* and *Teledensity*) variables are labelled with ln.

Table 4.1 demonstrates that several of the variables have a significant impact on growth at some point. In model 1, the dummies for length of coastline are included with *RoadDensity*. These variables are meant to measure the infrastructure and possibilities for freight, and were expected to be positive. As can be seen in model 1, they are negative, but not significant. In model 2, *UrbanPopulationGrowth* is included with the variables for infrastructure from model 1. Growth in urban population was expected to have a positive impact on growth, as growth in urban population might indicate a transition away from the primary sector. *UrbanPopulationGrowth* is both positive and significant at the 0.1 level. None of the variables from model 1 have become significant by the inclusion of *UrbanPopulationGrowth*. Both measures for human capital, *LiteracyDisparity* and *IMR* are included in model 3. In line with what was expected, they both have a negative sign. However, they are both insignificant. In model 4, *UrbanPopulationGrowth* is included with the variables for human capital. This variable was predicted to render both of the human capital variables insignificant. Instead, none of them are significant, indicating that they might measure too much of the same.

All control variables are included in model 5, except for *DummyCoastline*. As the dummies measuring the length of coastline are expected to be a better measure, *DummyCoastline* is left out. Two similar dummies would measure too much of the same, likely rendering both insignificant, as well as being problematic in terms of collinearity. As model 5 demonstrates, *Primary* is still significant at the 0.01 level, and has a higher coefficient than in the three previous models. This indicates that the effect of the primary sector is robust to the other variables in this analysis. However, the R^2 of model 5 indicates that all of the variables together only explain 8.1 percent of what affects economic growth among the Indian states. Along with *Primary*, both *RoadDensity* and *UrbanPopulationGrowth* become significant at the 0.05 level.

Model 6 summarize the initial findings in Table 4.1 and the presumptions from the theory chapter. The dummies measuring the length of coastline have been taken out, as they have failed to demonstrate a significant effect on growth. *LiteracyDisparity* is also taken out of the model, as it has not demonstrated any significant effect. This is likely due to *IMR*, expected to measure some of the same. *IMR* is also considered to be a better measure of development. All remaining variables in model 6 demonstrate a significant effect on growth. Even though model 5 has a better fit measured only by the R^2, it is not a significantly better model, as fewer variables are significant. Regardless, R^2 will increase as we include more variables, even if the model is not necessarily a better fit. *RoadDensity*, *UrbanPopulationGrowth* and *IMR* all have a significant impact on growth. *Primary* is still significant at the 0.01 level. Even though the coefficient has decreased from model 5, it still has a clear negative impact on economic growth.

4.2 Agriculture

As *Primary* is comprised of four industries depending on natural resources, it might be a too general measure of resource dependency. In Table 4.2, the four industries of which the primary sector consists are presented. These are *Agriculture*, *Mining*, *Forestry* and *Fishing.* The intention is to measure each of the industry's individual impact on growth, to better understand what *Primary* really measures.

Table 4.2 Primary sector's independent impact on economic growth

Variables	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Constant	8.438***	6.062***	5.605***	5.152***	9.340***
	(-0.672)	(-0.428)	(-0.250)	(-0.440)	(-0.855)
Agriculture L1	-0.109***				-0.112***
	(-0.025)				(-0.024)
Forestry L1		-0.535**			-0.327
		(-0.220)			(-0.301)
Mining L1			-0.101		-0.084
			(-0.092)		(-0.095)
Fishing L1				0.191	-0.169
				(-0.254)	(-0.216)
DummyGrowth	5.221***	5.256***	5.280***	5.248***	5.246***
	(-1.778)	(-1.756)	(-1.735)	(-1.749)	(1.779)
R^2	**0.068**	**0.047**	**0.042**	**0.042**	**0.073**
N	419	419	419	419	419

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Agriculture's impact on growth is the most significant, but *Forestry* demonstrates the highest coefficient. However, the effect *Forestry* has on growth disappears when included together with the other industries in model 5. *Mining* is insignificant even when included by itself, which is somewhat surprising considering the significant and negative impact minerals usually have on growth. *Fishing* is positive, but insignificance. In model 5, only *Agriculture* significantly affects growth. Its effect is negative. The explanation to Agriculture's significant impact on growth is likely found in the descriptive statistics in Table A.1 in Appendix A. The mean of *Agriculture* is far above the three other industries, indicating that *Agriculture* in general has a bigger impact on state economy. The same tendency can be read from the minimum and maximum values, which indicates that the other sectors constitutes a much smaller share of SGDP.

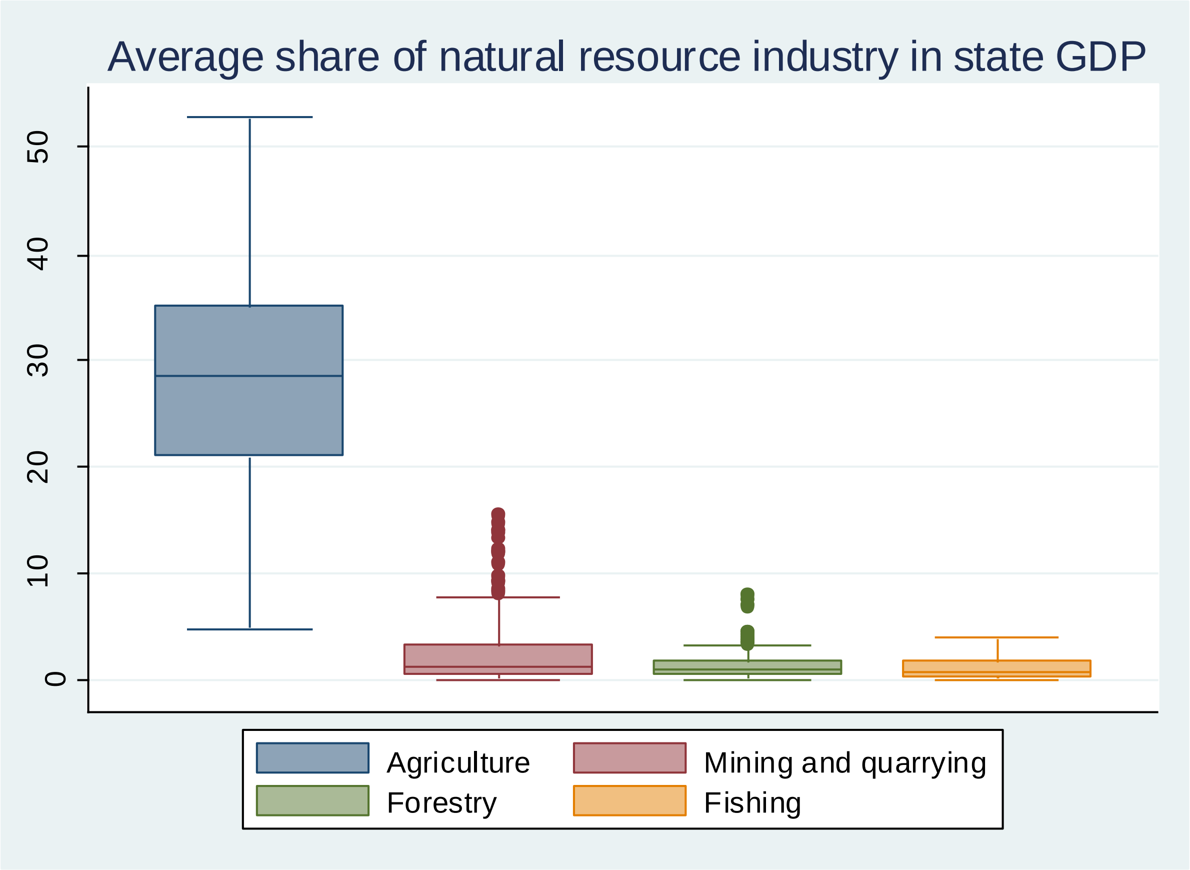

Figure 4.1 Average share of natural resource industry in state GDP

The tendency is confirmed by Figure 4.1, where *Agriculture* clearly is the most important industry within the primary sector. The other three industries have too small an impact to be expected to be significant.

As agriculture constitutes such a big share of the primary sector, it is interesting to see whether *Agriculture* is a better measure for natural resource dependence than *Primary*. *Agriculture* will exclude the three other industries entirely, but on the other hand, it might turn out to be a better fit of what is tried to measure. In Table 4.3, the same regression analyses are run as in Table 4.1, but with *Agriculture* instead of *Primary* as the main independent variable.

Table 4.3 Agricultures impact on economic growth

Variables	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Constant	10.120***	9.770***	8.945***	9.062***	17.750***	15.070***
	(-2.765)	(-2.735)	(-0.909)	(-0.900)	(-4.115)	(-2.605)
Agriculture L1	-0.131***	-0.129***	-0.085***	-0.089***	-0.114***	-0.092***
	(-0.027)	(-0.0289)	(-0.0257)	(-0.0257)	(-0.0281)	(-0.026)
DummyCoastA	-0.056	0.036			-0.519	
	(-0.652)	(-0.648)			(-0.498)	
DummyCoastB	-0.732	-0.561			-0.941	
	(-0.839)	(-0.850)			(-0.833)	
RoadDensity Ln L1	-0.128	-0.087			-0.859**	-0.710**
	(-0.311)	(-0.313)			(-0.429)	(-0.331)
UrbanPopGrowth L1		0.059		0.066	0.092**	0.105**
		(-0.046)		(-0.055)	(-0.045)	(-0.051)
LliteracyDisparity L1			0.014	0.017	-0.025	
			(-0.047)	(-0.048)	(-0.051)	
IMR L1			-0.023**	-0.023**	-0.038**	-0.036***
			(-0.011)	(-0.011)	(-0.015)	(-0.011)
DummyGrowth	5.218***	5.184***	5.270***	5.235***	5.310***	5.288***
	(-1.795)	(-1.812)	(-1.767)	(-1.788)	(-1.801)	(-1.796)
R^2	**0.070**	**0.070**	**0.071**	**0.072**	**0.077**	**0.075**
N	419	403	419	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

A comparison of the R^2 from Table 4.1 and Table 4.3 indicate that the models with *Primary* are a slightly better fit than the models with *Agriculture*. This effect is however very low, with the previous only being 0.2 percent better than the latter. On the other hand, it appears that using *Agriculture* yields more significant coefficients. However, using *Agriculture* excludes the three other primary industries from the analyses. Although these industries are small, the aim of this thesis is to investigate the relationship between economic growth and natural resources, not between economic growth and agriculture. If there had been a clear difference in the results, a change from *Primary* to *Agriculture* could have been justified.

4.3 The secondary and tertiary sector

To investigate hypothesis 4a, regression analyses equal to those in Table 4.1 are conducted for the *Secondary* and *Tertiary* sectors. Tables B.8 and B.9 with regressions of one variable at the time are attached in Chapter B.3 in Appendix B. Table B.8 demonstrates that *Secondary* is not as robust to the inclusion of control variables as *Primary*, as both the R^2 and the strength of the coefficients are lower. Table B.9 indicate that *Tertiary* is as robust as *Primary* by the same measures.

Table 4.4 Secondary sector's impact on economic growth

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Constant	0.660	0.402	6.336***	6.005***	12.240***	10.170***
	(-2.187)	(-1.980)	(-1.107)	(-0.915)	(-3.043)	(-1.876)
Secondary L1	0.077	0.085*	0.062**	0.076***	0.075*	0.062***
	(-0.049)	(-0.045)	(-0.029)	(-0.022)	(-0.037)	(-0.024)
DummyCoastA	0.738	0.781			-0.069	
	(-0.623)	(-0.604)			(-0.574)	
DummyCoastB	0.448	0.529			-0.030	
	(-0.748)	(-0.735)			(-0.676)	
RoadDensity ln L1	0.370	0.385			-0.709**	-0.473*
	(-0.290)	(-0.269)			(-0.304)	(-0.242)
UrbanPopGrowth L1		0.094		0.095	0.131**	0.125**
		(-0.058)		(-0.067)	(-0.061)	(-0.063)
LiteracyDisparity L1			-0.019	-0.022	-0.060	
			(-0.044)	(-0.041)	(-0.047)	
IMR L1			-0.033***	-0.031***	-0.041***	-0.048***
			(-0.006)	(-0.004)	(-0.008)	(-0.006)
DummyGrowth	5.213***	5.155***	5.307***	5.257***	5.316***	5.301***
	(-1.765)	(-1.786)	(-1.747)	(-1.771)	(-1.776)	(-1.778)
R^2	**0.053**	**0.056**	**0.062**	**0.064**	**0.066**	**0.065**
N	419	403	419	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

The secondary sector continues to demonstrate less significant results in Table 4.4. None of the variables except from *DummyGrowth* are significant in model 1. Model 6 is also the best model for *Secondary* with the highest R^2 second to model 5, which has more variables included. The same three control variables that were significant with *Primary* are also significant with *Secondary* in model 6.

Table 4.5 Tertiary sector's impact on economic growth

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Constant	1.627	1.800	1.054	1.594	9.084*	8.410**
	(-1.624)	(-1.561)	(-1.561)	(-1.506)	(-4.887)	(-3.750)
Tertiary L1	0.152***	0.138***	0.118***	0.110***	0.117***	0.118***
	(-0.018)	(-0.021)	(-0.023)	(-0.025)	(-0.023)	(-0.024)
DummyCoastA	0.286	0.444			-0.004	
	(-0.684)	(-0.645)			(-0.539)	
DummyCoastB	0.110	0.352			-0.007	
	(-0.558)	(-0.560)			(-0.601)	
RoadDensity ln L1	-0.401	-0.350			-0.913*	-0.850**
	(-0.280)	(-0.272)			(-0.503)	(-0.400)
UrbanPopGrowth L1		0.102**		0.0795*	0.125***	0.124***
		(-0.052)		(-0.042)	(-0.045)	(-0.046)
LiteracyDisparity L1			0.032	0.033	-0.015	
			(-0.044)	(-0.048)	(-0.054)	
IMR L1			-0.023**	-0.023*	-0.033*	-0.034***
			(-0.010)	(-0.012)	(-0.017)	(-0.013)
DummyGrowth	5.468***	5.418***	5.443***	5.406***	5.486***	5.481***
	(-1.663)	(-1.686)	(-1.669)	(-1.691)	(-1.695)	(-1.689)
R^2	**0.073**	**0.071**	**0.073**	**0.072**	**0.075**	**0.075**
N	419	403	419	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table 4.5 indicates that the tertiary sector is more robust to the inclusion of control variables than the secondary sector. *Tertiary* is significant at the 0.01 level throughout the table, and the R^2 indicates that the tertiary sector explains almost as much as the primary sector does. Overall, Table 4.5 has the highest number of significant variables so far. *UrbanPopulationGrowth* is significant throughout, and the same is *IMR*. Model 6 is also the best fit for the tertiary sector. The coefficients are the most significant and indicate the strongest effect on economic growth of all sectors.

4.4 Inclusion of interaction terms

Hypotheses 3, 4b and H5a and b presuppose that the effect of a variable is not equal for all the units. Hypothesis 3 assumes that resource abundant states will have poorer scores on human capital and infrastructure, implicitly assuming that states more dependent on secondary and tertiary sector have higher scores. The effect that human capital and infrastructure have on economic growth will therefore not be the same for all states. Hypothesis 4b assumes that states with a higher share of the tertiary sector in SGDP will also have a higher teledensity. The access to freight is also assumed to affect the sector composition in SGDP, and hypothesis 5a assumes that states with a higher share of secondary sector will have better possibilities for freight. States with better infrastructure and coastline are therefore expected to have a higher share of secondary sector in SGDP. The same is the case for education. As the secondary sector is dependent on skilled labor, hypothesis 5b assumes that better scores on human capital will have a more positive impact on economic growth in states with a higher secondary sector. To test for these hypotheses, we need to include interaction terms. The interaction terms are introduced to the model with best fit so far, model 6. This will be done for *Primary*, *Secondary* and *Tertiary* separately.

Table 4.6 Primary with interaction terms

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	21.70*** (-5.991)	15.00*** (-2.438)	14.27*** (-2.670)	14.25*** (-3.753)
Primary L1	-0.347* (-0.209)	-0.116*** (-0.025)	-0.104*** (-0.031)	-0.032 (-0.064)
RoadDensity ln L1	-1.662** (-0.839)	-0.621** (-0.308)	-0.599* (-0.328)	-0.803** (-0.352)
UrbanPopGrowth L1	0.126** (-0.060)	0.109** (-0.054)	0.700*** (-0.227)	0.135** (-0.058)
IMR L1	-0.024** (-0.011)	-0.031 (-0.024)	-0.022* (-0.012)	-0.021* (-0.013)
LiteracyDisparity L1				0.072 (-0.121)
Primary*RoadDensity	0.036 (-0.029)			
Primary*IMR		0.000 (-0.001)		
Primary*UrbanGrowth			-0.022*** (-0.009)	
Primary*Literacy				-0.004 (-0.004)
DummyGrowth	5.315*** (-1.773)	5.329*** (-1.775)	5.259*** (-1.761)	5.356*** (-1.758)
R^2	**0.078**	**0.077**	**0.078**	**0.079**
N	403	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table 4.6 demonstrates the interaction terms included with *Primary*. *Primary* is not as robust to the inclusion of interaction terms as it is for control variables. However, it is insignificant when included with interaction terms that are insignificant as well. The only significant interaction term is *Primary***UrbanPopulationGrowth*. It is significant at the 0.01 level, as are the coefficient of each of the variables included in the interaction term.

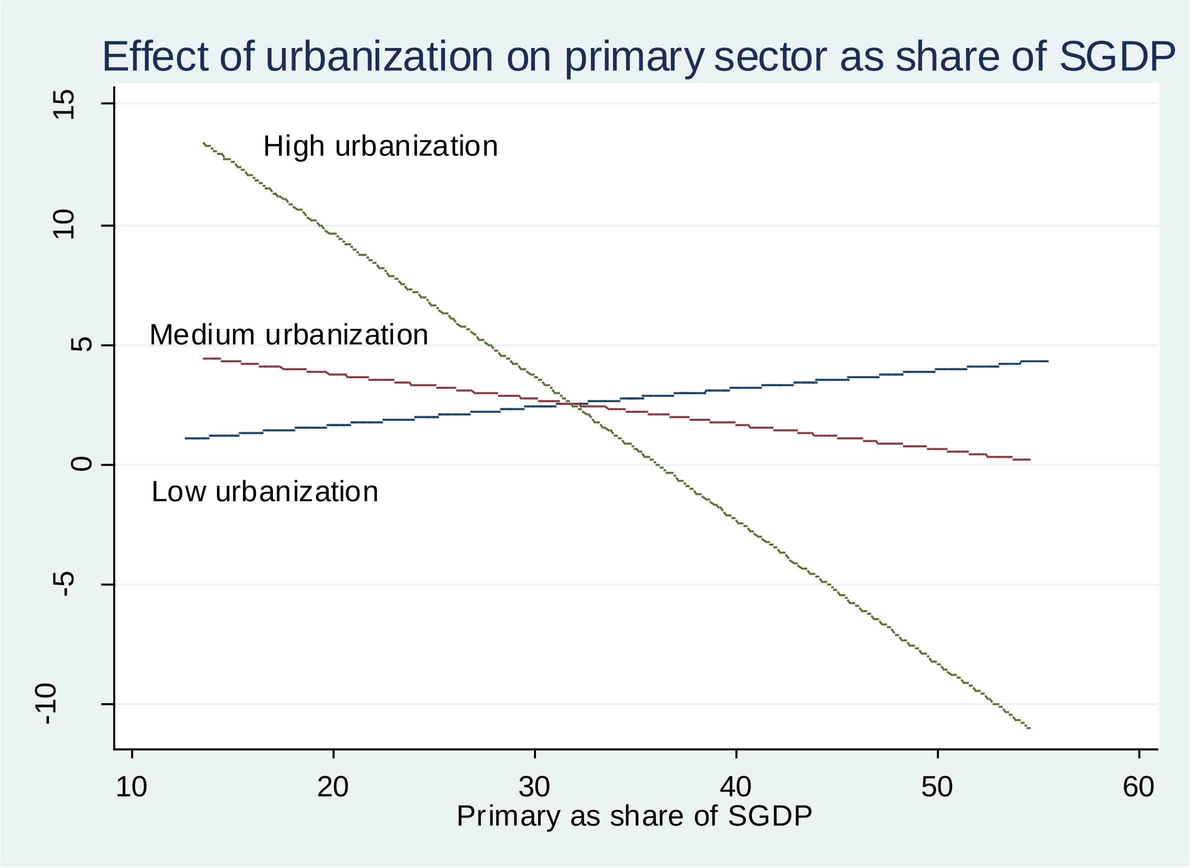

Figure 4.2 Effect of urbanization on primary sector as share of SGDP

Figure 4.2 is a graphical presentation of the interaction term. It clearly demonstrates that the effect of growth in urban population varies with the share of primary sector in SGDP. A high urbanization rate does not lead to higher economic growth when the share of the primary sector in SGDP is also high. Although growth in urban population has a positive effect on economic growth, this positive effect diminishes rapidly as the share of primary sector in SGDP increases. A low urbanization rate affects economic growth positively as the share of primary sector in SGDP increases.

Table 4.7 Secondary and interaction terms

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	9.426*** (-2.062)	5.253 (-4.334)	7.543*** (-2.753)	10.18*** (-1.896)
Secondary L1	0.078 (-0.084)	0.273 (-0.169)	0.204*** (-0.072)	0.0613** (-0.024)
UrbanPopGrowth L1	0.126** (-0.058)	0.129** (-0.063)	0.111* (-0.067)	-0.0342 (-0.233)
RoadDensity ln L1	-0.445** (-0.217)	0.264 (-0.645)	-0.584*** (-0.218)	-0.478* (-0.246)
IMR L1	-0.046*** (-0.008)	-0.050*** (-0.007)	0.008 (-0.027)	-0.048*** (-0.006)
DummyCoastline	0.884 (-1.917)			
Secondary*Coastline	-0.031 (-0.091)			
Secondary*RoadDensity		-0.031 (-0.024)		
Secondary*IMR			-0.002** (-0.001)	
Secondary*UrbanGrowth				0.007 (-0.011)
DummyGrowth	5.287*** (-1.778)	5.294*** (-1.784)	5.335*** (-1.772)	5.294*** (-1.778)
R^2	**0.065**	**0.065**	**0.066**	**0.065**
N	403	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

In Table 4.7, interaction terms are included with *Secondary*. The interaction term *Secondary**IMR* is significant at the 0.05 level. The coefficient of *Secondary* is significant, but that of *IMR* is not. As the coefficient is very low, at 0.008, it is probably a too small effect to have any significant impact on economic growth. The interaction term is difficult to read from the coefficients alone, but the graphical demonstration in Figure 4.3 gives a better indication of the effect.

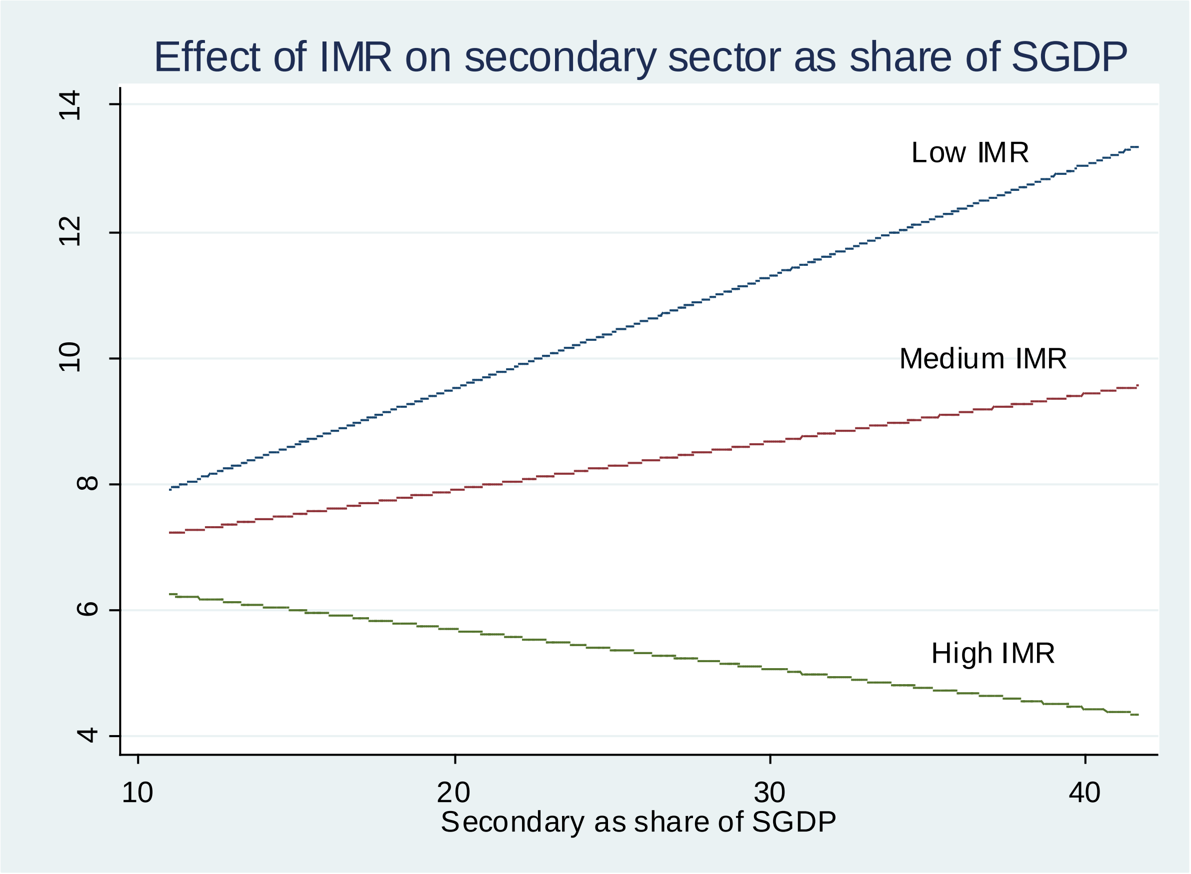

Figure 4.3 Effect of IMR on secondary sector as share of SGDP

This demonstrates that low infant mortality rates become more positive for the economic growth as the share of the secondary sector in SGDP increases. The same effect is demonstrated for medium infant mortality rates. It affects economic growth positively when the share of secondary sector increases. States with higher infant mortality rates have a lower economic growth, despite a high share of the secondary sector in SGDP. Figure 4.3 demonstrates how the effect of IMR varies with dependence on the secondary sector.

Table 4.8 Tertiary and interaction terms

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	1.722 (-6.681)	8.439** (-3.722)	9.336** (-4.213)	7.633* (-4.042)
Tertiary L1	0.241*** (-0.077)	0.115*** (-0.024)	0.091** (-0.046)	0.133*** (-0.026)
UrbanPopGrowth L1	0.149*** (-0.043)	-0.206 (-0.558)	0.124*** (-0.048)	0.130*** (-0.047)
RoadDensity ln L1	-0.510 (-0.380)	-0.841** (-0.408)	-0.819** (-0.391)	-0.818* (-0.419)
IMR L1	-0.053** (-0.021)	-0.034*** (-0.013)	-0.051 (-0.031)	-0.034** (-0.017)
Teledensity ln L1	4.658* (-2.601)			
DummyCoastline				1.624 (-1.379)
Tertiary*Teledensity	-0.101** (-0.051)			
Tertiary*UrbanGrowth		0.006 (-0.010)		
Tertiary*IMR			0.000 (-0.001)	
Tertiary*Coastline				-0.038 (-0.036)
DummyGrowth	6.520*** (-1.592)	5.467*** (-1.683)	5.488*** (-1.694)	5.510*** (-1.713)
R^2	**0.123**	**0.075**	**0.076**	**0.076**
N	225	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table 4.8 further supports the findings from Table 4.5 that the tertiary sector has a robust impact on economic growth. *Tertiary* remains significant throughout the models, while *Primary* and *Secondary* does not. The interaction term *Tertiary*Teledensity* is significant at the 0.05 level, and it is negative. However, the coefficient of Teledensity is highly positive. This indicates that Teledensity initially has a positive impact on economic growth, but this diminishes as the share of the tertiary sector in SGDP increases. This is demonstrated by Figure 4.4:

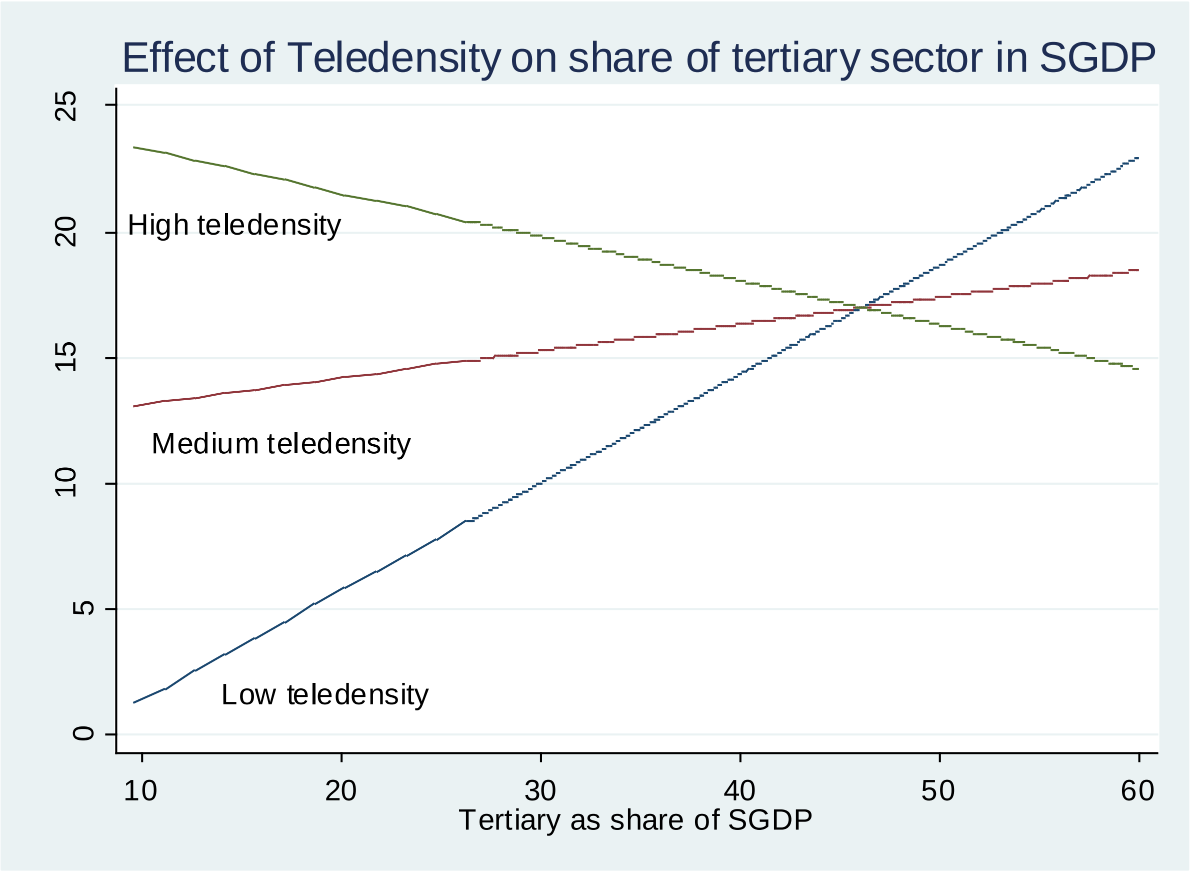

Figure 4.4 Effect of Teledensity on share of tertiary sector in SGDP

It becomes evident that states with a high teledensity have a higher level of economic growth when the share of tertiary sector in SGDP is low. Low teledensity have a positive impact on economic growth when the tertiary sector is high. This is opposite of what was expected, as higher teledensity was expected to have a purely positive impact on the tertiary sector. As the tertiary sector is believed to increase economic growth, this negative impact is puzzling.

5 Analyses

This chapter is based on the regression analyses presented in the previous chapter. The discussion will revolve around the hypotheses presented in Chapter 2.5.

5.1 Natural resources' impact on economic growth

The main hypothesis in this study, hypothesis 1, is that natural resource dependence is expected to have a negative impact on economic growth. This study seeks to investigate whether the divergent growth rates among the Indian states can be explained by the resource curse literature. The initial findings indicate that this is the case. *Primary* has a negative coefficient in model 6 in Table 4.1 with -0.104, demonstrating a negative relationship between natural resource dependence and economic growth. States that are more dependent on natural resources is thereby predicted to have a lower economic growth.

The coefficient of *Primary* indicates that for every one percent increase of the primary sector as share of SGDP, the economic growth declines with 0.1 percent. Similarly, states that are one percent less dependent on the primary sector will experience an economic growth of 0.1 percent. The coefficient of -0.104 might seem small, but if the share of natural resources in SGDP declines with ten percent instead of one, the economic growth will rise with 1 percent.

By applying this to one of the Indian states, the impact of a decrease in natural resource dependence can more easily be understood: Table 2.1 in Chapter 2.4 shows the states average growth in SGDP. The lowest growth rate belongs to Madhya Pradesh, with a 3.7 percent average annual growth in SGDP from 1980 to 2009. Knowing that Madhya Pradesh had a total SGDP of 14 249 993 million rupees in 2007-2008 (Indiastat no date), a one percent increase in SGDP equals an increased income of 142 499 rupees. Further, Table 2.2 in Chapter 2.4 demonstrates that the primary sector for the same state and time period constituted of an average of 38.4 percent of the SGDP. A decline of ten percent, to 28.4 percent, is not unrealistic, as many Indian states have achieved numbers lower than this. Thus, a ten percent decline in natural resource dependence and a predicted rise of economic growth of 1 percent should be a realistic scenario. Even though *Primary* initially seems to demonstrate a low effect on growth, its real impact is not as minute as it seems at first glance when it is applied to a real-case scenario.

Hypothesis 2 states that the effects of human capital and infrastructure have a significant impact on economic growth. *RoadDensity* works as a proxy for infrastructure, due to lack of data on other infrastructure more detrimental to freight, such as railroads. The effect of *RoadDensity* in model 6 is somewhat surprising. According to the literature, access to freight is vital to economic growth, and increased possibilities for freight should therefore have a positive impact on economic growth. However, *RoadDensity* demonstrates a negative effect on growth in Table 4.1. The negative effect seems quite strong, with a coefficient of -0.601. However, the variable is power transformed, so the interpretation of its coefficient is not straight forward. A one percent increase in the coefficient of *RoadDensity* is associated with 1 /100 times the coefficient change in the dependent variable. To correctly interpret the impact *RoadDensity* has on economic growth, Equation 6 is used:

$$0.01*\beta_{RoadDensity} = 0.01*(-0.601) = -0.006. \qquad \text{Eq. (9)}$$

A one percent change in the road density is associated with a decrease in the economic growth by -0.006 percent. This also indicates a rather small effect, but it is important to remember how the variable is constructed. It measures the road length in kms per 100 km^2 of area, and a one percent increase in road length is not necessarily a huge amount of new road. The impact will naturally increase when more roads are built. However, the variable's negative coefficient is still puzzling, even though its significance is not. The descriptive statistics in Table A.1 in Appendix A demonstrates a wide spread in the values on *RoadDensity* before it is power transformed, giving reason to believe that some extreme values might influence it. Even though the variable used in the regression has been transformed to avoid problems with skewness, Figure 5.1 demonstrates that the problem with extreme values is not completely eliminated.

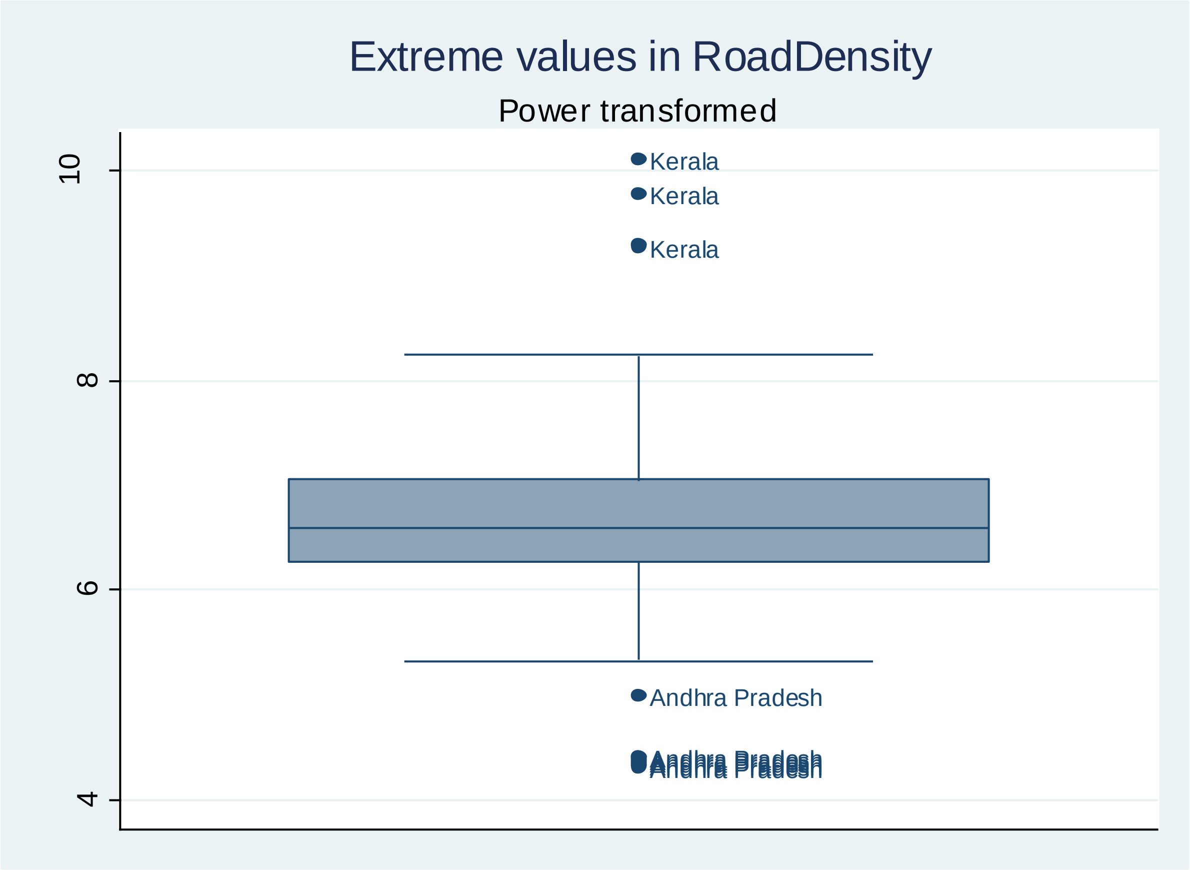

Figure 5.1 Extreme values in RoadDensity

Both Kerala and Andhra Pradesh have some extreme values on the variable. When Kerala and Andhra Pradesh are taken out of the regression, the coefficient of *RoadDensity* becomes more negative, but also insignificant. However, when Andhra Pradesh is kept in the regression and only Kerala is taken out, the variable is significant at the 0.1 level. Regression analyses with without Kerala, and without Kerala and Andhra Pradesh, are attached in Table B.13 in Appendix B.

The infant mortality rate is used as a proxy for the effect human capital has on economic growth. Model 6 in Table 4.1 demonstrates that the impact of *IMR* on economic growth is negative and significant, as expected. Higher values of infant mortality rates leads to lower economic growth. The coefficient of -0.024 indicates that an increase of one value on *IMR* decreases the economic growth by 0.02 percent. This means that every infant diseased within the first year of living for each 1000 infants born reduces the economic growth by 0.02 percent. The impact of an increase or decrease in *IMR* by one does not have a very big impact on economic growth. However, there are huge differences in the infant mortality rate among the states. Again, an example can be given to illustrate how change in the values on *IMR* can have an impact on growth: Kerala has a mean of *IMR* of 22 for the entire period, while Madhya Pradesh has a mean of 104. This indicates that 82 more infants die per 1000 infants born in Madhya Pradesh than in Kerala. With an infant mortality rate of 104, the economic growth will be reduced by

$$104*\beta_{IMR} = 104*(-0.024) = -2.5 \text{ percent} \qquad \text{Eq. (10)}$$

If the infant mortality rate is 22, the economic growth will be reduced by

$$22*\beta_{IMR} = 22*(-0.024) = -0.53 \text{ percent} \qquad \text{Eq. (11)}$$

If Madhya Pradesh manages to reduce its infant mortality rates to the same level of Kerala, its economic growth is predicted to rise by approximately two percent.

IMR is not significant in the initial models in Table 4.1, but it becomes significant in model 6, where also *UrbanPopulationGrowth* is included. The assumption that *UrbanPopulationGrowth* would render *IMR* insignificant thus seems to be wrong. As both variables were included together in model 4 without any of them being significant, the exclusion of *LiteracyDisparity* from model 4 to model 6 must be part of the explanation. Even though *LiteracyDisparity* and *IMR* do not have a problem with collinearity, they still measure enough of the same to render each other insignificant. Both are a measure of gender inequality in the society, and thus also a measure of development. As *IMR* measures infant mortality rates, which is expected to be determined partly of education, *IMR* might account for the effect of *LiteracyDisparity* once this is taken out. Neither *LiteracyDisparity* nor *IMR* measure enough of the variance in economic growth by themselves. Each variable explain too little of the variance to have a significant impact on economic growth. When *IMR* is included alone, it adopts enough of *LiteracyDisparity*'s explanatory power to account for enough of the variance in growth to become significant by its own.

It should however be noted that *IMR* becomes insignificant when Kerala is taken out[4]. This indicates that the values Kerala has on *IMR* are high enough to make it significant. As already mentioned, this will be a topic of interest in the robust tests later on. For now, Kerala remains in the analyses, and *IMR* is significant.

---

[4] Again, see model 1 in Table B.13 in Appendix B.

The growth in state-wise urban population is a proxy for the urbanization rate. Urbanization is not explicitly part of any hypothesis in this thesis, but based on previous studies it is expected to have an effect on economic growth. Sachs et al. (2002), among others, have argued that economic growth for the most part takes place within urban areas. Sachs et al. (2002:10) have also found human capital indicators to be rendered insignificant once measures for urbanization are accounted for. If human capital and infrastructure lose their significance by the inclusion of urbanization, it is because urbanization accounts for their effect. Thus, urbanization is of interest in explaining hypothesis 2, that human capital and infrastructure are expected to have a significant impact on growth.

In model 6 in Table 4.1, *IMR* is still significant, even after including *UrbanPopulationGrowth*. *UrbanPopulationGrowth* is positive and significant, indicating that growth in urban population leads to higher economic growth. The variable is also positive and significant even after Kerala is excluded from the sample.

A one percent increase in urban population will increase the economic growth by 0.1 percent. This effect seems small. Looking at the descriptive statistics, the spread of the variable is not very big. The mean is low at only 0.01. This indicates that the general growth of urban population is not high enough for *UrbanPopulationGrowth* to have a big impact on the economic growth. However, it still has a significant positive impact on economic growth. A reason for the low coefficient might be explained by *IMR*, which accounts for the gender differences in education and access to health care services. A lower value on *IMR* along with a higher value on *UrbanPopulationGrowth* indicates a higher economic growth. Thus, they both contribute to higher economic growth in the way that was expected.

The R^2 is a measure of how much of the change in Y can be explained by the model. In this case, it demonstrates how much of the economic growth is explained by the control variables *RoadDensity*, *UrbanPopulationGrowth* and *IMR*. To investigate hypothesis 3, that resource abundant states are expected to have poorer scores on human capital and infrastructure, a set of interaction terms are included in Table 4.6. Interaction terms are included when there is reason to assume that the effect of one X on Y is dependent on the effect of another X on Y. For example, hypothesis 3 assumes that states with higher share of primary resources as share of SGDP will have a poorer score on human resources or infrastructure.

The interaction term *Primary*UrbanPopulationGrowth* in model 3 is significant at a 0.01 level. As demonstrated in Figure 4.1, a high urbanization rate has a positive effect on growth, as long as the share of primary sector in SGDP is low. Higher dependence on natural resources drastically reduces the positive effect of high urbanization rates. Similarly, low urbanization rates initially have a negative impact on economic growth, as expected. However, when the primary sector accounts for more than approximately 35 percent of SGDP, low urbanization rates has a positive impact on economic growth.

This result runs contrary to what was expected. It was assumed that high urbanization rates are associated with better infrastructure and lower infant mortality rates. As *Primary* has demonstrated a robust and negative impact on economic growth, it might be the case that the negative impact of natural resource dependence is higher than the positive impact of urbanization. As the share of primary sector in SGDP increases, the urbanization rates become less detrimental to economic growth.

5.2 The secondary and tertiary sector's impact on economic growth

Besides investigating whether the divergent growth in India can be explained by the resource curse, this thesis also seeks to investigate which of the other two sectors explain more of the economic growth. Some studies on the divergent growth rates among the Indian states emphasize the need to change from agriculture towards manufacturing to create sustainable economic growth (Kurian 2000:541; Sachs et al. 2002:8). As demonstrated by Figure 4.1, agriculture is by far the most important industry within the primary sector. A shift from agriculture is virtually the same as a shift from the primary industry. Although manufacturing and the secondary sector are frequently highlighted as detrimental to sustained economic growth, others have found the tertiary sector to be more important to the economic development in India. The tertiary sector usually accelerates after the secondary sector is well established and the economy has attained a fairly high level of development (Bhattacharya and Sakthivel 2004:13). These features of the Indian development are therefore unusual.

As this pattern of sector composition is somewhat remarkable, hypothesis 4a seeks to investigate whether the secondary or the tertiary sector explain more of the economic growth in India. Based on literature on regional economic growth in India, it is expected that the tertiary sector have a bigger impact on economic growth than the secondary sector. Regression analyses to investigate hypothesis 4a are equal to those done on the primary sector in Table 4.1. The most noticeable features of Table 4.4 and Table 4.5 are that *Tertiary* stay highly significant throughout, while *Secondary* does not. The numbers of significant variables are also higher for *Tertiary* than for *Secondary*. The coefficients of the significant variables are also stronger. In model 6, the coefficient for *Tertiary* is almost twice as high as that for *Secondary*. This indicates that if both the tertiary and secondary sector increases by one percent as a share of SGDP, the tertiary sector will render a higher economic growth rate than the secondary sector. Even though the growth as a share of SGDP is equally high, the return from the tertiary sector will be higher. As *Tertiary* has a coefficient of 0.118, a ten percent increase of the tertiary sector as share of SGDP would give a predicted economic growth of 1.18 percent. *Secondary* has a coefficient of 0.062, and a ten percent increase of the secondary sector in SGDP would give a predicted economic growth of 0.62 percent. The impact of having a bigger tertiary sector on economic growth is evident.

The R^2 is also notably higher for model 6 including *Tertiary* than *Secondary*. This further confirms that the tertiary sector has a bigger impact on economic growth in Indian states than the secondary sector. However, it should be noted that the explanatory power of both models are still low, with 6.5 and 7.5 percent of the economic growth explained by the variables included. It cannot be ruled out that the R^2 for the secondary sector would be higher than that for the tertiary sector by the inclusion of more control variables. The same is the case for the coefficients. Including more control variables could alter the coefficients, giving the secondary sector a higher impact on economic growth than the tertiary sector. These findings are therefore not conclusive.

The control variables describe the same as they did in Table 4.1, but the size of the coefficients varies. *RoadDensity* has a higher negative impact on growth when included with *Tertiary* than with *Primary*. As it is also more significant, it might indicate that *Primary* measure some of the effect that is also captured by *RoadDensity*. As *RoadDensity* is both stronger and more significant when included with *Tertiary*, *Tertiary* measures less of the same effect as *RoadDensity*. Thus, *RoadDensity* becomes stronger and more significant. The same can be assumed to be the case for *UrbanPopulationGrowth*, which also demonstrates stronger and more significant coefficients when included with *Tertiary*.

IMR demonstrates the strongest impact on economic growth when included with *Secondary* in Table 4.4. At the same time, the impact of *RoadDensity* is lower. As mentioned, this could indicate that some of the effect of *RoadDensity* also is covered by *Secondary*, thus giving it a smaller explanatory power. As *RoadDensity* explains less, it allows *IMR* to explain more.

Hypotheses 4b, 5a and 5b assume that the values on some of the control variables vary between the units. To test these hypotheses, interaction terms are included. Hypothesis 4b assumes that states that are more dependent on the tertiary sector have higher teledensity. In Table 4.8, the interaction term *Tertiary*Teledensity* is significant on the 0.05 level. The coefficients of both variables are significant and positive, indicating that both have a positive impact on economic growth. However, the interaction term's coefficient is negative, indicating that this relationship decreases as the values becomes higher. Figure 4.3, displaying the interaction term, demonstrates this relationship. High teledensity has a higher positive impact on growth when the dependence on tertiary sector is lower. As the tertiary sector in SGDP increases, teledensity has a negative impact on growth. Similarly, low scores on teledensity have a negative impact on economic growth as the share of tertiary sector is low. But when the share of the tertiary sector increases, low teledensity affects economic growth positively. This is an unexpected finding, and not easily explained. One explanation might be that there are other industries within the tertiary sector that have a higher positive impact on economic growth than the IT industry. If this is the case, teledensity is not necessarily a good measure for which conditions are detrimental to growth in the tertiary industry. The measures for coastline are not significant either, making the tourist sector a poor fit for explaining growth in the tertiary industry.

According to hypothesis 5a, good possibilities for freight will have a positive impact on economic growth in states with a higher share of secondary sector in SGDP. Hypothesis 5b further assumes that states with higher share of secondary sector in SGDP will have a better score on human capital. Interaction terms to test for both hypotheses are included in Table 4.7. As demonstrated, only the interaction term *Secondary*IMR* in model 3 is significant. As mentioned, the coefficient of *IMR* is insignificant. A coefficient of 0.008 is likely too small to have any significant impact on economic growth. However, the coefficients of the interaction term and *Secondary* are both significant. The coefficient of *Secondary* is positive, meaning that a high share of secondary sector in SGDP has a positive impact on growth. As the interaction term is negative, this indicates that for every value *IMR* increases, the positive effect of having a big secondary sector diminishes. This is further confirmed by Figure 4.2. As expected, low infant mortality rates increase the economic growth, and this positive effect gets higher as the share of secondary sector in SGDP increases. States with a high infant mortality rate have a lower economic growth, and this negative effect increases when the state is more dependent in the secondary sector. As *IMR* is a proxy for both education and fertility rates, this is in line with hypothesis 5b that human capital will have a more positive impact on economic growth in states with a higher secondary sector.

Table 5.1 List of which sectors each state is more dependent on, whether they have an economic growth above average, and whether or not they have a coastline

State	Coastline	High Primary	High Secondary	High Tertiary	High Growth
Andhra Pradesh	X			X	X
Assam		X			
Bihar		X		X	
Goa	X		X	X	X
Gujarat	X		X		X
Haryana		X	X		X
Karnataka	X		X	X	X
Kerala	X		X	X	X
Madhya Pradesh		X			
Maharashtra	X		X	X	X
Punjab		X			
Rajasthan		X	X		X
Tamil Nadu	X		X		
Uttar Pradesh		X			
West Bengal	X			X	

All states with a coastline are highlighted in blue

Although none of the coastline dummies are significant in any of the analyses, Table 5.1 demonstrates a striking relationship between the states with a coastline and those without. None of the states with the highest share of primary sector in SGDP have a coastline. On the other hand, almost all the states with a high share of secondary and/or tertiary sector in SGDP have a coastline. Among the eight states with the highest average growth rate for the period 1980 to 2009, only two do not have a coastline. Similarly, only two of the states with a coastline do not rank among the top eight states when it comes to economic growth.

Four of the states that are most dependent on the primary sector do not have a high share in any of the other sectors. It seems like the states with the highest economic growth are the ones who have above average share of two sectors in SGDP. This indicates that a specialized sector yields higher economic growth rates. Among the eight states with highest economic growth, only two have above average share in one sector. These are Andhra Pradesh, most dependent on the tertiary sector, and Gujarat, most dependent on the secondary sector.

Table 5.1 does not yield any exact numbers, but it does demonstrate an evident trend among the Indian states. States with access to coastline have a) higher economic growth rates and b) higher share of secondary and tertiary sector in SGDP. It is also striking that *all* of the states that are most dependent on natural resources reside inland. By a retrospective glance at Chapter 2.4.2, the prediction that Assam, Uttar Pradesh and Bihar would be among the states with lowest economic growth still hold in Table 5.1. The lack of coastline and below average share of secondary and tertiary sectors in SGDP seems to be disadvantages to economic growth. Of the three states expected to have the lowest economic growth, Assam and Uttar Pradesh both have a share of the primary sector above average. They also have a share of the secondary and tertiary sectors below average, and none have a coastal access. The initial expectations still hold.

6 Discussion

This thesis is mostly based on the comprehensive work of Gylfason (2001), Sachs and Warner (1995; 2001) and Papyrakis and Gerlagh (2003; 2004). Through their work, a general consensus of what causes the resource curse arises. The aim of this thesis is to examine whether this literature fits the Indian case of divergent growth. Due to large geographical and natural habitat diversity, the occurrences of natural resources are unevenly distributed across the states (Sachs et al. 2002). Previous studies who have found the resource curse to fit the case of divergent economic growth both across the US states and across counties (Freeman 2009; Gerard 2011; James and Aadland 2011; Papyrakis and Gerlagh 2007); this gives reason to assume that similar findings can be found in the Indian case. To explain the occurrence of the resource curse, four transmission channels have been presented. The aim of these is to provide explanations to the causes for and solutions to the paradox of slow growth in resource rich areas. These transmission channels were mapped out in Chapter 2.2, and it is soon time to revisit them. But first, the results presented in this thesis should be scrutinized further.

6.1 “One size fits all?”

The findings from the regression analyses in Table 4.1 and Table 4.6 on the primary sector’s impact on economic growth, indicate that natural resource dependence has a negative impact on economic growth across the Indian states. The results seem robust to the inclusion of other control variables. However, these results should be taken with a pinch of salt, as there are several features with both the data and the results that need commenting on.

The first feature is the low explanation power given by these models. According to the R^2, the best model for the inclusion of *Primary* is model 3 in Table 4.6. Here, the three control variables *RoadDensity*, *UrbanPopulationGrowth* and *IMR* are included, as well as the interaction term *Primary*UrbanGrowth*. Although this model is the best fit with the highest explanatory power, it explains only 7.9 percent of the variation in economic growth across the states. Needless to say, there are several omitted variables expected to explain more than the variables included here. This is not surprising, as the data material in this thesis is scarce.

However, some tendencies can be spotted despite the lack of sound and comprehensive data material.

The inclusions of the control variables seem to make the coefficient of *Primary* less negative. This gives support to the findings of Papyrakis and Gerlagh (2003; 2004; 2007), who found natural resources to impact economic growth positively when all of the transmission channels were accounted for in their analyses. Further, it has also been demonstrated that several resource rich countries have escaped the curse entirely. Gylfason (2001) is among those who point out Norway and Botswana in this context, both who have turned their natural resources into economic riches. The resource curse literature repeatedly points out the paradox that natural resources tend to harm growth, as natural resources should be assets that could be turned into economic growth (Bulte et al. 2005:1029, Gerard 2011:1; James and Aadland 2011:440). However, examples presented in the literature, both of states that have escaped the curse, and those who have fallen prey to it, tend to revolve around extractive resources such as petroleum and minerals. Figure 4.1 demonstrated that in the case of India, the primary sector is comprised almost exclusively of agriculture. Table 4.2 also tested for agriculture as a substitute for the primary sector as a whole, and demonstrated almost similar results. This further indicates that the effect of the primary sector on economic growth is driven almost exclusively by the agricultural industry.

The four transmission channels presented in Chapter 2.2 has been highlighted as the most important features to explain how the resource curse works, and how it can be avoided. As there are no variables measuring the four transmission channels in this thesis, it is important to discuss their fit to the Indian case.

6.1.1 Do the transmission channels fit the Indian case?

Corruption and rent-seeking are more likely to occur in the wake of resource dependence (Leite and Weidmann 1999:5; Sachs and Warner 1995:4, among others). One hypothesis proposed as an explanation is that powerful interest groups emerge following natural resources, seeking to gain access to the easy riches of the natural resources (Sachs and Warner 1995:4). Another hypothesis is that weak institutions promote the bribing of officials and the allocation of resources to less productive areas of the economy (Leite and Weidmann 1999:5).

The natural resources in India consist almost exclusively of agriculture. As there are no big instant riches occurring along with it, feeding frenzies and strong interest groups are not likely to arise in the wake of agriculture.

Mehlum et al. (2006) argue that victims to the resource curse are determined by institutional quality. Countries with strong legal and bureaucratic institutions are more likely to escape the resource curse. Robinson et al. (2006) compliment these findings with a political model of resource extraction. Low quality institutions make it easier and more attractive for politicians to redistribute rents from the natural resources to influence the outcome of elections (Robinson et al. 2006:465-466). The model created by Robinson et al. use *extractive resources* as a point of departure to explain the effect of poor institutions on economic growth. Agricultural activity can hardly be deemed an extractive industry, and this political model of institutional quality seems like a poor fit to the Indian case.

However, weak legal and bureaucratic institutions can hardly be said to harm growth only in the countries with natural resources. Even though India’s natural resources are not as prone to feeding frenzies and bribing as other natural resources, the states can still fall victim to corruption and poor institutions. Corruption and poor institutions can be just as destructive for economic productivity where the occurrences of extractive natural resources are low. Although countries with a higher share of extractive resources might be more prone to corruption and rent-seeking, the effect of weak institutions does not have to be linked to natural resource endowment. Auty (2001:361) finds that domestic economic policy is more important than natural resources in driving economic growth. Robinson et al. (2006) investigate the impact that natural resources, and resource booms in particular, have on economic growth. They find that sound political institutions are crucial to escaping the resource curse. Again, the theory is based on extraction of natural resources, and on resource booms. Robinson et al. (2006:447) find that “the overall impact of resource booms on the economy depends critically on institutions since these determine the extent to which political incentives map into policy outcomes” (Robinson et al. 2006:447).

Agriculture might be expected to be volatile in terms of drought and failing crops, but it is reasonable to assume that it will be less volatile in terms of resource booms. Despite the fact that resource booms are a poor fit to the Indian case of natural resource dependency, corruption and weak institutions can be just as destructive to the economic growth. Corruption might be due to poverty and not natural resources. Weak institutions in poor states might make bribing a feasible option for state officials looking for easy money. As poor states also are the ones more resource dependent, measures of corruption might seem to be caused by natural resources, while it in fact is caused by other factors. Therefore, it cannot be ruled out that corruption and weak institutions play a part in the divergent economic growth across the Indian states.

The second transmission channel presented in the literature, openness and terms of trade, is not of much relevance to explain divergent economic growth at the state level. As previously mentioned, most of the measures to improve openness and terms of trade have to be implemented at the federal level of government. Although it is difficult for states to enhance their openness towards the global market, some measures can be taken to enhance their attractiveness to investments. Both foreign and domestic investors seek to invest where conditions favour investments, such as areas with sound infrastructure and skilled labuor. In this aspect, state governments can seek to invest in factors such as these to promote their investment climate.

This leads to the third transmission channel: investment. This was found by Papyrakis and Gerlagh to have the biggest impact on economic growth of all four transmission channels (2003:11; 2004:189). Sachs and Warner (2001) found that countries with higher resource dependence had lower export in manufactures. They link this to the higher price level for production caused by the higher prices of domestic primary resources, rendering the manufactured goods uncompetitive on the global market (Sach and Warner 2001:834). Papyrakis and Gerlagh further argue that as natural resources generate a continuous income of wealth for the foreseeable future, investment in manufacturing to ensure future welfare seem less important to the government (2007:1020). Also here, the fit to the Indian case seems poor. Considering that the primary sector primarily constitutes of agriculture, the price level of primary resources as input in the manufacture sector is hardly the main reason why it performs below expectations.

The hypothesis that the income from natural resources leads the governments to fail to invest in future welfare as the current and future economic prospect are good, does not seem to be a good explanation either. Agriculture hardly creates such a sound future prospect for economic welfare that investments seem superfluous. Similarly, the Indian government identifies the problems of poverty, inequality and the rural – urban divide as serious societal and economical problems (Planning Commission of India 2006:8-10). Failing to invest due to sound future prospects might be the case in countries where poverty and inequality are not fundamental issues, such as the US. But this does not fit the case of India very well, as the future economic prospects hardly can be deemed as good, especially not due to the income from the primary sector.

The Indian government further points out that infrastructure and education are the biggest obstacles to improve growth in manufacturing. As argued earlier, investment in these factors might be as detrimental to economic growth in India as investment directly in the manufacturing sector seems to be in the US. It can be argued that the results presented here demonstrate this. Both *IMR* and *UrbanPopulationGrowth* demonstrate that improved values on these variables predict higher economic growth. Especially the interaction term *Secondary*IMR* seems to confirm this. Low infant mortality rates are associated with higher share of secondary sector in SGDP, and a higher economic growth. A conclusion to whether this is caused by targeted investments in infrastructure and human capital cannot be drawn. However, considering the divergent growth rates in India along with the positive impact of low infant mortality rates on economic growth, investment in human capital and infrastructure should be considered as well-aimed growth-promoting investments by all states.

In this matter, the case of failed investment is also a not-so-good fit to the Indian case of divergent growth. Poverty, inequality and the rural – urban divide are of serious political concern, and they are recognized as harmful to future economic growth. Investments to tackle these issues and to raise the income from the secondary and tertiary sector are identified as the most important growth-promoting measures that can be done. It is unlikely that the high levels of poverty and inequality, as well as the rural – urban divide are caused by government failure to realize the need for future investments due to the riches produced by natural resources. The interaction term *Secondary*IMR* demonstrates that investment in human capital in fact leads to higher economic returns to the secondary sector.

The fourth and final transmission channel is schooling, or education. Education is widely viewed as an engine for societal development and economic growth. However, Papyrakis and Gerlagh (2003; 2004) find education to have the least impact on economic growth of the four transmission channels. The hypothesis presented by Sachs and Warner is that easy riches lead to sloth (1995:4). Once again, the hypothesis presented in the comprehensive literature fit the Indian case poorly. It is hardly likely that the government fail to invest in education because of the easy riches that occur in the wake of agricultural industry. On the other hand, the argument that the need for skilled labor is lower within the primary sector should fit the Indian case well. Agriculture is not dependent on skilled labor, giving families and residents within areas highly involved with agriculture fewer incentives to send their children to school.

Better education for a higher share of the population is an explicit goal from the Indian government, as better health and education are important pre-conditions for sustained long-term economic growth (Planning Commission of India 2006:2). Education has been demonstrated to have a positive effect on economic growth in several studies. Gylfason points out that education shifts the comparative advantage away from the primary sector and towards the secondary and tertiary sector (2001:856). Guio et al. (1995) find economic growth to be tied to education at a more fundamental level. Education leads to lower fertility rates, lower infant mortality rates, prolonged life, gender equality and empowered women. Education is thus a means in itself (Guio et al. 1995:747), but it is also a means to higher economic growth.

The interaction term *Secondary***IMR* can be used as evidence also here. Lower infant mortality rates are positively associated with growth. Figure 4.3 does not demonstrate the causality of the two variables. But as *IMR* is a proxy for education and human capital, it might indicate that lower infant mortality rates, and thus higher education, in fact contribute to raising the share of secondary sector in SGDP. It does however confirm that low infant mortality rates increases the economic returns to the secondary sector.

Overall, the literature on the resource curse does not pose as a good fit for the Indian case of divergent economic growth. Although some features of the literature fit the Indian conditions quite well, this is not likely due to the dependence on natural resources. Corruption and weak institutions are devastating to economic growth everywhere. Failing investments in education and infrastructure are more likely due to the low level of development rather than to natural resource dependence. The pre-conditions for economic growth are not yet in place in Indian states. It therefore seems that there are other reasons for the divergent economic growth rates in India than the resource dependence.

6.2 Growth promoting factors in India

The variables in this thesis mostly derive from growth-promoting factors identified by the convergence literature in India. Due to the development level of India, education is measured as the gender differences between men and women. However, the infant mortality rates have been demonstrated (both through the literature and through the analyses presented in this study) to be a better measure of development, equality and education. The effect of education, through the proxy of infant mortality rates, has been demonstrated to have a strong and robust impact on growth. The interaction term *Secondary*IMR* further confirms this tendency.

Urbanization has a positive impact on economic growth. The findings in this study indicate that the secondary and tertiary sectors are the main engines of growth. These sectors are mainly located where infrastructure and skilled labour are present. As these factors are mostly located in urban areas, the positive impact of growth in urban population was expected.

Although the dummies for coastline are not significant in any of the analyses, Table 5.1 demonstrates a connection between states with a coastline and high economic growth. It is also interesting that all the states that are most dependent on natural resources are located inland. The economic growth thus seems to happen in states with coastal access. The states with the high shares of two sectors in SGDP are also more likely to have higher economic growth. A specialization of the economy thus seems to have a positive impact on economic growth.

The sector composition of the economies does have a big impact on economic growth. Sachs et al. (2002:9) claim that the secondary sector is likely to be important in states with a coastline, but that the tertiary sector more successfully can be established inland. Table 5.1 demonstrates that growth in the tertiary sector also seems to be located in states with coastlines.

Teledensity was assumed to be an important feature in the growth of the tertiary sector. Surprisingly, it is only significant in the interaction term in Table 4.8 in Chapter 4. It is further surprising that it demonstrates a negative effect on growth as the teledensity becomes higher. Figure 4.3 shows that high teledensity in fact has a negative impact on growth when the tertiary sector constitutes of a higher share of SGDP. An explanation for this puzzling relationship might be that Goa is excluded from the sample. The numbers for the variable are drawn from state-wise and service area-wise data. As Goa is such a small state, it is included in Maharashtra. Goa also has very high economic growth and a very high share of the tertiary sector in SGDP. When Goa is left out of the analyses as missing, the values of Goa's economic growth is left out as well. If observations for Goa had been included, it might have increased changed the impact teledensity has on growth.

6.3 Hypotheses, revisited

Through this thesis, seven hypotheses have been investigated. The main hypothesis is that natural resource dependence is expected to have a negative impact on economic growth. This is the point of departure for this study, trying to apply the resource curse literature to the case of divergent growth rates in India. The results from a set of time-series cross-section regression analyses indicate that natural resources have a negative and significant impact on economic growth. The results seem robust to the inclusion of a small set of control variables. Although the regression analyses seem to fit into the frame of a resource curse quite well, analyses of the literature on the subject theme gives reason to believe that other factors might stand behind these numbers.

First of all, India’s primary sector is highly dependent on agriculture. Agriculture does not lead to many of the negative impacts characteristic for other natural resources such as petroleum and minerals. Second, the characteristic that all states with the highest resource dependence reside inland and lack a coastline is an intriguing trait. Although the dummy variables for coastline are insignificant, the general characteristics of these states would be interesting to investigate further. This is albeit outside the scope of this thesis, but are indeed noteworthy.

Hypothesis 1, that natural resource dependence is expected to have a negative impact on economic growth, cannot be confirmed nor declined. According to the numbers presented by the regression analyses, the hypothesis should be confirmed. However, the resource curse literature indicates otherwise. There are too many features in the literature that do not fit the Indian case of divergent economic growth. Although the numbers indicate a resource curse, this might be a spurious relationship caused by other underlying factors than the resource curse.

The case of endogeneity from Chapter 3.1.2 arises again. Gerard (2011:2-3) argues that in countries with low economic growth, the primary sector will be the last one standing, as it requires less skilled labor than the secondary and tertiary sectors. Thus, the causality will run the other way. Applied to the Indian case, poor states will be more dependent on the primary sector as other developmental features detrimental to economic growth are lacking. Therefore, what seems like a resource curse might in reality be low economic growth due to lacking or failed developmental policies.

Hypothesis 2 states that human capital and infrastructure are expected to have a significant impact on economic growth. Poor infrastructure and human capital are throughout this thesis highlighted as pre-conditions for economic growth, and this is also demonstrated by the regression analyses. IMR tend to be negative and significant throughout. Lower shares of infant mortality rates affect economic growth positively. As infant mortality rates also poses as a proxy for education and gender equality, the effect of education can also be said to have a significant impact on economic growth. The interaction term in Table 4.7 and Figure 4.2 demonstrate this well. Low infant mortality rates have a positive effect on growth, and seem to lead to a higher share of secondary sector in SGDP. As the secondary sector frequently is highlighted as one of the main engines of sustained economic growth, both federal and state government should promote human capital as growth promoting strategies.

Infrastructure demonstrates an unexpected effect on economic growth, through the proxy of road density. This was expected to have a positive impact on economic growth, but is surprisingly negative. There are no obvious reasons as to why this affects economic growth negatively, expect the possibility of extreme values. This will be investigated in Chapter 7. Hypothesis 2 can therefore be confirmed only half way. Human capital has a significant and expected impact on economic growth, but infrastructure display an unexpected and negative relationship.

According to hypothesis 3, resource abundant states are expected to have poorer scores on human capital and infrastructure. The interaction term in model 3 in Table 4.6 demonstrates a significant relationship between primary sector and growth in urban population. High urbanization rate affects growth positively when the share of primary sector in SGDP is low. Urbanization is by Sachs et al. (2002:10) demonstrated to have a positive impact on economic growth. Investments are primarily located in areas where conditions favour industry, and pre-conditions for investments such as infrastructure and skilled labor are mainly located in urban areas. When the shares of primary resources are low, high urbanization clearly affects economic growth positively. However, when the share of primary resources increases, higher urbanization rates are expected to lead to lower economic growth. In states with higher share of secondary and tertiary sector in SGDP, and thus probably lower rates of primary sector, high urbanization rates affect the economic growth positively, and low urbanization rates affect the economic growth negatively.

Hypothesis 3 can neither be confirmed nor discarded by the findings, as none of the interaction terms directly including human capital and infrastructure are significant. However, urbanization affects economic growth positively in states with lower shares of the primary sector in SGDP. In states more dependent in natural resources, high urbanization harms growth. The underlying causes for this are not evident. It might be the case that the urban areas within these states fail to develop a big secondary or tertiary sector to create economic growth.

Table 5.1 lends some support to this. The four states Assam, Madhya Pradesh, Punjab and Uttar Pradesh all lack a coastline, are highly dependent on the primary sector, have a share of the secondary and tertiary sector below average, as well as an average economic growth below average. This indicates that although there might be high urbanization rates in these states, the secondary and tertiary sector in the urban areas in these states are not high enough to create sustainable economic growth in any of the sectors.

Although the results from the regression analyses presented here does not clearly demonstrate such an effect, hypothesis 3 should be assumed to be true. It is demonstrated that the resource abundant states are the states with low economic growth. It is also demonstrated that high infant mortality rates have a significant and negatively impact on economic growth. Also, the inclusion of the measure of human capital, *IMR*, impacts the coefficient of *Primary*, making it less negative. Accounting for infant mortality rates thus makes the relationship between resource dependence and economic growth less negative, indicating that states with better human capital are less harmed by the presence of natural resources.

Hypothesis 4a indicates that the tertiary sector has a higher explanatory power on economic growth than the secondary sector. All findings in this thesis support this hypothesis. The literature on sector composition in India noted this unusual tendency for the tertiary sector to be more important than the secondary, even before the secondary sector is fully developed. The findings in Table 4.4 and Table 4.5 confirm this, as the tertiary throughout yield more positive and more significant coefficients than the secondary sector.

States with a higher share of the tertiary sector in SGDP will also have a higher teledensity, according to hypothesis 4b. The interaction term *Tertiary***Teledensity* is included in model 1 in Table 4.8, and it is significant. As the teledensity has a negative impact on economic growth as the dependence on tertiary sector increases, the hypothesis cannot be confirmed.

According to hypothesis 5a, states with good possibilities for freight will have a higher share of secondary sector in SGDP. None of the regression analyses demonstrates such an effect. Road density is surprisingly negative, and none of the dummies for coastline is significant either. However, Table 5.1 indicates that there in fact is a relationship between having a coastline and a big secondary sector. Only two of the eight states with a coastline do not have a high economic growth rate on average from 1980 through 2009. Further, only two of the eight states with the highest share of secondary sector in SGDP do not have a coastline. These two states, Haryana and Rajasthan, still have achieved high economic growth. Hypothesis 5a should be discarded based on the regression analyses, but kept if Table 5.1 is emphasized.

The last hypothesis, 5b, assumes that states with a higher share of secondary sector in SGDP will have better scores on human capital. The interaction term *Secondary*IMR* cannot confirm this, as it tests how human capital affects the economic returns from the secondary sector. What it does confirm, however, is that high share of human capital has a positive impact on growth when the share of secondary in SGDP is low, and that this positive impact on economic growth grows as the share of secondary sector rises. It can therefore be concluded that human capital affects economic growth positive when the share of secondary sector is low, and that this effect rises when the share of secondary sector in SGDP grows.

7 Conclusion

The aim of this thesis has been to apply the existing literature on the resource curse to the case of India, to see whether it can explain some of India’s divergent growth rates. As similar studies on the US states have found the literature to fit, it was expected to be found here as well. The results from a set of time-series cross-section regression analyses indicates that resource dependence, measured as the share of primary sector in SGDP, has a significant and negative impact on economic growth. The results are robust to the inclusion of the (few) control variables created. However, most of the literatures revolve around extractive natural resources, but India’s primary sector is mostly dependent on agriculture. Therefore, the literature does not seem to fit as well as the results from the regression analyses indicate.

Initially, seven hypotheses were presented. The main hypothesis in this study is that natural resource dependence is expected to have a negative impact on economic growth. The results from the regression analyses indicate that this is the case. Four transmission channels from the existing literature were presented as explanations of how natural resources impact economic growth. Through the discussion of these, the relatively poor fit to the Indian case became evident. Although corruption and weak institutions might have a negative effect on the economic growth in India, this is likely to be caused (primarily) of other factors than the presence of natural resources.

Lacking investments are also not likely caused by natural resources. Investments in manufacturing have been demonstrated to be important to create economic growth through the secondary sector. However, the general development level in India is far below that of e.g. the US. Investments in human resources and infrastructure are assumed to be as detrimental to growth in India as investments in the manufacturing sector are to growth in the US. The Indian government acknowledge that poverty, inequality and the rural-urban divide preclude sustained long-term economic growth. Further, they seek to take measures to invest in human capital and infrastructure to raise all boats, lifting more people out poverty and into productive parts of the economy. The theory that governments fail to realize the need for investment caused by the allegedly sound economic future prospects caused by the natural resources is also a poor fit for two reasons: a) agriculture hardly creates a sound economic future prospect, and b) governments acknowledge the need for investment in growth-promoting factors to be of highest importance.

Despite the fact that education has been found to have the least impact of the transmission channels, it is the one with the best fit to the Indian case. The hypothesis that “easy riches lead to sloth” does however not fit well. Again, it is hardly likely that the government fail to invest in education because of the easy riches that occur in the wake of the agricultural industry. On the other hand, agriculture is likely to be the industry where the need for skilled labour are lowest, and areas attached to agriculture are therefore likely to be least educated. The connection between the dependence on agriculture and low education to slow economic growth can be made.

The lack of sound and comprehensive data for this study should be considered. Similar studies on the subject should strive to include more control variables into the analyses, to improve the explanatory power of the models. As the best model in this study only explains 7.8 percent of the economic growth, it is needless to say that future research should aspire to raise this. Drawing conclusions from models with such a low explanatory force is risky, and I have therefore tried to make use of the existing literature in the discussion to better understand how the results should be interpreted.

Although the regression results fit the description of a resource curse, the literature does not. It can therefore not be concluded that the low economic growth is due to natural resource dependence and not other underlying causes. The main hypothesis, H1, cannot be confirmed nor rejected. However, it is possible that the poor fit of the literature is due to the composition of the primary sector, consisting primarily of the agriculture industry. The agriculture industry is a poorer fit to the resource curse than e.g. extractive resources such as petroleum and minerals.

The other hypotheses are well accounted for in Chapter 6.3, with the most noteworthy being the confirmation of the tertiary sector as the most detrimental to economic growth in India. This confirms previous research on divergent growth rates. It also supports the empirical findings that the tertiary sector has bypassed the secondary at an earlier stage in the development of the economy than what is theoretically anticipated.

References

Agriculture Research Data Book (2002). *Length of coastline and population of coastal States and Islands*. Retrieved the 27th of May 2014 from http://iasri.res.in/agridata/02data%5Cchapter%204%5Cdb2002tb4_5.htm

Auty, R.M. (2001). *Resource Abundance and Economic Development*. Oxford University Press, Oxford.

Bajpai, N., & Sachs, J. D. (2011). India's Decade of Development: Looking Back at the Last 10 Years, and Looking Forward to the Next 20.

Barro, R. J. (2001). Human capital and growth. *American Economic Review*, 12-17.

Beck, N. (2001). Time-series-cross-section data: what have we learned in the past few years?. *Annual review of political science, 4*(1), 271-293.

Beck, N. L., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American political science review, 89*(3), 634-647.

Bhattacharya, B. B., & Sakthivel, S. (2004). Regional growth and disparity in India: Comparison of pre-and post-reform decades. *Economic and Political Weekly, 39*(10), 1071-1077.

Bulte, E. H., Damania, R., & Deacon, R. T. (2005). Resource intensity, institutions, and development. *World Development, 33*(7), 1029-1044.

Dasgupta, D., Maiti, P., Mukherjee, R., Sarkar, S., & Chakrabarti, S. (2000). Growth and interstate disparities in India. *Economic and Political Weekly, 35*(27), 2413-2422.

Dholakia, R. H. (1994). Spatial dimension of acceleration of economic growth in India. *Economic and Political Weekly, 29*(35), 2303-2309.

Eikemo, T. A. & Clausen, T. H. (2007). *Kvantitativ analyse med SPSS. En praktisk innføring i kvantitative analyseteknikker.* Trondheim: Tapir Akademisk Forlag

Freeman, D. G. (2009). The ‘Resource Curse’and regional US development. *Applied Economics Letters, 16*(5), 527-530.

Gerard, B. (2011). A Natural Resource Curse: Does it Exist Within the United States?.

Guio, A. C., Murthi, M., & Drèze, J. (1995). Mortality, fertility, and gender bias in India: A district-level analysis. *Population and development review, 21*(4), 745-782.

Gylfason, T. (2001). Natural resources, education, and economic development. *European Economic Review, 45*(4), 847-859.

Hamilton, L. C. (1992). *Regression with graphics. A second course in applied statistics*. Duxbury Press.

Hamilton, L. C. (2006). *Statistics with STATA. Version 12.* University of New Hampshire.

Indexmundi (no date). *India – Consumer Price Index*. Retrieved March 28th from http://www.indexmundi.com/facts/india/consumer-price-index

Indiastat (no date). *Front page.* http://www.indiastat.com/default.aspx

Jakobsen, T.G & Jakobsen, J. (2007). 13.1 Paneldata. In Eikemo, T. A. & Clausen, T. H. (eds.) *Kvantitativ analyse med SPSS. En praktisk innføring i kvantitative analyseteknikker.* Trondheim: Tapir Akademisk Forlag.

James, A., & Aadland, D. (2011). The curse of natural resources: An empirical investigation of US counties. *Resource and Energy Economics*, *33*(2), 440-453.

Klasen, S. (2002). Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development. *The World Bank Economic Review*, *16*(3), 345-373.

Kurian, N. J. (2000). Widening regional disparities in India: some indicators. *Economic and Political Weekly*, *35*(7), 538-550.

Leite, C. A., & Weidmann, J. (1999). Does mother nature corrupt? Natural resources, corruption, and economic growth. *Natural Resources, Corruption, and Economic Growth (June 1999). IMF Working Paper*, (99/85).

Listhaug, O. & Ringdal, K. (2008). Trust in political institutions. In Ervasi, H., Fridberd, T., Hjerm, M. & Ringdal, K. (Eds.) *Nordic social attitudes in a European perspective*. Northampton Massachussets, Edward Elgar Publishing.

Mehlum, H., Moene, K., & Torvik, R. (2006). Institutions and the resource curse*. *The Economic Journal*, *116*(508), 1-20.

Miller, A. & Listhaug, O. (1999). Political Performance and Institutional Trust. In Norris, P. (Eds.) *Critical Citizens*. New York, Oxford University Press.

Mo, P. H. (2001). Corruption and economic growth. *Journal of Comparative Economics*, *29*(1), 66-79.

Papyrakis, E., & Gerlagh, R. (2003). *Natural resources: a blessing or a curse?* (No. 2003.8). Fondazione Eni Enrico Mattei.

Papyrakis, E., & Gerlagh, R. (2004). The resource curse hypothesis and its transmission channels. *Journal of Comparative Economics*, *32*(1), 181-193.

Papyrakis, E., & Gerlagh, R. (2007). Resource abundance and economic growth in the United States. *European Economic Review*, *51*(4), 1011-1039.

Planning Commission of India (2006). Towards Faster and More Inclusive Growth: An Approach to the Eleventh Five Year Plan (2007-2012). *Government of India*. Retrieved May 8th 2014 from http://planningcommission.nic.in/plans/planrel/app11_16jan.pdf

Purfield, C. (2006). *Mind the Gap-Is Economic Growth in India Leaving Some States Behind?*. International Monetary Fund.

Rao, M. G., Shand, R. T., & Kalirajan, K. P. (1999). Convergence of incomes across Indian states: A divergent view. *Economic and Political Weekly*, *34*(13), 769-778.

Robinson, J. A., Torvik, R., & Verdier, T. (2006). Political foundations of the resource curse. *Journal of development Economics*, *79*(2), 447-468.

Sachs, J. D., & Warner, A. M. (1995). *Natural resource abundance and economic growth* (No. w5398). National Bureau of Economic Research.

Sachs, J. D., & Warner, A. M. (2001). The curse of natural resources. *European economic review*, *45*(4), 827-838.

Sachs, J. D., Bajpai, N., & Ramiah, A. (2002). Understanding Regional Economic Growth in India. *Asian Economic Papers*, *1*(3), 32-62.

The United Nations (2014a). *The System of National Accounts (SNA)*. Retrieved the 15th of May from https://unstats.un.org/unsd/nationalaccount/sna.asp

The United Nations (2014b). *System of National Accounts 1993 - 1993 SNA*. Retrieved the 15th of May from https://unstats.un.org/unsd/nationalaccount/sna1993.asp

Topalova, P. (2008). India: is the rising tide lifting all boats?. *IMF Working Papers*, 1-29.

Wooldridge, J. M. (2006). *Introductory Econometrics. A Modern Approach*. Thomson South-Western.

World Bank (no date, a). *World Data Bank*. Retrieved March 27th 2014 from http://databank.worldbank.org/data/views/reports/tableview.aspx

World Bank (no date, b). *India Overview*. Retrieved March 27th 2014 from http://www.worldbank.org/en/country/india/overview

Appendix A

A.1 Descriptive statistics

Table A.1 Descriptive statistics. All case statistics are given for the variable with one year lag.

Variable	**N**	**Min.**	**Max.**	**Mean**	**St. dev.**
Dependent:					
Growth	419	-45.370	42.420	5.751	6.730
DummyGrowth	435	0	1	0.007	0.254
Sectors:					
Primary	420	12.307	55.541	33.090	10.458
Secondary	420	10.699	41.974	25.107	6.111
Tertiary	420	9.148	60.473	41.593	8.544
Agriculture	420	4.918	52.974	28.155	10.260
Mining	420	0.006	15.358	2.285	2.658
Forestry	420	0.090	7.852	1.283	0.990
Fishing	420	0.027	4.140	1.168	1.069
Control Variables:					
UrbanPopGrowth	406	-8.205	22.330	0.010	2.080
LiteracyDisparity	422	4.960	34.550	21.508	6.464
IMR	422	13.273	133.200	64.242	21.663
RoadDensity*	437	74.940	31 166.870	1 215.766	2 186.067
RoadDensity ln	422	4.317	10.098	6.703	0.741
Teledensity*	241	0.150	71.813	10.974	15.041
Teledensity ln	227	-1.897	4.124	1.322	1.350
Dummies:					
DummyCoastline	437	0	1	0.526	0.500
Coastline A	437	0	1	0.265	0.442
Coastline B	437	0	1	0.261	0.440
Coastline C	437	0	1	0.474	0.500

*Variables used to create the power transformed variables Teledensity ln and RoadDensity ln. Not lagged.

A.2 List of variables used in the regressions

Dependent variable

Growth: State-wise annual growth in state gross domestic product (SGDP) for each year between 1980 and 2009. All data retrieved from Indistat.com, with the heading “Gross State Domestic Product (GSDP) at Factor Cost by Industry of Origin in *State* {(At Current Prices)} (*year* to *year*)}”. The data is divided into three time periods, 1980-1981 to 1992-1993, 1993-1994 to 1999-2000 and 1999-2000 to 2009-2010[5]. Data are re-calculated from current prices to constant 2009 prices using (SGDP Year_t * Indian CPI Year 2009)/Indian CPI Year_t).

DummyGrowth: As the data is not directly comparable across the time periods, I have included a dummy variable to control for the abnormal growth rates the first year in the new time period. 1993 and 2000 have the value 1, all the other years 0. This is always included with *Growth*.

Sector Variables

Primary: Measure of state-wise dependence of natural resources as the share of SGDP as percent for each year between 1980 and 2009. The primary sector consists of agriculture, mining and quarrying, forestry and fishing. Data for each state from the same files as *Growth*.

Secondary: Measure of state-wise share of secondary sector as percentage of SGDP for each year between 1980 and 2009. The secondary sector consists of manufacturing, electricity, gas and water supply and construction. Data for each state from the same files as *Growth*.

---

[5] The data on Madhya Pradesh is from four time periods: 1980-1981 to 1988-1989, 1989-1990 to 1997-1998, 1993-1994 to 1999-2000 and 1999-2000 to 2007-2008. In the data from 1989-1990 to 1997-1998, only the data until 1993-1994 have been used. This is also discussed in the text.

Tertiary: Measure of state-wise share of tertiary sector as percentage of SGDP for each year between 1980 and 2009. The tertiary sector consists of trade, hotel and restaurants, railways, transport by other means, communication, storage, banking and insurance, real estate, ownership, business and legal, public administration and other services. Data for each state from the same files as *Growth*.

Agriculture: Measure of state-wise share of agriculture as percentage of SGDP for each year between 1980 and 2009. Data for each state from the same files as *Growth*.

Mining and quarrying: Measure of state-wise share of mining and quarrying as percentage of SGDP for each year between 1980 and 2009. Data for each state from the same files as *Growth.*

Forestry: Measure of state-wise share of forestry as percentage of SGDP for each year between 1980 and 2009. Data for each state from the same file as *Growth*.

Fishing: Measure of state-wise share of fishing as percentage of SGDP for each year between 1980 and 2009. Data for each state from the same file as *Growth*.

Dummy variables

Coastline: Dummy to separate states with coastline from those without. States with coastline=1, states without coastline=0. Data from Central Statistical Organization, Compendium of Environment Statistics, 2002.

Coastline A: Dummy to identify the states with the longest coastline. Includes Gujarat, Kerala, Andhra Pradesh and West Bengal, whom all have a coastline longer than 900 km., all valued=1. Other=0. Data from Central Statistical Organization, Compendium of Environment Statistics, 2002.

Coastline B: Dummy to identify the states with shorter coastline. Includes Maharashtra, Tamil Nadu, Karnataka and Goa with coastlines varying from 840km. to

300km. All valued=1. Other=0. Data from Central Statistical Organization, Compendium of Environment Statistics, 2002.

Coastline C: Dummy to identify the states without coastline. Includes Assam, Bihar, Haryana, Madhya Pradesh, Punjab and Uttar Pradesh, all valued=1. Other=0. Data from Central Statistical Organization, Compendium of Environment Statistics, 2002.

Control variables

Literacy rate: Measure of state-wise literacy rate, measured as the number of literate per hundred citizens. Data are retrieved from Indiastat.com with the heading "State-wise literacy rate in India (1951-2011 Census)". Data from 1951, 1961, 1971, 1981, 1991, 2001 and 2011 are interpolated using STATA. Numbers from 1981-2009 are included in the variable used in the analysis.

Literacy gender: Measure of state-wise gender disparity in literacy rates, measured by percent points. Numbers from the census from 1961, 1971, 1981, 1991 and 2001 are retrieved from Indiastat.com with the headings "State-wise Effective Literacy Rate by sex in India 1961, 1971, 1981, 1991 and 2001 Part I", "State-wise Effective Literacy Rate by sex in India 1961, 1971, 1981, 1991 and 2001 Part II" and "State-wise Effective Literacy Rate by sex in India 1961, 1971, 1981, 1991 and 2001 Part III". Provisional numbers for 2011 are retrieved from "State-wise Provisional Literacy Rate by sex in India (as per 2011 Census)". The numbers are calculated by Literacy rates *men* – Literacy rates *women* = Literacy disparity. The literacy disparity numbers are interpolated using STATA. Numbers from 1981-2009 are included in the variable used in the analysis.

IMR: State-wise Infant Mortality Rates is a measure for number of infants under the age of 1 per 1000 live births. Numbers are retrieved from Indiastat.com with the headings "State-wise Infant Mortality Rates in India (1961, 1971 and 1981)", "State-wise Infant Mortality rate by Residence in India (1995 to 1999)" and "State-wise Infant Mortality Rates by Sex and Residence in India

(2010)”. Data from 1961, 1971, 1981, 1995, 1999 and 2010 are interpolated using STATA. Numbers from 1981-2009 are included in the variable used in the analysis.

PopUrban: Measure of state-wise urban population, measured in 1000s. Retrieved from Indiastat.com with the headings “State-wise Progress in Urban Population (excluding Jammu & Kashmir) of India (1901 to 1991)”, Statewise Projected Urban Population by sex (As on 1st March , 2002-2004)” and “Statewise Projected Urban Population by sex (As on 1st March, 2008-2010)”. The data from 1971, 1981, 1991, 2002, 2003, 2004, 2008, 2009 and 2010 are interpolated using STATA. The numbers from 1901 to 1991 are not in 1000s, but this was adjusted previous to the interpolation. Numbers from 1981-2009 are included in the variable used in the analysis.

UrbanGrowth: Measure of state-wise percentage growth in urban population. This variable is based on *PopUrban* and is calculated as (*PopUrban* Year$_t$ – *PopUrban* Year $_{t-1}$) / *PopUrban* Year$_t$ * 100.

Roaddensity: Measure of state-wise road density, measured as road length in kms per 100 km^2 of area. The data are retrieved from Indiastat.com with the headings “State-wise Road Density in India (1971-72, 1981-82, 1991-92 & 1996-1997)” and “State-wise Road Length in Relation to Area and Population in India (As on 31st March, 2011)”. Data from all years are interpolated using STATA. Numbers from 1981-2009 are included in the variable in the analysis.

Teledensity: Measure of number of telephone connections for every hundred individuals. The numbers are based on service area and/or state, replicating the states rather well. However, as Goa is included in the service are of Maharashtra, there are no observations for Goa in this variable. The data are retrieved from Indiastat.com with the headings “State-wise Tele-Density in India (as on 31.1.1992)”, “State-wise Teledensity in India (1999-2000)”, “State-wise Teledensity in India (2002 and 2003)” and “Service Area-wise Rural and Urban Teledensity in India (As on 31.03.2012)”. As there are no data previous to 1992, there are no observations for the ten first years. The data is interpolated, and numbers from 1992-2009 are included in the variable used in the analysis.

Appendix B

B.1 Dependent variable

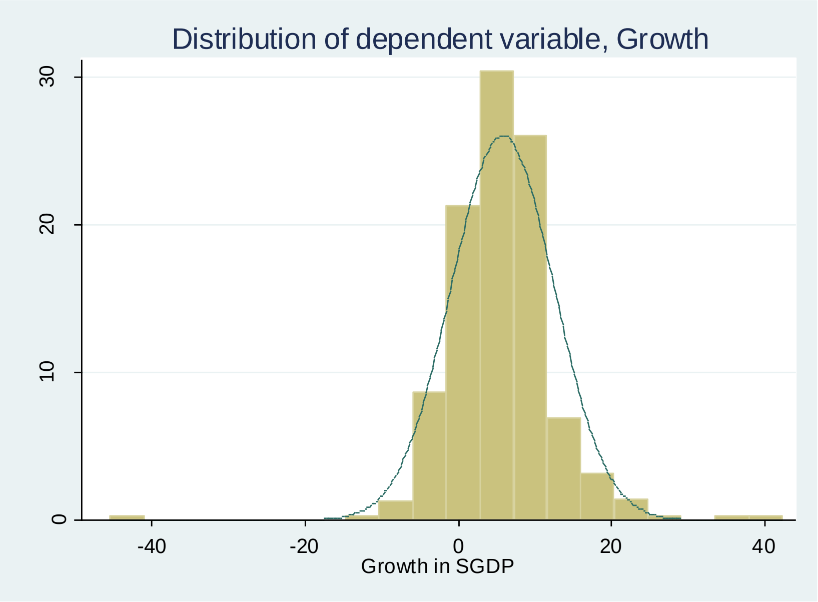

Figure B.1 Normal distribution of dependent variable, economic growth

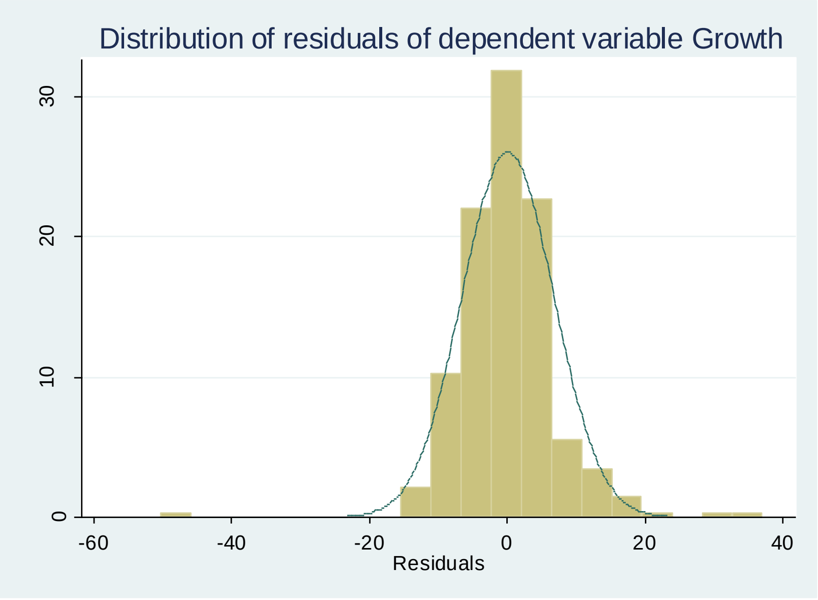

Figure B.2 Normal distribution of residuals of dependent variable, economic growth.

B.2 Tests for multicollinearity

Table B.1 Tolerance test of all variables

Variable	Tolerance
Primary	**.006**
Secondary	**.041**
Tertiary	**.034**
Agriculture	**.011**
Mining	**.146**
Forestry	.353
Fishing	**.144**
Literacy	**.109**
LiteracyDisparity	**.080**
IMR	**.187**
UrbanPopGrowth	.795
RoadDensity	.338
Teledensity	.239
Coastline A	**.153**
Coastline B	.227
DummyGrowth	.850

Table B.2 Tolerance test of Sector

Variables	Tolerance
Primary	**.039**
Secondary	**.065**
Tertiary	**.099**

Table B.3 Tolerance test of the primary industries with and without Primary

Variables	Tolerance	Tolerance
Primary	**.019**	
Agriculture	**.020**	.860
Mining	.212	.850
Forestry	.490	.904
Fishing	.560	.783

Table B.4 Tolerance test of control variables before lagging

Variable:	**Tolerance**	**Tolerance w/ Primary**	**Tolerance w.o./ LiteracyRate**
LiteracyRate	.253	**.141**	
LiteracyDisparity	.298	**.175**	.326
IMR	.279	.268	.274
Coastline A	.662	.469	.504
Coastline B	.678	.363	.414
RoadDensity	.719	.719	.723
Teledensity	.711	.580	.560
UrbanPopGrowth	.833	.831	.844
Primary		**.178**	.317

Table B.5 Tolerance test of control variables after lagging

Variable	**Tolerance**	**Tolerance w/Primary**	**Tolerance w.o/ LiteracyRate**
LiteracyRate	**.189**	**.142**	
LiteracyDisparity	.259	**.136**	**.183**
IMR	.293	.284	.287
Coastline A	.552	.321	.324
Coastline B	.688	.290	.307
RoadDensity	.403	.374	.387
Teledensity	.453	.269	.269
UrbanPopGrowth	.840	.837	.854
Primary		**.129**	**.169**

B. 3 Regression results

Table B.6 Inclusion of one variable at a time with Primary

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	9.216***	10.83***	10.18***	11.40***	9.334***	9.559***	9.412***	9.152***
	(-0.862)	(-1.422)	(-1.280)	(-2.605)	(-0.847)	(-0.940)	(-0.822)	(-1.731)
Primary L1	-0.116***	-0.152***	-0.136***	-0.125***	-0.118***	-0.110***	-0.102***	-0.136***
	(-0.024)	(-0.034)	(-0.029)	(-0.027)	(-0.023)	(-0.023)	(-0.030)	(-0.044)
DummyCoastA		-0.346						
		(-0.608)						
DummyCoastB		-1.342						
		(-0.952)						
DummyCoastline			-0.618					
			(-0.638)					
RoadDensity ln L1				-0.280				
				(-0.299)				
UrbanPopGrowth L1					0.074			
					(-0.052)			
LliteracyDisparity L1						-0.025		
						(-0.038)		
IMR L1							-0.010	
							(-0.011)	
Teledensity ln L1								0.318
								(-0.356)
DummyGrowth	5.293***	5.310***	5.301***	5.298***	5.263***	5.307***	5.311***	5.884***
	(-1.743)	(-1.757)	(-1.752)	(-1.749)	(-1.760)	(-1.740)	(-1.743)	(-1.807)
R^2	**0.073**	**0.077**	**0.075**	**0.074**	**0.074**	**0.074**	**0.074**	**0.099**
N	419	419	419	419	403	419	419	226

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table B.7 Inclusion of one variable at a time with Agriculture

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	8.438*** (-0.672)	9.164*** (-1.050)	8.820*** (-0.974)	9.615*** (-2.505)	8.598*** (-0.677)	8.841*** (-0.942)	9.063*** (-0.754)	7.970*** (-1.575)
Agriculture L1	-0.109*** (-0.025)	-0.127*** (-0.026)	-0.117*** (-0.023)	-0.114*** (-0.023)	-0.113*** (-0.025)	-0.102*** (-0.022)	-0.085*** (-0.026)	-1.121*** (-0.046)
DummyCoastA		-0.059 (-0.659)						
DummyCoastB		-0.751 (-0.856)						
DummyCoastline			-0.275 (-0.651)					
RoadDensity ln L1				-0.155 (-0.332)				
UrbanPopGrowth L1					0.0599 (-0.051)			
LliteracyDisparity L1						-0.0283 (-0.042)		
IMR L1							-0.0203** (-0.010)	
Teledensity ln L1								0.468 (0.398)
DummyGrowth	5.221*** (-1.778)	5.218*** (-1.790)	5.219*** (-1.784)	5.221*** (-1.784)	5.183*** (-1.797)	5.241*** (-1.770)	5.273*** (-1.764)	5.969*** (-1.802)
R^2	**0.068**	**0.070**	**0.068**	**0.068**	**0.069**	**0.069**	**0.071**	**0.095**
N	419	419	419	419	403	419	419	226

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table B.8 Inclusion of one variable at a time with Secondary

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	2.845***	3.180***	3.251***	-0.041	2.655***	4.867***	6.340***	2.916***
	(-0.748)	(-0.910)	(-0.842)	(-2.560)	(-0.559)	(-1.102)	(-1.072)	(-0.663)
Secondary L1	0.101***	0.071	0.068	0.095***	0.112***	0.102***	0.057*	0.058*
	(-0.030)	(-0.051)	(-0.048)	(-0.029)	(-0.023)	(-0.025)	(-0.032)	(-0.034)
DummyCoastA		0.869						
		(-0.618)						
DummyCoastB		0.722						
		(-0.790)						
DummyCoastline			0.816					
			(-0.667)					
RoadDensity ln L1				0.453				
				(-0.358)				
UrbanPopGrowth L1					0.109			
					(-0.081)			
LiteracyDisparity L1						-0.095**		
						(-0.043)		
IMR L1							-0.037***	
							(-0.009)	
Teledensity ln L1								0.841**
								(-0.336)
DummyGrowth	5.204***	5.218***	5.219***	5.204***	5.148***	5.263***	5.308***	5.912***
	(-1.784)	(-1.764)	(-1.761)	(-1.776)	(-1.809)	(-1.757)	(-1.746)	(-1.789)
R^2	**0.049**	**0.052**	**0.052**	**0.052**	**0.052**	**0.057**	**0.062**	**0.084**
N	419	419	419	419	403	419	419	226

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table B.9 Inclusion of one variable at a time with Tertiary

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	-0.380	-0.347	-0.279	1.545	-0.121	-0.313	1.518	-1.293
	(-0.788)	(-0.589)	(-0.639)	(-1.872)	(-0.793)	(-1.211)	(-1.500)	(-2.526)
Tertiary L1	0.138***	0.135***	0.133***	0.156***	0.133***	0.137***	0.115***	0.135**
	(-0.016)	(-0.014)	(-0.015)	(-0.018)	(-0.017)	(-0.016)	(-0.023)	(-0.056)
DummyCoastA		0.274						
		(-0.709)						
DummyCoastB		0.025						
		(-0.581)						
DummyCoastline			0.166					
			(-0.608)					
RoadDensity ln L1				-0.401				
				(-0.289)				
UrbanPopGrowth L1					0.074			
					(-0.046)			
LiteracyDisparity L1						-0.002		
						(-0.043)		
IMR L1							-0.015	
							(-0.011)	
Teledensity ln L1								0.466
								(-0.368)
DummyGrowth	5.446***	5.443***	5.439***	5.474***	5.404***	5.446***	5.446***	6.281***
	(-1.657)	(-1.667)	(-1.668)	(-1.651)	(-1.678)	(-1.658)	(-1.665)	(-1.666)
R^2	**0.071**	**0.072**	**0.071**	**0.073**	**0.069**	**0.071**	**0.073**	**0.097**
N	419	419	419	419	403	419	419	226

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table B.10 Interaction terms with Primary

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	21.70***	15.00***	14.27***	14.25***
	(-5.991)	(-2.438)	(-2.67)	(-3.753)
Primary L1	-0.347*	-0.116***	-0.104***	-0.0315
	(-0.209)	(-0.025)	(-0.031)	(-0.064)
RoadDensity ln L1	-1.662**	-0.621**	-0.599*	-0.803**
	(-0.839)	(-0.308)	(-0.328)	(-0.352)
UrbanPopGrowth L1	0.126**	0.109**	0.700***	0.135**
	(-0.060)	(-0.054)	(-0.227)	(-0.058)
IMR L1	-0.024**	-0.031	-0.022*	-0.021*
	(-0.011)	(-0.024)	(-0.012)	(-0.013)
LiteracyDisparity L1				0.072
				(-0.121)
Primary*RoadDensity	0.036			
	(-0.029)			
Primary*IMR		0.000		
		(-0.001)		
Primary*UrbanGrowth			-0.022***	
			(-0.009)	
Primary*Literacy				-0.004
				(-0.004)
DummyGrowth	5.315***	5.329***	5.259***	5.356***
	(-1.773)	(-1.775)	(-1.761)	(-1.758)
R^2	**0.078**	**0.077**	**0.078**	**0.079**
N	403	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table B.11 Interaction terms with Secondary

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	9.426***	5.253	7.543***	10.18***
	(-2.062)	(-4.334)	(-2.753)	(-1.896)
Secondary L1	0.078	0.273	0.204***	0.0613**
	(-0.084)	(-0.169)	(-0.072)	(-0.024)
UrbanPopGrowth L1	0.126**	0.129**	0.111*	-0.0342
	(-0.058)	(-0.063)	(-0.067)	(-0.233)
RoadDensity ln L1	-0.445**	0.264	-0.584***	-0.478*
	(-0.217)	(-0.645)	(-0.218)	(-0.246)
IMR L1	-0.046***	-0.050***	0.008	-0.048***
	(-0.008)	(-0.007)	(-0.027)	(-0.006)
DummyCoastline	0.884			
	(-1.917)			
Secondary*Coastline	-0.031			
	(-0.091)			
Secondary*RoadDensity		-0.031		
		(-0.024)		
Secondary*IMR			-0.002**	
			(-0.001)	
Secondary*UrbanGrowth				0.007
				(-0.011)
DummyGrowth	5.287***	5.294***	5.335***	5.294***
	(-1.778)	(-1.784)	(-1.772)	(-1.778)
R^2	**0.065**	**0.065**	**0.066**	**0.065**
N	403	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Figure B.12 Interaction terms with Tertiary

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	1.722	8.439**	9.336**	7.633*
	(-6.681)	(-3.722)	(-4.213)	(-4.042)
Tertiary L1	0.241***	0.115***	0.091**	0.133***
	(-0.077)	(-0.024)	(-0.046)	(-0.026)
UrbanPopGrowth L1	0.149***	-0.206	0.124***	0.130***
	(-0.043)	(-0.558)	(-0.048)	(-0.047)
RoadDensity ln L1	-0.510	-0.841**	-0.819**	-0.818*
	(-0.380)	(-0.408)	(-0.391)	(-0.419)
IMR L1	-0.053**	-0.034***	-0.051	-0.034**
	(-0.021)	(-0.013)	(-0.031)	(-0.017)
Teledensity ln L1	4.658*			
	(-2.601)			
DummyCoastline				1.624
				(-1.379)
Tertiary*Teledensity	-0.101**			
	(-0.051)			
Tertiary*UrbanGrowth		0.006		
		(-0.010)		
Tertiary*IMR			0.000	
			(-0.001)	
Tertiary*Coastline				-0.038
				(-0.036)
DummyGrowth	6.520***	5.467***	5.488***	5.510***
	(-1.592)	(-1.683)	(-1.694)	(-1.713)
R^2	**0.123**	**0.075**	**0.076**	**0.076**
N	225	403	403	403

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

B.4 Robust tests

Table B.13 Regression without Kerala, and without Kerala and Andhra Pradesh

Variable	**Model 1**	**Model 2**
Constant	15.240***	17.920***
	(-3.321)	(-5.054)
Primary L1	-0.115***	-0.113***
	(-0.035)	(-0.038)
RoadDensity ln L1	-0.766*	-1.100*
	(-0.418)	(-0.634)
UrbanPopGrowth L1	0.775***	0.940**
	(-0.280)	(-0.458)
IMR L1	-0.015	-0.021
	(-0.015)	(-0.018)
Primary*UrbanGrowth	-0.025**	-0.032*
	(-0.011)	(-0.019)
DummyGrowth	4.763***	4.224**
	(-1.823)	(-1.843)
R^2	**0.069**	**0.063**
Observations	376	350

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

Model 1: Without Kerala

Model 2: Without Kerala and Andhra Pradesh

Table B.14 Regression with Primary without 1993 and 2000

Variable	**Model 1**	**Model 2**
Constant	16.320***	16.300***
	(-2.822)	(-2.823)
Primary L1	-0.135***	-0.135***
	(-0.031)	(-0.031)
RoadDensity ln L1	-0.874**	-0.874**
	(-0.356)	(-0.359)
UrbanPopGrowth	0.185***	0.270
	(-0.064)	(-0.363)
IMR L1	-0.009	-0.009
	(-0.013)	(-0.013)
Primary*UrbanGrowth		-0.003
		(-0.012)
R^2	**0.061**	**0.061**
N	373	373

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

Table B.15 Regression with Secondary without 1993 and 2000

Variable	**Model 1**	**Model 2**
Constant	9.877***	8.302***
	(-2.513)	(-2.804)
Secondary L1	0.099***	0.184***
	(-0.033)	(-0.068)
RoadDensity ln L1	-0.671***	-0.736***
	(-0.260)	(-0.263)
UrbanPopGrowth L1	0.205***	0.197***
	(-0.048)	(-0.045)
IMR L1	-0.038***	-0.004
	(-0.009)	(-0.027)
Secondary*IMR		-0.001
		(-0.001)
R^2	**0.036**	**0.036**
N	373	373

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

Table B.16 Regression with Tertiary without 1993 and 2000

Variable	**Model 1**	**Model 2**
Constant	9.710**	8.840**
	(-3.991)	(-3.837)
Tertiary L1	0.119***	0.085**
	(-0.025)	(-0.038)
RoadDensity ln L1	-1.109**	-0.901**
	(-0.492)	(-0.447)
UrbanPopGrowth L1	0.198***	0.114**
	(-0.068)	(-0.045)
IMR L1	-0.0285*	-0.026
	(-0.015)	(-0.017)
Teledensity ln L1		-0.481
		(-1.258)
Tertiary*Teledensity		0.017
		(-0.027)
R^2	**0.045**	**0.081**
N	373	198

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

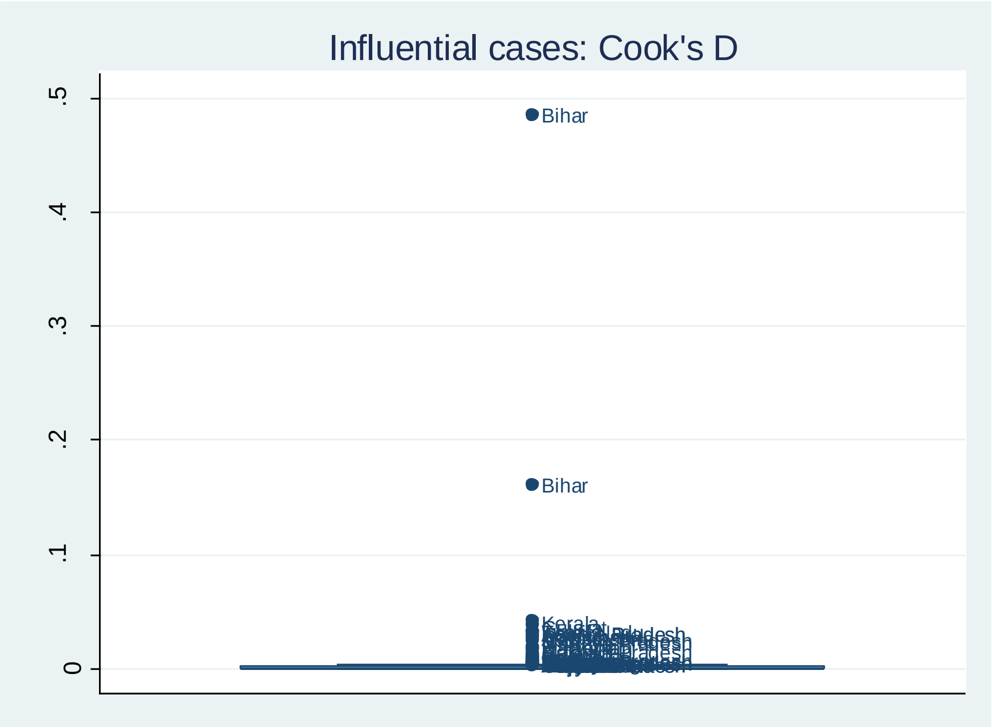

Figure B.3 Influential units, Cook's D

Table B.17 Regression with Primary without influential cases

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	14.730***	16.950***	14.550***	16.860***
	(-2.457)	(-2.710)	(-2.536)	(-2.753)
Primary L1	-0.094***	-0.129***	-0.094***	-0.129***
	(-0.026)	(-0.029)	(-0.025)	(-0.028)
RoadDensity ln L1	-0.672**	-0.878***	-0.669**	-0.876***
	(-0.289)	(-0.324)	(-0.303)	(-0.328)
UrbanPopGrowth L1	0.124**	0.166***	0.655***	0.438
	(-0.050)	(-0.053)	(-0.223)	(-0.331)
IMR L1	-0.026**	-0.025**	-0.024**	-0.024**
	(-0.011)	(-0.011)	(-0.011)	(-0.011)
Primary*UrbanGrowth			-0.020**	-0.010
			(-0.008)	(-0.012)
DummyGrowth	6.157***	6.413***	6.089***	6.398***
	(-1.712)	(-0.874)	(-1.710)	(-0.869)
R^2	**0.106**	**0.129**	**0.108**	**0.129**
N	401	383	401	383

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

Model 1: Absolute cut off value, cook's D <0.1

Model 2: Size-adjusted cut off value, cook's D <0.01

Model 3: With interaction term, absolute cut off value, cook's D<0.1

Model 4:With interaction term, size-adjusted cut off value cook's D<0.01

Table B.18 Regression with Primary, two year lag

Variable	**Model 1**	**Model 2**
Constant	16.180***	16.330***
	(-2.852)	(-2.875)
Primary L2	-0.052**	-0.052**
	(-0.025)	(-0.026)
RoadDensity ln L2	-0.872**	-0.874**
	(-0.361)	(-0.361)
UrbanPopGrowth L2	0.153***	-0.307
	(-0.041)	(-0.438)
IMR L2	-0.048***	-0.045***
	(-0.012)	(-0.013)
Primary*UrbanGrowth L2		0.017
		(-0.017)
DummyWeirdGrowth	5.284***	5.288***
	(-1.740)	(-1.740)
R^2	**0.067**	**0.068**
N	387	387

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

Table B.19 Regression with Secondary, two year lag

Variable	**Model 1**	**Model 2**
Constant	15.170***	9.854***
	(-3.016)	(-3.592)
Secondary L2	0.006	0.305***
	(-0.034)	(-0.093)
RoadDensity ln L2	-0.856**	-1.111***
	(-0.357)	(-0.376)
UrbanPopGrowth L2	0.158***	0.126***
	(-0.040)	(-0.037)
IMR L2	-0.063***	0.053
	(-0.009)	(-0.037)
Secondary*IMR L2		-0.005***
		(-0.002)
DummyGrowth	5.259***	5.304***
	(-1.721)	(-1.714)
R^2	**0.063**	**0.069**
N	387	387

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

Table B.20 Regression with Tertiary, two year lag

Variable	**Model 1**	**Model 2**
Constant	11.340**	13.950***
	(-4.486)	(-4.694)
Tertiary L2	0.107***	0.042
	(-0.029)	(-0.040)
RoadDensity ln L2	-1.112**	-1.057**
	(-0.440)	(-0.462)
UrbanPopGrowth L2	0.163***	0.116***
	(-0.057)	(-0.040)
IMR L2	-0.043***	-0.051***
	(-0.014)	(-0.019)
Teledensity ln L2		-1.811
		(-1.289)
Tertiary*Teledensity L2		0.038
		(-0.028)
DummyGrowth	5.439***	5.695*
	(-1.685)	(-3.348)
R^2	**0.074**	**0.099**
N	387	211

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

Table B.21 Regression with Primary, three year lag

Variable	**Model 1**	**Model 2**
Constant	15.230***	15.380***
	(-2.867)	(-2.882)
Primary L3	-0.092***	-0.092***
	(-0.028)	(-0.028)
RoadDensity ln L3	-0.648*	-0.648*
	(-0.359)	(-0.361)
UrbanGrowth L3	0.103	-0.398
	(-0.074)	(-0.565)
IMR L3	-0.034**	-0.037***
	(-0.014)	(-0.013)
Primary*UrbanGrowth L3		0.018
		(-0.021)
DummyGrowth	5.191***	5.190***
	(-1.696)	(-1.695)
R^2	**0.075**	**0.076**
N	371	371

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in brackets.

Table B.22 Regression with Secondary, three year lag

Variable	**Model 1**	**Model 2**
Constant	12.540***	9.774***
	(-2.427)	(-2.986)
Secondary L3	0.025	0.188***
	(-0.027)	(-0.072)
RoadDensity ln L3	-0.565*	-0.716**
	(-0.326)	(-0.324)
UrbanPopGrowth L3	0.096*	0.075
	(-0.056)	(-0.054)
IMR L3	-0.058***	0.003
	(-0.008)	(-0.032)
Secondary*IMR L3		-0.003**
		(-0.001)
DummyGrowth	5.170***	5.228***
	(-1.691)	(-1.699)
R^2	**0.065**	**0.066**
N	371	371

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.

Table B.23 Regression with Tertiary, three year lag

Variable	**Model 1**	**Model 2**
Constant	8.735*	5.593
	(-4.785)	(-5.923)
Tertiary L3	0.142***	0.207***
	(-0.034)	(-0.062)
RoadDensity ln L3	-0.996**	-1.066*
	(-0.481)	(-0.613)
UrbanPopGrowth L3	0.138	0.064
	(-0.095)	(-0.075)
IMR L3	-0.036**	-0.032
	(-0.015)	(-0.024)
Teledensity ln L3		0.635
		(-1.618)
Tertiary*Teledensity L3		-0.019
		(-0.038)
DummyGrowth	5.446***	6.057*
	(-1.627)	(-3.370)
R^2	**0.082**	**0.120**
N	371	184

***p<0.01, **p<0.05, *p<0.1. Standard errors in brackets.