Bachelor's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

Iver Andreas Ugelvik, Ida Nesheim, Magnus Gilje,
Lene Quille

Structural Change and Economic Growth in India

Bachelor's thesis in Economics
Supervisor: Hildegunn Ekroll Stokke
May 2024

NTNU
Norwegian University of
Science and Technology

Iver Andreas Ugelvik, Ida Nesheim, Magnus Gilje,
Lene Quille

Structural Change and Economic Growth in India

Bachelor’s thesis in Economics
Supervisor: Hildegunn Ekroll Stokke
May 2024

Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

NTNU
Norwegian University of
Science and Technology

Summary

This thesis examines whether structural change from agriculture to the service sector in India has contributed to economic growth. In the coming years, India is predicted to surpass large economies like Japan and Germany. For a long time, India followed a protectionist economic strategy. In 1991, the policymakers initiated a liberalization program that deregulated the private sector and allowed for more trade and investment. From 1991 to 2010, the employment in agriculture fell from 64 % to 52%, which was accompanied by a roughly equivalent increase in employment in the service and industry sector. This thesis examines this structural change using the Lewis model, and analyses if these changes could have contributed to economic growth in the period 1991 to 2019. The thesis concludes that there exists a positive statistical correlation between the GDP growth and the employment in the service sector at a 0.05 significance level. Further research should focus on increasing the sample size and expand the analysis to include other countries in the region.

Table of Contents

1 Introduction

In the mid-18th century, India stood for approximately a fourth of the manufactural production in the world. As the country approached the 19th century, manufacturing decreased to about two percent due to the industrial revolution in Europe (Cowen, 2024). In the years prior to India's independence, entrepreneurs started to move their capital from rural to urban areas (Bayly, 1985, pp. 13). In the decades after the independence, the policymakers in India employed a protectionist economic strategy, a decision which remained unquestioned until other Indo-Pacific countries experienced rapid growth by opening up their economy (Mukherji, 2008, pp. 316). In 1991, the Indian government initiated a liberalization program which opened up for more trade and foreign investment and emphasized a greater focus on exports. Additionally, these policies involved a deregulation of the private sector, putting increased faith in market forces (Ahluwalia, 2002, pp. 67). By looking at some indicators for economic development discussed in the next paragraphs, it is possible to observe a positive development in India during the last decades. This has been accompanied by a structural change in the labour market, where the service sector takes up an increased proportion of the labour market. Our thesis uses the Lewis model to understand if structural changes in employment contribute to economic growth. Gérard Roland defines structural changes in an economy as expansion in one sector, when another shrinks (Roland, 2016, pp.114).

Growth in an economy contributes to increased consumption in a society. As the income of workers increases, they can consume more goods. When income rises, it should have a positive effect on the gross domestic product (GDP) of an economy. In addition, it can alleviate poverty (Bhalla and Bhasin, 2024). For instance, it can increase focus on establishing a better infrastructure providing access to more clean water, food and toilets, thus reducing sickness and poor health. This has been the case for India where they have improved access of piping of water from 16,8% to 74,7% in rural areas (Bhalla and Bhasin, 2024). The factors could explain how India lowered its poverty rate from 63,47% in 1977 to 12,15% in 2021 (World Bank, 2017).

From 1950 to 2000, India experienced a structural production shift from agriculture into modern sectors such as industry and service (Joshi, 2004, pp. 4175). During this period the service sector became one of the major drivers of economic growth in India. Unni and Naik (2011) argue that the service sector has contributed to the rapid economic growth. Additionally, the IT-sectors started to grow in 1997 and grew rapidly in the year 2000. This thesis focuses on the shift from agriculture to service sector. The sector covers a wide range of different jobs, which can have a low or a high productivity. It is also a discussion if this sector can manage to grow (Joshi, 2004, pp. 4176).

The average economic growth rate has increased in India during the period 1990 to 2019, as can be observed in Figure 4.1. Furthermore, India is predicted to become the world's third largest economy by 2027. For instance, they are expected to surpass large economies such as Japan and Germany (Laker, 2024). In an article from Forbes, Laker (2024) writes that a new phase is about to start in India. Private investment, both domestic and foreign, will be the focus of the next phase. This will be a change for the Indian market economy because the focus has been on infrastructure, innovation and advancing skills which already has a positive effect on foreign direct investment (FDI). In 2021-2022 India experienced the highest annual FDI inflow up with $84,8 billion (Laker, 2024).

1.1 Research Question

India has undergone substantial changes in the past decades, and the thesis uses the Lewis-model to examine if the structural changes in employment contribute to economic growth. In our literature research we found surprisingly few articles using the Lewis model as a framework when studying economic growth and structural changes in India. The thesis therefore wants to investigate this research gap further. In order to answer our research question, we used several multiple regression analysis based on the variables: Service employment, total employment, degree of openness, and foreign direct investment. The variables are discussed in Chapters 4.

It is predicted that India will become one of the world's largest economic powers (Laker, 2024). Therefore, we want to understand which factors contributed to this change. Knight (2021) argues that the Lewis model is a suitable framework for providing insight into the development of a country. The increasing growth rate makes India an interesting agent when it comes producing products for other big brands. An example of this is Apple which produces 7 % of iPhones in India, which is predicted to increase to 25 % in 2025 (Travelli, 2024). Moreover, it is interesting to look at India due to their impressive economic background, where they changed from a closed economy to an open and liberal economy.

India has seen a reduction in the proportion of workers employed in agriculture, from 64 % in 1991 to 52 % in 2010 (World Bank, 2024a). This has been accompanied by a roughly equal increase in employment in both the industry and service sector. This is an unusual development, as most economies tend to grow their industry before expanding services. Some researchers speculate that this could be detrimental to India's long-term economic growth, as the potential for productivity growth in the service sector is smaller than in industry (Ghose, 2021). Our research question is as follows: *Can structural changes from the agriculture sector to the service sector contribute to economic growth?*

The next chapter gives an introduction to the literature on economic growth as well as papers applying the Lewis model as a framework. The thesis therefore continues with a chapter explaining the Lewis model as our theoretical framework. Afterwards, the empirical method and variables will be discussed. The thesis ends with a chapter with the results from the analysis, a discussion, and a conclusion.

2 Literature

This chapter summarizes literature on the topic of economic growth done by other authors. We will look at which factors have been attributed as causes of growth in previous research, to provide context and support for our research question. The first part of the chapter aims to present well known literature covering the broader theme of economic growth. Further, the chapter goes on to systematize research papers employing the Lewis model as their framework, before the chapter concludes with a brief introduction to the academic debate surrounding our research topic.

2.1 Economic Growth

Our topic of research in this thesis whether structural change has affected economic growth in India. In the book *Development Economic* by Gérard Roland, he identifies three crucial factors that help promote economic growth: Technology progress, population growth, and increases in physical and human capital (Roland, 2016, pp. 102). There are numerous factors that can accumulate economic growth, and Roland defines economic growth as the measurement of GDP (Gross Domestic Product) (Roland, 2016, pp. 584). The coming sections will shed light upon previous research on the topic of economic growth, as well as papers applying the Lewis model to interpret economic growth.

Wu et al. (2017) looked at the economic growth of China. In their research they present estimates that can have contributed to efficient changes in Chinese economy, like productivity growth and innovation (Wu et al., 2017, pp. 213). Their book concludes that productivity has a positive relationship with economic growth, and this factor is the most important in both the traditional sectors like agriculture and the modern sectors, such as service, to accumulate growth (Wu et al., 2017, pp. 22). There exists a lot of research on the economic miracles in Asia, some of which look at the structural changes in an economy, and how economic growth is affected. A paper written by J. Laitner uses Engels law to observe how structural changes affect the measured savings rate in an economy based on historical statistic from United states and United Kingdom after industrialization (Laitner, 2000, pp. 545-547). The findings in the paper suggests that when income is low, agriculture has a more significant role for national income in a country. If the income increases, the used share in agriculture goods declines, and at the same time the now reproduced capital will displace the importance of land. The article argues that the evidence shows that an increase in income goes hand in hand with an increase in saving rate (Laitner, 2000, pp. 559).

Mackie (1964) analyzes the effect of agriculture on economic growth and development (Mackie, 1964, pp. 1). He argues that agriculture contributes to economic growth through progression in the sector that will increase employment in other sectors than agriculture. He also argued that the more income and production will lead increased balanced in prices, as well as more progression in non-agricultural sectors because of higher demand in non-farm products (Mackie, 1964, pp. 9).

Bosworth and Collins (2008) compare the economic growth for both China and India and examines what lies behind the growth (Bosworth and Collins, 2008, pp. 45). The article states that India stands out for the share of the population without schooling in contrast to other Asian economies (Bosworth and Collins, 2008, 51-52). When looking at the sectoral growth, the paper from Bosworth and Collins finds that employment in agriculture has increased, but at a reduced rate. They point out that this could have a relationship with the population growth (Bosworth and Collins, 2008, pp. 55). They also write that the biggest share of the service sector in India, accounted for 60 percent, was wholesale, transportation, and retail trade (Bosworth and Collins, 2008, pp. 56). The article write that China have had a remarkable growth in the last 25 years, and pointing out that India has a growth rate like East-Asian industrializing economies. At the same time, they point out that the growth in India comes from the service sector, and that Chinas growth comes from agriculture, service, and industry (Bosworth and Collins, 2008, pp. 62). The paper concludes that the prospect on the supply-side of growth for both India and China in form of labor, reallocates over the sectors, and the physical capital looks positive for the future (Bosworth and Collins, 2008, pp. 64).

As mentioned in the introduction India has had a rapid economic growth rate the last decades. Basu and Maertens (2007) study several factors that could explain the economic growth in India. They analyze the transition of India from having an overall weak economic performance to overtaking many other economies in terms of growth rate. At the same time, they argue that employment rate has gone down, and the wages are not increasing as much as the income per-capita (Basu and Maertens, 2007, pp.162). Basu and Maertens conclude that if India wishes to continue the rapid economic growth, they must address the *"bottlenecks in Indian economy"* (Basu and Maertens, 2007, pp.164). They identify these bottlenecks as bankruptcy regulation, corruption in bureaucracy, infrastructure, and labor (Basu and Maertens, 2007, pp.164). In their article they address only some of these factors, but conclude that inequality can lead to more poverty which could destabilize India's rapid growth continued (Basu and Maertens, 2007, pp.164). Acharyya (2009) looks at two key effect of FDI inflow, which Acharyya identifies as environmental degradation and GDP growth (Acharyya, 2009, pp. 43). At the same time Acharyya writes that there is no agreement that FDI has a positive correlation with economic growth (Acharyya, 2009, pp. 43-44). In a few decades, the paper argues that India has through liberal trade increased their inflow of foreign direct investment, after having a fare more restrictive policies (Acharyya, 2009, pp. 45). Acharyya conclude that the impact of FDI in India have had a small positive relationship on economic growth (Acharyya, 2009, pp. 56).

2.2 Former Data

In an article authored by Nazrul Islam and Kazuhiko Yokata (2008), they try to understand the industrialization in China considering the Lewis model (Islam and Yokota, 2008, pp. 359), and found that China largely followed the predictions of the Lewis model better than most other economies, and concluded that China gradually was approaching the Lewis turning point[1], but at the same time were far from it (Islam and Yokota, 2008, pp. 391). The article also identified some weaknesses with the implication of the Lewis model and wrote that it was difficult to find empirical sectors that were like the theoretical duality in the Lewis model. They also argue that it was laborious to know which wage to examine as well as the marginal product of labor (Islam and Yokota, 2008, pp. 392). Islam and Yokata conclude that the industrialization of China could be better understood using the Lewis model (Islam and Yokota, 2008, pp. 392).

Another paper commenting the Lewis model is the article authored by John Knight. The paper examines the labor markets of China and South Africa, and uses the Lewis model as a framework (Knight, 2021, pp. 143). He states that the model gives a good framework, but that the theory needs modification. The reason is that the empirical findings like relative price mechanism do not congruous with the theory in the Lewis model (Knight, 2021, pp. 168). The paper claims that both China and South Africa have segmentation and inflexibility in their labor markets. Knight concludes that South Africa has low investors' confidence and a low growth rate. He also states that social instability, like crimes, has a root in unemployment. He also points out that the spread of corruption gives lower trust for the businesses and investment which leads to a pessimistic view of the South African economy. On the other hand, Knight looks at China, which in contrast to South Africa, has a high growth rate. The paper writes that the high rate is a combination of many factors, but the elimination of inefficiency associated with the planned economy and the reemployment of part-time employees and openness for trade as factors. Knight also states that the huge investment in infrastructure and confidence in investment contributed to further economic growth (Knight, 2021, pp. 169). The article concludes that the Lewis model can be used to understand the development in the countries, but that the model does not cover enough (Knight, 2021, pp. 170).

2.3 Academic Debate

There is a large amount of literature on economic growth in India, but remarkably few papers use the Lewis model as their chosen framework. The literature previously presented in this chapter provides examples where the Lewis model has been applied to explain economic growth.

In a paper written by John Knight, he mentions that the Lewis model could be used as an adequate framework for interpreting the economic growth of developing countries, but also states that the theory needs modifications. Both John Knight and Islam & Yokata makes use of the Lewis model with the objective of understanding economic development and structural changes in each country, and are mentioning that the model alone is not enough. Knight states that the model does not cover enough, and Islam & Yokata identified some weaknesses with the model. In other words, the framework of the

[1] A Lewis turning point is reached when there is no longer a surplus of labor in agriculture. This leads to an increase in both agricultural and industrial wages as well as reduced profits and investment in industry. (Das and N'Diaye, 2013).

Lewis model might be a good explanation of structural change, but the assumption can make it difficult to apply to the reality.

The thesis will on the other hand argue that the Lewis model gives a superior framework to analyze economic growth considering structural changes, and will argues that this topic requires further analysis, hence our research question. It will also use the Lewis model as a framework to attain a greater understanding of the causal factors of India's rapid economic growth. The theory chapter provides a summary of the core aspects of the Lewis model and its framework.

3 Theory - Economic Growth and Lewis Model

The topic of economic growth has received great attention from economical researchers for the last few decades as showcased in Chapter 2. This chapter begins with an introduction to economic growth theory building on production functions. Afterwards, the Lewis model is presented with a focus on the analysis of economic growth.

3.1 Production Function

A production function is used for modelling the relationship between a country or a firms output Y and the production factors capital K and labor N. This relationship is modelled in equation (3.1).

$$Y = F(K, N) \tag{3.1}$$

In a firm, there are multiple factors that play a part in production. Firstly, a sufficient number of qualified workers are needed. They also rely on other firms to provide raw materials and services necessary for producing the relevant goods (Holden, 2016, pp. 80-81). The production function shows how these factors contributes to the total production level. If there is an increase in real capital or labor, it will have a positive effect on the GDP, as the variable Y is increased.

The relationship between the output and production factors is the foundational basis of economic growth. In order for a firm to produce goods, it need access to both real capital and labor. They can then contribute to economic growth by increasing their production and thus output. Compared to the Lewis model, which explains the structural changes from agriculture to a modern sector (Lewis, 1954), the production function highlights the necessities for a firm to contribute to growth in GDP (Holden, 2016, pp. 80). Therefore, we can see that a structural change can result in an increased modern sector, where there is potential for firms to achieve higher output levels, thus leasing to economic growth.

3.2 The Lewis Model

The theoretical framework used in this thesis is the Lewis model, and we have chosen economic growth as the main dependent variable. Economic growth is defined in Section 2.1. The Lewis model is a dual sector model which explains the structural change which transfers labour from agriculture to the modern sector (Roland, 2016, pp.116). For the rest of this section, agriculture will be denoted as the first sector, while the modern sector is denoted as the second sector. The section will first present the agriculture sector, then present the modern sector and finally combine them into a complete model.

Agriculture - Sector One

The production function TP_A describing the agriculture sector is shown in equation (3.2). Here $\bar{K}_A$ indicate the land/capital in agriculture, and $\bar{t}_A$ indicate the technology in agriculture. A visualization of the model can be seen in Figure 3.1. Here MP_{LA} stands for the marginal product of labor in agriculture, AP_{LA} is the average product of labor in agriculture, W_A is the wages in agriculture, L_A is the current labor in agriculture, and Q_{LA} stands for quantity labor in agriculture.

$$TP_A = F(L_A, \bar{K}_A, \bar{t}_A) \tag{3.2}$$

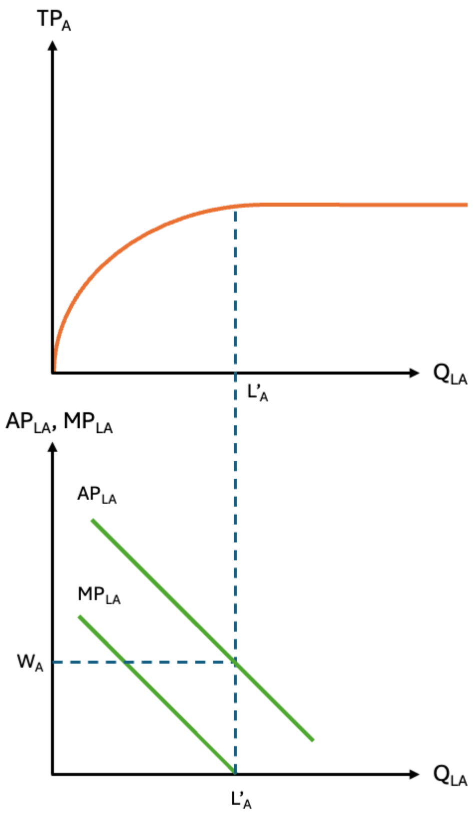


Figure 3.1: Productivity in the agricultural sector.

The model explaining the agriculture relies on some important assumptions. The first is surplus in labor which can be seen in equation (3.3). This means that given that the industrial sector needs more labor, the surplus of workers in agriculture can be reallocated to meet their required need of labour force. The reallocation of workers will not affect the agriculture sector, and their production will persist (Lewis, 1954, pp. 141).

$$MP_{\mathrm{LA}} = 0, L_A < L_A \tag{3.3}$$

The second assumption states that the wages in agriculture is equal to the average product of labor in agriculture as seen in equation (3.4). This means that there is a non-competitive market. (Lewis, 1954, pp. 142).

$$W_A = AP_{LA} = \frac{TP_A}{L_A} \tag{3.4}$$

The Modern Sector - Sector two

The production function in the service and industry sector TP_M can be seen in equation (3.5). The function contains L_M which stands for Labor in the modern sector, $\bar{K}_M$ stands for capital/land in the modern sector, while $\bar{t}_M$ is the technology in the modern sector. A visualization of this sector can be seen in Figure 3.2, where S_L indicates labor supply and D_1 indicates labor demand.

$$TP_M = F(L_M, \bar{K}_M, \bar{t}_M) \tag{3.5}$$

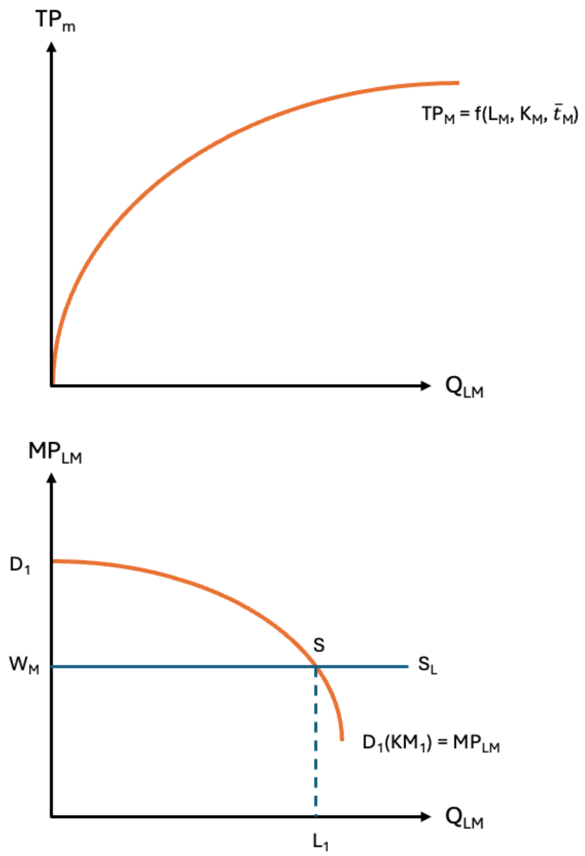


Figure 3.2: Productivity in the modern sector.

The model explaining the modern sector relies on some assumptions. The first one states that supply of labor is perfectly elastic because there is an unlimited access to workers from the agriculture sector (Roland, 2016, pp.118). The cheap work and surplus of labor results in a high profit (Lewis, 1954, pp. 149). The second assumption states that there is a competitive market. Here the wage equals to the marginal product, with the marginal product being greater than zero (Lewis, 1954, pp. 150-151), as seen in the modern sector as seen in equation (3.6).

$$W_M = MP_{LM}, MP_{LM} > 0 \tag{3.6}$$

The third assumptions is that the labor demand decrease and is equal to the marginal product of labor in the modern sector (Lewis, 1954, pp. 150). The last assumption is important for the economic growth. All profits are reinvested and resulting in an increase in capital/land K. This creates a shift in the production function and in the marginal product of labor function, meaning both curves are experiencing a shift upwards. When all the profits are reinvested into new capital/land, we are left with a high capital accumulation and an inflow of workers from agriculture(Lewis, 1954, pp. 152).

Combining Both Sectors

When the modern sector provides a profit, this profit is reinvested and ensures a positive shift in the model. First from TP_{M1} to TP_{M2}, and then from TP_{M2} to TP_{M3} and so on. This is a reaction from the increase in demand for labor after reinvesting in capital/land. Now the workers are more productive, and they have a higher profit, therefore the demand for labor will also have an increase from F to G. This is where the wage is equal to the marginal product of labor (MP_{LM} = D), and the new demand for labor from L_1 to L_2. The wage persist because of the perfectly elastic supply of labor.

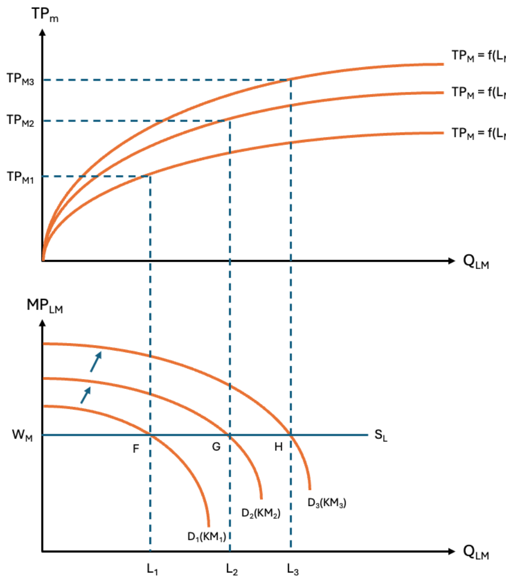


Figure 3.3: An overview of the effect of shifts in the Lewis model.

As the marginal productivity of labor approaches zero, workers in agriculture will be encouraged to change occupation and seek employment in a more productive sector, like industry. This will lead to farm workers migrating to urban areas where the new industrial jobs are to be found. This transition persists until there no longer is an excess of labor employed in agriculture. The rate at which this happens depends upon the demand for labor in industry, which is determined by the marginal product of industrial labor.

The wage level for industrial workers is determined by the productivity in agriculture. To attract workers, industries need to provide wages that are at least as high as the income as they would get by remaining in agriculture. The wage in agriculture is therefore equal to the average product of labor in agriculture. This results in the condition seen in equation (3.7)

$$W_A = AP_L A = \frac{TP_A}{L_A} \tag{3.7}$$

3.3 Weaknesses of the Lewis Model

Although the Lewis model can explain structural changes, it might exist cases where the model struggles at explaining the situation in a given poor country. In reality, there are many poor countries that do not experience these structural changes, as the Lewis model explains. It is because some of the assumptions in the model are not fulfilled. In general, there is not a reallocation of labor from the traditional sector to the modern sector, due to a lack of capital (Roland, 2016, pp.120). The Lewis model states that capital is to be reinvested to achieve continuous growth. However, a sizable amount of capital is needed to be able to establish a modern sector and grow the economy. In practice, many poor countries do not obtain enough capital to initiate this process, and if they do, it is not given that the profits are reinvested domestically instead of abroad. Additionally, the profits earned might not be enough for it to have an impact on the country's development (Roland, 2016, pp. 121). Certain weaknesses with the models framework are mentioned in the article by Islam and Yokata, where they argue that "*there are generic problems concerning the application of the Lewis model, such as difficulty in finding empirical sectors that correspond to the theoretical duality, ambiguity regarding the type of wage to examine*" (Islam and Yokota, 2008, pp.392).

4 Dataset, Variables and Empirical Method

This chapter introduces the dataset used in this thesis. The chapter starts by discussing the dataset and an introduction to our selected variables. Then, it continues with an overview and discussion of how each variable is measured. Finally, the chapter ends with a brief introduction to the method applied in Chapter 5.

4.1 Dataset Overview

The dataset is gathered by using the world development indicators provided by the World Bank Open Data Catalogue (World Bank, 2024c). The data bank contains data from 1960 to today and consists of 167 distinct variables measured in different frequencies. In this application, the dataset has been reduced by extracting a subset of the original variables. The reduction is necessary as most of the variables are not relevant when applying the Lewis model for analyzing economic growth in India. Furthermore, the data prior to 1990 has been discarded as it is out of scope for the research question, and because many data points are missing in this period. This leads us to a dataset from 1990 to 2019.

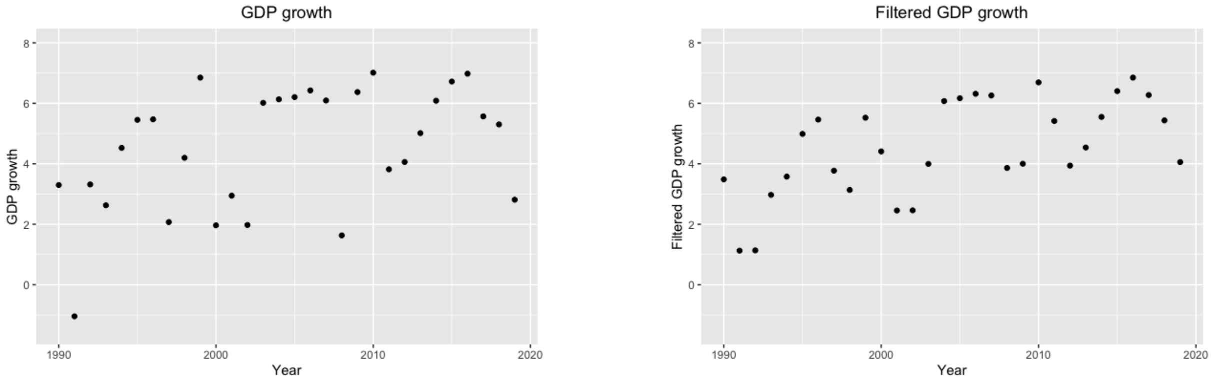
4.2 Selected Variables

This section introduces the subset of variables selected for the analysis as well as an overview of how each variable is measured. The variables are either chosen to reflect the suggestions in the Lewis model as discussed in Section 3.2 or to account for other factors which could lead to a rapid economic growth. The dependent variable is GDP growth, the main independent variable is the employment in service sector and the rest of the introduced variables are control variables.

4.2.1 GDP Growth as Dependent Variable

As the purpose of our research is to investigate the relationship between structural change and economic growth, it thus follows that we use GDP growth as our dependent variable. The GDP growth is measured by the World Bank (World Bank, 2024c) and is defined as the annual percentage growth rate of GDP given a constant local currency, where the constant prices are fixed at the 2015 price levels. The GDP summarizes all gross value added by all producers in the economy and taxes minus subsidies.

This thesis uses a pre-processed version of the GDP growth based on a moving average filtering method discussed in Section 4.5.3. The raw and pre-processed data is showcased in Figure 4.1, where the raw data is seen on the left in Figure 4.1a and the filtered version can be seen in Figure 4.1b. It can be seen from the figure that the average growth rate increases over time. A summary of descriptive statistic for this variable can be seen in Table 4.1.



(a) Raw GDP growth (b) Filtered GDP growth

Figure 4.1: Overview of the raw and filtered GDP growth

	Mean	Median	Minimum	Maximum
GDP Growth	4.590	5.302	-1.045	7.013
Filtered GDP Growth	4.562	4.538	1.126	6.851

Table 4.1: Descriptive statistics for GDP growth and filtered GDP growth.

4.2.2 Employment in Service

The Lewis model predicts that a structural change from agricultural labor to the modern sector generates economic growth. This thesis interprets the service sector as the modern sector in India, which is illustrated in Figure 4.2. The variable is obtained from the World Bank Development Indicator (World Bank, 2024c) and contains every person in working age engaged in wholesale, retail trade or works in service jobs such as in hotels, transportation, storage, real estate, IT and business services. The figure shows a steady increase in employment in the service sector from 22 percent and up to approximately 33 percent. Additionally, a small dip in service employment can be observed in the period from 2006 to 2010.

Figure 4.3 shows that the increased employment in the service sector is due to a large reduction in agriculture employment, which is reduced from 64 percent to 41 percent. Additionally, the industry sector has an increase in the same period from 15 percent to 25 percent. Based on the data, it is not possible to assume that most of the agriculture employment is moved directly to the service sector, as it is possible that people move from agriculture to industry and then to the service sector, with an equal flow of people between the agriculture to industry and from industry to service. However, the figure do show a clear decrease in agriculture employment and an increase in service employment. An overview of the descriptive statistics for all three sectors can be seen in Table 4.2.

	Mean	Median	Minimum	Maximum
Service Employment	26.48	26.83	22.02	33.34
Agriculture Employment	54.16	53.75	41.30	63.41
Industry Employment	19.36	18.22	14.57	25.37

Table 4.2: Descriptive statistics for the service, agriculture and industry sector.

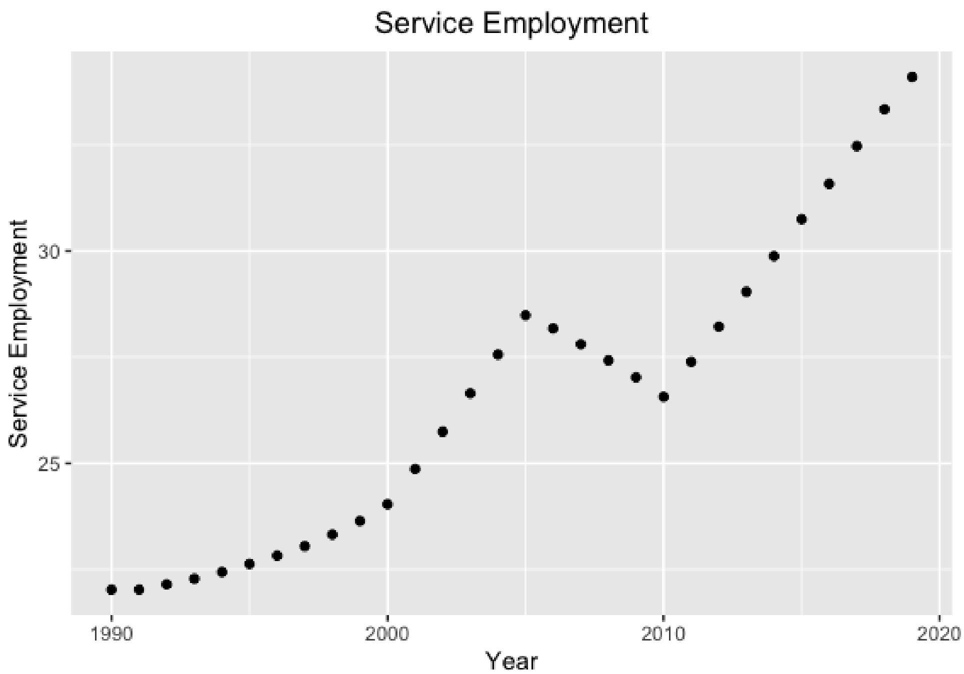
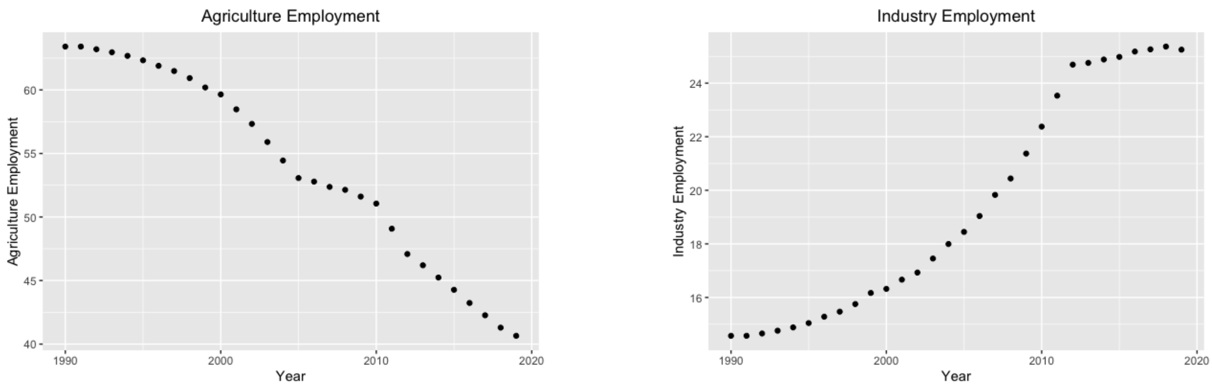


Figure 4.2: Employment in the service sector relative to total employment.



(a) Employment in the agriculture sector relative to total employment.

(b) Employment in the industry sector relative to total employment.

Figure 4.3: Overview of employment in the agriculture and industry sector

4.2.3 Degree of Openness

Degree of openness is defined as imports I plus exports X as a share of GDP. Increasing trade allows for each nation to specialize in goods which they can produce more efficiently relative to other goods. This means that resources (whether it be land, labour or capital) will be allocated more efficiently to yield higher production at a lower expense, while still meeting market demand. As GDP can be viewed as an aggregate of all domestic production, this causes GDP to rise. It is worth noting that the degree of openness is not directly present in the dataset. However, this factor can be derived based on the import and export relative to gross domestic product GDP as seen in equation (4.1).

$$Openness = I_r + X_r = \frac{I}{GDP} + \frac{X}{GDP} \tag{4.1}$$

The degree of openness is visualized in Figure 4.4, where it is shown that the variable varies in magnitude between about 18 percent and reaches a maximum value of approximately 56. This value is a combination of the import and export percentage of GDP, and thus the degree of openness can be interpreted as a percentage.

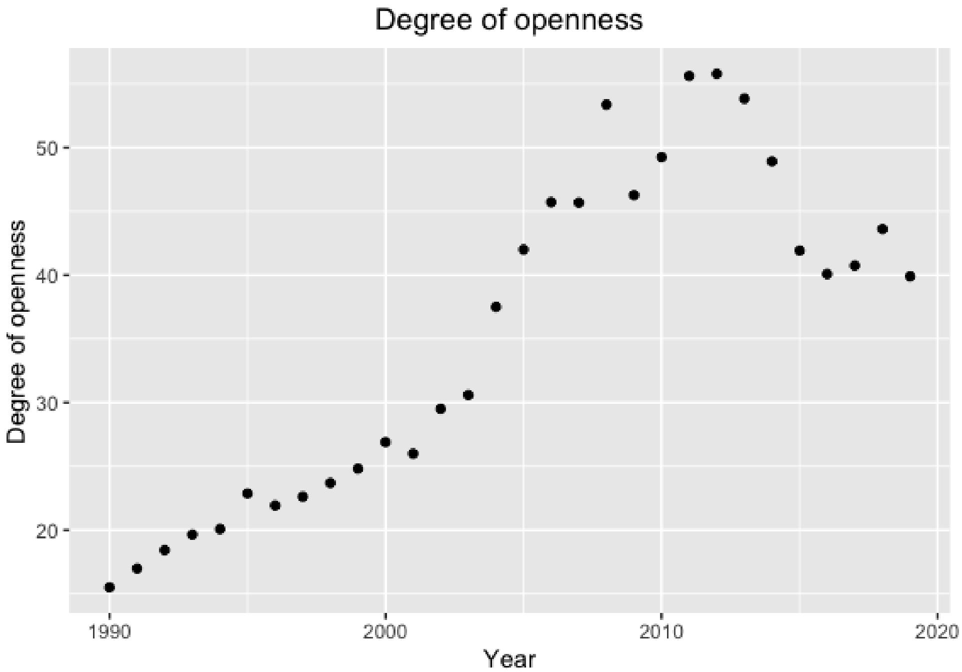
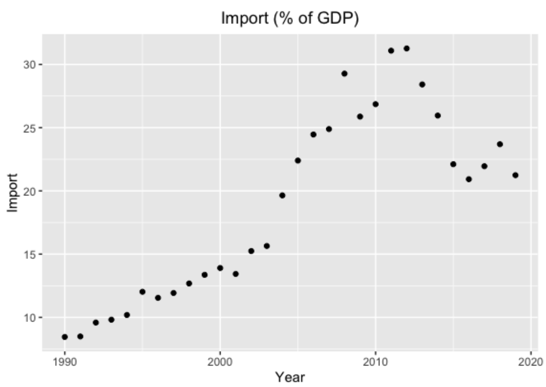
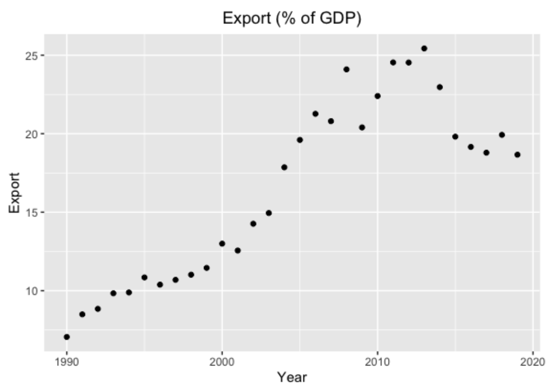


Figure 4.4: Degree of openness to trade

Figure 4.4 shows that the degree of openness steadily increases from the start of 1990 and until 2010, before it starts decaying. The decay in openness is a result of a decay in both the import and export from India in the same time period as observed in Figure 4.5.



(a) Import as percentage of GDP



(b) Export as percentage of GDP

Figure 4.5: Overview of the import and export of India between 1990 and 2020

The degree of openness consist of both relative import I_r and relative export X_r, illustrated in Figure 4.5 and provided by the data bank (World Bank, 2024c). The relative import represents all goods and services received from the rest of the world relative to the GDP of India. This includes merchandise, freight, insurance, transport, and other services. However, this excludes salaries for employees and investment incomes from the rest of the world. The export represents all goods and services provided to the rest of the world. This contains the same factors as seen in the relative import, but in this the goods and services exported from India in a given year. Table 4.3 summarises the descriptive statistics for the degree of openness, import and export data.

	Mean	**Median**	**Minimum**	**Maximum**
Degree of Openness	35.17	37.50	15.51	55.79
Import	18.796	19.645	8.453	31.259
Export	16.375	17.859	7.053	25.431

Table 4.3: Descriptive statistics for degree of openness, import and export.

4.2.4 Foreign Direct Investments

As production is dependent on capital, increasing investments could be contributing factor to economic growth (Holden, 2016, pp. 80-81). Since developing countries tend to have a plentiful supply of labour and a lack capital, an inflow of investments from foreign countries has the potential to yield large increases in production. This phenomena is known as foreign direct investments, which is visualized in Figure 4.6. As the figure shows, foreign direct investments (FDI) are steadily increasing in the period from 1990 to 2005, but rapidly increases in 2006 to stabilize at a higher level for the rest of the period. This quantity is measured based on the balance of payment data reported to the international monetary fund (World Bank, 2024c). This consists of equity investments, enterprise investments, debt, and reverse investments (IMF, 2009). A summary of the descriptive statistics for FDI can be found in Table 4.4.

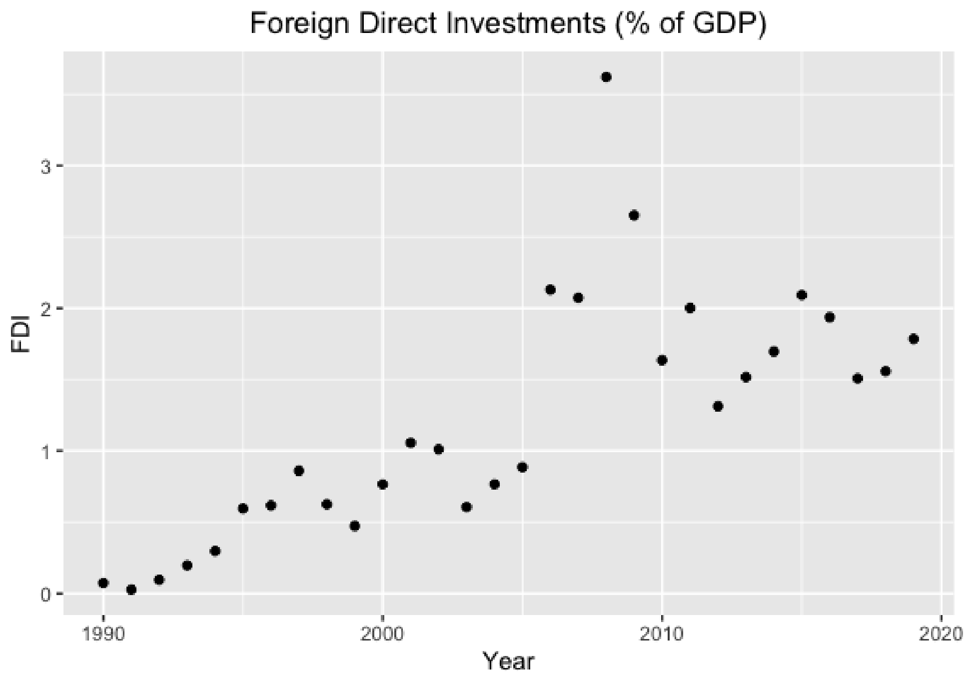


Figure 4.6: Foreign direct investments as percentage of GDP over time

It is worth noting that Figure 4.6 shows a large spike in FDI as share of GDP from about 0.9 percent to 3.7 percent in the years 2005 to 2008. This behavior may aligns with the financial crisis in 2007 and might be due to a decrease in GDP.

	Mean	**Median**	**Minimum**	**Maximum**
FDI	1.19605	1.01157	0.02723	3.62052

Table 4.4: Descriptive statistics for foreign direct investments.

4.2.5 Total Employment

Changes in the total employment relative to the population can explain the GDP growth, as an increase in the workforce can increase the GDP of the country. The total employment relative to the population is retrieved from the World Bank (World Bank, 2024c) and is measured as people of working age that produce goods or services for payment relative to the total potential population that could participate in this activity. The variable is illustrated graphically in Figure 4.7, where the total employment is shown for each year measured in percentages.

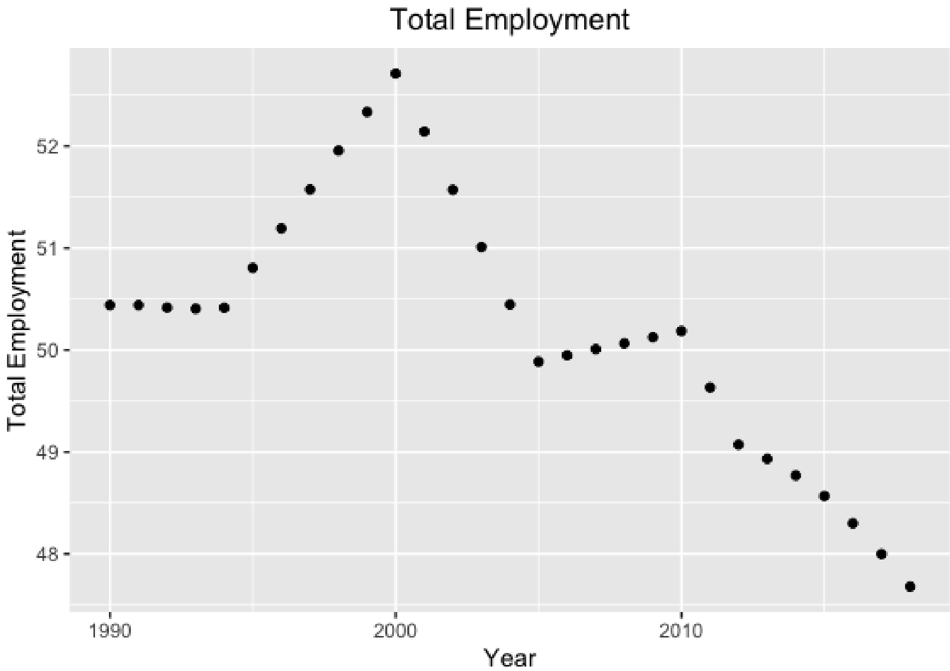


Figure 4.7: Total employment relative to the population

The figure shows that the total employment increases rapidly from the start to year 2000, and afterwards decreases from about 53 percent and down to less than 48 percent when approaching 2020. An overview of descriptive statistics for total employment can be found in Table 4.5.

	Mean	Median	Minimum	Maximum
Total employment	50.24	50.30	47.68	52.71

Table 4.5: Descriptive statistics for total employment.

4.3 Linear Regression - Overview

The rest of the chapter is based on the theory of econometrics presented in Wooldridge (2019). Linear regression is a tool for analyzing the relationship between two or more variables. The model assumes that the relationship between two variables can be described as a linear equation as seen in equation (4.2). In the equation y is known in the literature as the dependent variable, while x is the independent variable. Furthermore, β_0 is known as the intercept parameter and β_1 is the slope parameter, describing the direct relationship between x and y.

$$y = \beta_0 + \beta_1 x \tag{4.2}$$

In order to perform estimates using linear regression, it is assumed that y is a result of a realization of the stochastic variable Y as seen in equation (4.3). Equation (4.3) and (4.2) are quite similar, except the introduction of the error term u. This error term is assumed to have an expectation equal to zero and a constant variance σ^2.

$$Y = \beta_0 + \beta_1 x + u \tag{4.3}$$

The parameters β_0 and β_1 can be estimated by minimizing the least squares error. This framework aims at creating a line with the optimal fit to the data. This is done by choosing the parameter values which minimizes the squared distance from each observation to the resulting line. This optimization problem is summarized in equation (4.4), and an derivation of how this is applied to estimate the parameters can be seen in Wooldridge (2019).

$$LS = \sum_{i}^{n} (y_i - Y_i)^2 = \sum_{i}^{n} (y_i - \beta_0 - \beta_1 x_i - u_i)^2 \tag{4.4}$$

4.3.1 Extensions to Multiple Linear Regression

This model is beneficial given an analysis of the relationship between two variables. However, the coming analysis requires an extension of the simple linear regression model into the multiple linear regression. This generalization is necessary as the model introduced in the analysis in Section 5.1 utilizes multiple independent variables. The multiple linear regression (MLR) is formulated mathematically in equation (4.5).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{4.5}$$

Similarly to the simple linear regression, the multiple linear regression estimates the regression parameters by finding the set of parameters which minimizes a least squares error similar to the one introduced in equation (4.4).

4.4 Linear Regression - Assumptions and Robustness

The linear regression framework is based on a few assumptions, to guarantee an unbiased result with low variance. These are known as the Gauss-Markow assumptions and collectively guarantees the best linear unbiased estimator (BLUE). Each of the assumptions, and how to check for the conditions, are discussed in more details in the next few sections.

4.4.1 Linear in Parameters

This assumption states that the model is linear in the parameters β_i. The requirement implies that the resulting model should be on the form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{4.6}$$

4.4.2 Random Sampling

Random requires that observations should be randomly sampled from the population. Random sampling is accomplished in this thesis by using a single sample from each year over a long period of time.

4.4.3 No Perfect Co-linearity

No perfect co-linearity means that none of the independent variables in the dataset should have perfect linear relationships. The requirement do however allow for co-linearity between the variables. In the case of this thesis, where the structural change from agriculture to service is analyzed, the linear model cannot include both the current percent of people working in agriculture and service. This is because they are perfectly linear dependent on each other, thus invalidating the assumption.

4.4.4 Zero Conditional Mean

Zero conditional mean states that the expected value of the error term in the stochastic model seen in equation (4.3) should equal to zero. This is checked after the analysis is done, by analyzing the error term and checking that average error is close to zero.

4.4.5 Variance - Homoskedasticity and Multicollinearity

The final assumption guarantees a minimal variance of the estimator given that the variance of the dependent variable is constant, often referred to as homoskedasticity. This can be checked after the analysis by plotting the error residuals for the dependent variable relative to the fitted values, or by visually inspecting the plot of the dependent variable.

Furthermore, it is desirable to avoid the multicollinearity of variables. This is because a high correlation in between the independent variables makes it difficult to discern the individual effect of each predictor on the dependent variable, and thus the validity of the regression result. This issue can be accounted for using a variance inflation factor (VIF) test. A VIF result less than 10 is a good indication of low levels of multicollinearity.

4.5 Pre-processing of Data

This thesis analyses the GDP growth over time for India between 1990-2019. The GDP growth can be seen to have a higher average value in the period from about 2000 to 2010 4.1. However, the variable is quite noisy, thus making it difficult for an MLR model to actually predict the values. Additionally, there exists a few outlier samples with quite low growth compared to the other values at the adjacent samples. These are mainly in the early 90s and 2007 and is explained by the financial crises that occurred in those years. However these outliers reduce the performance of MLR, as there are no variables actually accounting for financial crises.

This thesis aims to showcase that the structural change from agriculture to the modern sector over time leads to a period with higher GDP growth as postulated by the Lewis model discussed in Section 3.2. Hence the exact value of the samples is not really relevant, but rather that the trend of a lower or higher GDP growth is related to the structural change. Thus, it would be acceptable to manipulate the dependent variable in order to showcase the trend of the GDP growth rather than the actual values. There are multiple methods for achieving this as discussed in the next few subsections.

4.5.1 Dummy Encoding

In order to showcase the trend of the GDP growth, it is possible to encode the values for GDP growth into a dummy variable. In this case this means encoding the values into a 0 or 1, where 0 could mean low growth, and 1 could mean high growth. However this task requires the authors to find a lower limit for high growth, which might not reflect the actual result. Additionally, a lot of information will be lost as it is not possible to distinguish between the levels of low growth or even negative growth seen during some of the financial crises. This method is avoided in the thesis due to mentioned limitations, and because the method does not handle the outliers in the years with financial crises.

4.5.2 Removing Outliers

Another approach for handling the outliers in the dataset would be to remove the outliers entirely. This would ensure that the disturbance does not affect the final model. However, this would require a systematic approach for detecting the outliers. Additionally, the method would reduce the sample size, thus making it more difficult to draw any conclusions from the regression result. This approach would reduce the variance of the dependent variable, but it is omitted in this thesis as it would also remove valuable samples from the independent variable.

4.5.3 Moving Average Filtering

The approach used in this thesis for handling the variance is called moving average filtering. This method is used for reducing the variance of a variable, by stating that the current value x_t an average of an arbitrary amount of the previous and future values. Increasing the window size would lower the variance between each sample, but at the cost of filtered variable not representing the initial value. Moreover, increasing the window size reduces the sample size, thus it is desirable to keep the window size low. This approach was utilized in this thesis by making each sample an average of the current and previous sample. The result of the technique can be seen in Figure 4.1b. The figure shows that the periods with high and low GDP growth are clearer due to the lower variance of the GDP growth.

4.6 Hypothesis Testing and P-values

The analysis seen in Chapter 5 utilizes the concept of hypothesis testing and p-values for validating a statistical relationship between variables. This technique is used to distinguish the observed effects between random chance and a true underlying effect in the data. A hypothesis test is performed by proposing two hypothesis H_0 and H_1 as seen in equations (4.7) and (4.8). The null hypothesis H_0 represents the scenario where there is no significant relationship between two variables, while the alternative hypothesis H_1 describes the scenario where there is a statistically significant relationship. When working with the linear equations seen in the previous sections of this chapter, the null hypothesis is interpreted as a coefficient being equal to zero, which means that the there is no relationship between the independent variable x_i and the dependent variable y. The alternative hypothesis on the other hand reflects the situation where the coefficient is non-zero, thus indicating a relationship between an independent variable and the dependent variable.

$$H_0 : \beta_i = 0 \tag{4.7}$$

$$H_0 : \beta_i \neq 0 \tag{4.8}$$

In order to test this hypothesis, it is common to use a T-test based on a student t-distribution. This involves comparing the test observator seen in equation (4.9) with the critical values of the student t-distribution under the assumption of H_0 being true. In other words, if the value of the test observator is greater than a given critical value, then the null hypothesis is rejected in favor of the alternative hypothesis. The critical values are typically based on the significance levels of 0.1, 0.05 and 0.01, where the critical values for the student t-distribution under a given amount of samples can be retrieved from standardized critical value tables.

$$TS = \frac{\hat{\beta}_i - \beta_i}{se(\beta_i)} \tag{4.9}$$

A p-value is a common statistic used for quickly determining the significance of a variable in a statistical test. The mathematical definition of it can be seen in equation (4.10), and it measures the probability of observing a sample T larger than your result t_0 given that the null hypothesis is correct. In more common words, it describes how likely it is to observe your estimate, or a more extreme estimate, given that the null hypothesis is correct. The p-value can then be used for determining if H_0 should be rejected at a certain significance level based on if the value is less than that value. For instance, given an analysis of H_0 on a 0.05 significance level, then a p-value less than 0.05 would result in the rejection of the null hypothesis in favor of the alternative hypothesis.

$$p\text{-value} = P(T \geq t_0 | H_0 = \text{true}) \tag{4.10}$$

5 Results

This chapter presents the empirical results from the multiple linear regression. It then provides a brief discussion on the robustness of the results related to the assumptions of the MLR seen in Section 4.4. Finally, the chapter ends by providing an interpretation of the numbers seen in the empirical results.

5.1 Empirical Results

The results were obtained by using the multiple linear regression in the programming language R. The dataset contains 30 observations, as the results are based on the time period from 1990 to 2019. The dependent variable, GDP growth, is smoothed by using a moving average filter, where the smoothed value is given by the average of the current and previous value. The smoothing is based on the GDP growth data from 1989 to 2019 in order to keep the 30 samples. The model used in the analysis is seen in equation (5.1). Here the service employment is denoted as x_{SE}, degree of openness is x_{Open}, foreign direct investments is denoted as x_{FDI} and the total employment is x_{TE}. Additionally, the intercept coefficient equals β_0.

$$\text{GDP Growth} = \beta_1 x_{\text{SE}} + \beta_2 x_{\text{Open}} + \beta_1 x_{\text{FDI}} + \beta_1 x_{\text{TE}} + \beta_0 \tag{5.1}$$

The rest of this section is dedicated to showcase multiple regressions based on a subset of variables to gain a better picture of which effects are relevant for the economic growth. This starts with a regression between GDP growth and the service sector employment, and ends with a full linear regression including all variables in equation (5.1). Each subsection will contain a table with the regression coefficients, and an interpretation of the coefficients. Each table contains an overview of the estimated coefficients, the standard deviation for the estimate, a p-value and the significance level for this variable. The significance level is denoted using stars. A single star indicates a 0.1 significance level, two stars indicates a 0.05 significance level and three stars indicates a 0.01 significance level.

5.1.1 Regression With Service Sector Employment

The first model is based on a regression between GDP growth in India and the employment in the service sector as seen in equation (5.2). The regression estimates can be seen in Table 5.1.

$$\text{GDP Growth} = \beta_1 x_{\text{SE}} + \beta_0 \tag{5.2}$$

Table 5.1 shows that the regression coefficient for the service employment is positive, meaning that an increase in employment in the modern sector leads to an increase in the GDP growth. Moreover, it is clear that an increase in the service employment by 1 percent point leads to a 0.31 percent point increase in the GDP growth. Additionally, it is shown that the interception coefficient equals to -3.67, which means that under a condition with no employment in the service sector, the economic growth in India would be equal to -3.67 percent. Furthermore, it can be observed based on the p-value for the regression coefficients that service employment is significant at a 0.01 level. This states that it is more likely than not that an increase in the service sector employment leads to an increase in the GDP growth in India. The R squared value of this model equals 0.39, which means that this model explains 39 percent of the variance in the GDP growth.

	Estimate	**Standard error**	**P-value**	**Significance**
Intercept	-3.67	1.96	0.07	*
Service Employment (x_{SE})	0.31	0.07	0.0002	***

Table 5.1: Linear regression between GDP growth and service sector employment.

5.1.2 Regression with Service Sector Employment and Openness

This regression expands the simple regression model in equation (5.2) by controlling the result for the openness of India, which is described mathematically by the expansion seen in equation (5.3). The new regression estimates are observed in Table 5.2.

$$\text{GDP Growth} = \beta_1 x_{\text{SE}} + \beta_2 x_{\text{Open}} + \beta_0 \tag{5.3}$$

Table 5.2 presents the results of a linear regression model showcasing the relationship between GDP growth and the predictor variables service employment and openness. The regression coefficient for service employment is positive and equal to 0.26, thus indicating that an increase in employment in the service sector leads to an increase in GDP growth. Specifically, a one percent point increase in service sector employment is associated with a 0.26 percent point increase in GDP growth. Similarly, the regression coefficient for openness is also positive and equal to 0.019, suggesting that greater openness in the economy is associated with higher a GDP growth rate. However, the coefficient for openness is smaller compared to service employment, indicating that the impact of service employment on GDP growth is more substantial. Additionally, the intercept coefficient is -3.85 percent, meaning that if both service employment and openness are zero, the expected GDP growth would be equal to -3.85 percent. The R squared value of this model equals 0.37, which means that this model explains 37 percent of the variance in the GDP growth.

Furthermore, based on the p-values for the regression coefficients, service employment is statistically significant predictors of GDP growth. Service employment has a p-value of 0.0372 and significant at a 0.05 level, which suggests that it is more likely than not that changes in service employment have a significant impact on GDP growth, while openness is insignificant.

	Estimate	Standard error	P-value	Significance
Intercept	-3.85	2.36	0.24	
Service Employment (x_{SE})	0.26	0.12	0.0372	**
Openness (x_{Open})	0.019	0.03	0.527	

Table 5.2: Linear regression of GDP growth based on service sector employment and openness.

5.1.3 Regression With Service Sector Employment and Foreign Direct Investments

This regression, similarly to the previous section, is an expansion of the simple regression model seen in equation (5.2). However, this model controls for the effects of foreign direct investments instead of controlling for the openness in the economy. This relationship is summarized mathematically in equation (5.4), while the regression result can be seen in Table 5.3.

$$\text{GDP Growth} = \beta_1 x_{\text{SE}} + \beta_2 x_{\text{FDI}} + \beta_0 \tag{5.4}$$

Table 5.3 presents the results of a linear regression model showcasing the relationship between GDP growth and the predictor variables service employment and foreign direct investments. The regression coefficient for service sector employment is 0.29, indicating that an increase in employment within the service sector is associated with an increase in GDP growth. Specifically, a one percent point increase in service sector employment corresponds to a 0.29 percent point increase in GDP growth. Conversely, the coefficient for foreign direct investments (FDI) equals 0.4. This suggests that higher levels of FDI are also associated with higher GDP growth. The larger coefficient for FDI compared with service sector employment can either indicate that the FDI value has more impact on the GDP growth, however it is more likely that this difference is due to the FDI being defined between 0 and 1.6, while service employment has a larger range. Furthermore, the intercept coefficient is -3.19, suggesting that given zero foreign direct investments and service employment, then the GDP growth would equal -3.19. The R squared value of this model equals 0.38, which means that this model explains 38 percent of the variance in the GDP growth.

In terms of statistical significance, service sector employment demonstrates significance at a 0.01 level, denoted by the triple asterisks, suggesting a strong relationship with GDP growth. However, the p-value for FDI is 0.48, indicating that it statistically insignificant in predicting GDP growth at conventional significance levels.

	Estimate	Standard error	P-value	Significance
Intercept	-3.19	2.08	0.14	
Service Employment (x_{SE})	0.29	0.08	0.0016	***
Foreign Direct Investments (x_{FDI})	0.4	0.55	0.48	

Table 5.3: Linear regression of GDP growth based on service sector employment and foreign direct investments.

5.1.4 Regression With Service Sector Employment and Total Employment

This regression, similarly to the previous two models, expands the simple regression model seen in equation (5.2). However, this model controls for the effects of the total employment which is summarized mathematically in equation (5.5), The regression result can be seen in Table 5.4.

$$\text{GDP Growth} = \beta_1 x_{\text{SE}} + \beta_2 x_{\text{TE}} + \beta_0 \tag{5.5}$$

Table 5.4 illustrates the findings of a linear regression analysis investigating the relationship between GDP growth and the two predictor variables service sector employment and total employment. The regression coefficient for service sector employment is 0.39, indicating that an increase in employment within the service sector is associated with an increase in GDP growth. Specifically, a one percent point increase in service sector employment corresponds to a 0.39 percent point increase in GDP growth. In contrast, the coefficient for total employment is 0.25 suggesting that a rise in total employment (across all sectors) is also linked to higher GDP growth. However, the coefficient for total employment is smaller than that of service sector employment, implying that the service sector employment has more impact on GDP growth compared to total employment in this analysis.

The intercept coefficient is -18.14, indicating that if both service sector employment and total employment are zero, the expected GDP growth would be -18.14. This value has a larger magnitude compared to the two previous models. This observation can be explained by the fact that the total employment has a range from 47 to 53, which when scaled with the regression coefficient has a range between 11.75 and 13.25. The same effect is true for the service employment which exists between 22 and 33 percent, and scaled with the regression coefficient has a range of approximately 8 to 13. Thus, the large negative value for the intercept value is due to the independent variables being defined quite far from zero. Finally, the R squared for this model equals 0.38, indicating that 38 percent of the variance is captured by this model.

In terms of statistical significance, service sector employment demonstrates significance at a 0.01 level, denoted by the triple asterisks, suggesting a strong relationship with GDP growth. However, the p-value for total employment is 0.44, meaning it is statistically insignificant in predicting GDP growth.

	Estimate	**Standard error**	**P-value**	**Significance**
Intercept	-18.14	18.6	0.34	
Service Employment (x_{SE})	0.39	0.12	0.004	***
Total Employment (x_{TE})	0.25	0.32	0.44	

Table 5.4: Linear regression of GDP growth based on service sector employment and total employment.

5.1.5 Full Regression

This final regression model includes all the control variables used in the previous regressions in order to control for all the effects when analyzing the effect of the service employment on the GDP growth of India. Table 5.5 shows the regression results.

	Estimate	Standard error	P-value	Significance
Intercept	-15.77	18.44	0.49	
Service Employment (x_{SE})	0.34	0.15	0.043	**
Openness (x_{Open})	0.02	0.04	0.55	
Foreign Direct Investments (x_{FDI})	-0.12	0.51	0.80	
Total Employment (x_{TE})	0.22	0.31	0.48	

Table 5.5: Multiple linear regression parameters.

Table 5.5 summarizes the results of a multiple linear regression analysis examining the relationship between GDP growth and several predictor variables: Service employment, openness, foreign direct investments (FDI), and total employment. The intercept coefficient is -15.77, indicating that when all predictor variables are zero, the expected GDP growth would be -15.77. The large negative value of the coefficient is explained by the displacement from zero seen in some variables as discussed in Section 5.1.4.

Among the predictor variables, service employment has a positive coefficient of 0.34. This suggests that an increase in service sector employment is associated with an increase in GDP growth. Meaning that a one percent point increase in service employment is associated with a 0.34 percent point increase in GDP growth. The coefficient for service employment is statistically significant at a 0.05 level, denoted by the double asterisks, indicating a significant relationship with GDP growth. Openness and foreign direct investments have coefficients of 0.02 and -0.12, respectively. The positive coefficient for openness indicates that a greater openness in the economy is associated with a higher GDP growth, while the negative coefficient for FDI indicates that a higher levels of foreign direct investments are associated with lower GDP growth. However, neither of these variables is statistically significant at conventional levels, suggesting that their effects on GDP growth may not be statistically distinguishable from zero. Additionally, total employment has a coefficient of 0.22, suggesting that an increase in total employment (across all sectors) is associated with higher GDP growth. However, similarly to openness and FDI, the coefficient for total employment is not statistically significant. Additionally, the regression resulted in an R-squared of 0.33, which means that the model explains 33 percent of the variance in the data.

In summary, among the variables considered in the analysis, only service employment demonstrates a statistically significant and positive relationship with GDP growth. The positive coefficient for service employment suggests that an increase in service sector employment is associated with an increase in GDP growth, holding other variables constant. Additionally, the coefficient for service employment remains in the range of 0.26 and 0.39 for all the multiple linear regressions, thus suggesting that the estimates for the service employment remains robust combined with different control variables.

5.2 Robustness Of The Full Regression Model

In order to analyze the robustness of the full regression results in Section 5.1, it is necessary to revisit the assumptions for the best linear unbiased estimator discussed in Section 4.4. The first three assumptions already holds given the use of the linear model in equation (5.1) and the periodic sampling of values. As seen in Section 4.4, it is necessary to check for an zero conditional mean, heteroskedasticity and multicollinearity to further strengthen the validity of the results.

5.2.1 Zero Conditional Mean

Under the assumptions of a zero conditional mean, it is assumed that the error averages out to zero. In this case the error has a mean of 0.1. The assumption of an average error of zero can be further analyzed by analyzing a QQ plot. The QQ plot shows the distribution of the residual errors relative to a standard normal distribution. If the assumptions of normality of the residuals holds, then the data should align in a straight line, covering the entire range of the QQ-plot. The QQ-plot for the model seen in 5.1 can be seen in Figure 5.1.

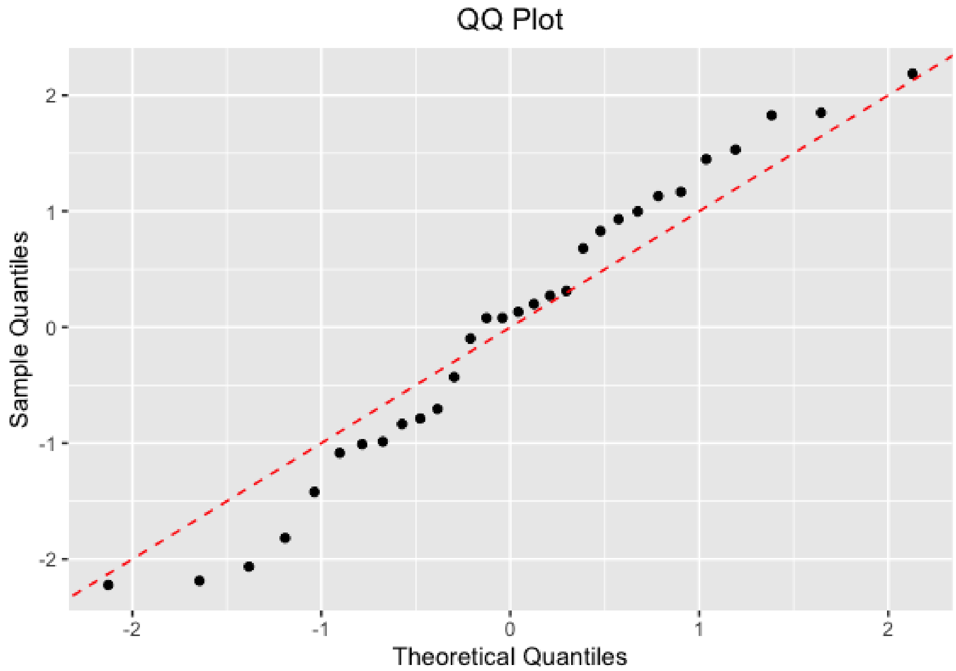


Figure 5.1: The QQ-plot for the regression model

The figure shows that the residuals almost make up a straight line, which suggests that the residuals are close to normally distributed. However, it can be seen that the residuals are above the line from about -0.5 to 2, while being below the line from -2 to -0.5. This aligns with the regression result stating that the average residual error computes to 0.1. This does not necessarily invalidate the regression result, even though the average residual error is non-zero. The dataset only contains about 30 samples, which the literature suggests is a too small sample size to obtain a perfectly normal distributed model. Thus, this line indicates an underlying normality in the residuals which might have had an average error closer to zero given a larger sample size.

5.2.2 Homoscedasticity

The homoscedasticity can be analyzed by analyzing the residuals of the estimated line for each fitted value. This relationship are showcased in Figure 5.2. Here the blue line represents how the residuals changes over time, the shaded grey area the variance and the dots is the residual value for a given fitted value of GDP. Based on the figure, the variance, marked with the shaded area, stays relatively even over time. However, the average residual, seen as the blue line, shows that the residuals oscillate around the zero line, with a maximum amplitude of about 0.5, but for the most part stays close to the zero line. Ideally, the blue line should stay completely flat for a truly normally distributed error. However, given that the dataset only contains about 30 samples, it could be argued that the sample size is too low to have an exact 0 error.

Since the variance stays relatively even across the fitted values, it can be concluded that the model holds the assumption of homoskedasticity.

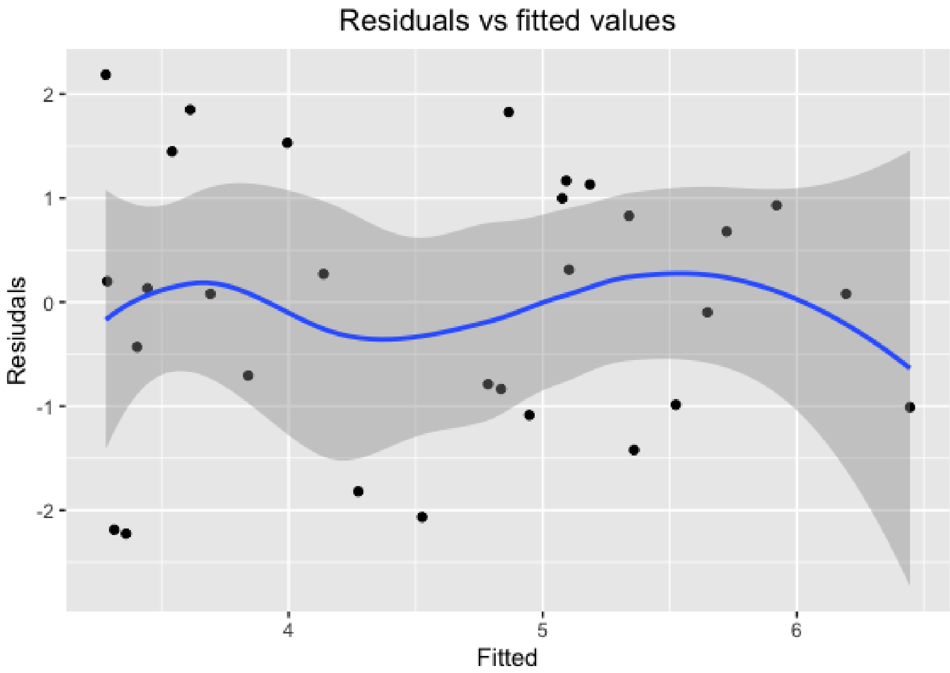


Figure 5.2: Residuals vs fitted values

5.2.3 Multi Co-Linearity

As discussed in 4.4, it is necessary to use a VIF test in order to ensure a low level of multi co-linearity. In this analysis the VIF test resulted in the coefficients seen in Table 5.6. Given that all tests results in a value less than 5, it can be seen that the result contains an acceptable amount of multi co-linearity.

	Service Employment	**Openness**	**Foreign Direct Investments**	**Total Employment**
VIF	4.34	4.59	3.22	2.69

Table 5.6: VIF test results

Additionally the correlation matrix for the four variables can be seen in Table 5.7. Here a positive coefficient means that the variables are positively correlated, and a negative variable indicates a negative correlation. Positive correlation means that when one of the variables increases, then the other variable tends to increase. Negative correlation means that when one of the variable increases, then the other variable tends to decrease.

	Openness	**FDI**	**Service Employment**	**Total Employment**
Openness	1.0	0.82	0.78	-0.61
FDI	0.82	1.0	0.66	-0.46
Service Employment	0.78	0.66	1.0	-0.78
Total Employment	-0.61	-0.46	-0.78	1.0

Table 5.7: Correlation matrix.

5.3 Results

Among the independent variables we utilized, service employment was the only one that was statistically significant. This could indicate that structural changes in the economy is accompanied by economic growth, as predicted by Lewis (1954). While the Lewis model has typically been used to describe the transformation from an agricultural economy to an industrial one, our findings provide evidence that its overarching growth theory still holds true for the service-based economy of India.

We find that foreign direct investment is not statistically significant. This contradicts the findings in Acharyya (2009). Although it should be noted that our choice of time period is different, there is still some overlap, with our research investigating the period 1990 to 2020, and that of Acharyya (2009) looking at 1980 to 2004. The lack of significance of FDI could mean that the exogenous increases in capital needed to kickstart the economic growth happened prior to the time period we chose to look at. Alternatively, it could be because the effects of FDI is difficult to measure without adding a lag variable.

Degree of openness was not statistically significant in our analysis. Since exports make up a relatively low amount of GDP in India compared to many other developing countries (World Bank and OECD, 2021), this was not surprising. While trading with other countries generally provides positive outcomes, it does make them reliant of global supply and demand. Having a low dependence on trade could therefore potentially be beneficial during turbulent times.

Total employment was used to control for changes in employment, since an increase in the size of the workforce could yield higher production. However, this variable was not statistically significant, which indicates that the productivity of the group of workers that have entered and left the workforce is overall low compared to the rest. This could be a sign of the low marginal productivity in agriculture described by the Lewis model.

The appendix contains two regressions that were done using employment in agriculture and combined employment in industry and service as independent variables. These two variables are the inverse of each other. They are statistically significant, but less so than service employment. This reinforces our assertion that it is employment in the service sector in particular that is tied to economic growth, not just non-agricultural jobs in general.

5.4 Limitations

In our thesis we have experienced limitations in our choice of time period and in our dependent variable. The rest of the chapter is dedicated to analyzing how this has affected the results.

5.4.1 Method

A multiple linear regression model is used in the thesis, and the model works well when only observing economic trends over a time period but is less responsive to economic shock as the finance crisis in 2008 or Covid-19 pandemic. This is the reason we chose to use the period 1990 to 2019 because it reduced the impact from exogenous shock like the pandemic and made it easier to observe economic trends. Simultaneously, the difficulty of finding data before 1990 is something that might have led to empirical weakness, the reason is that with a small sample size it could be difficult to draw conclusion from the results. At the same time, the thesis looks at the uptick in economic growth in the millennial shift, and by using numbers from before and after 2000, we could observe the economic trends. We started with a period that went from 1960 to 2022, but due to the lack of data and economic shock had to shorten this period to 1990-2019.

The choice of the main independent variable can also be interpreted as a limitation of the regression. The results were based on utilizing the service sector employment as the dependent variable for modelling the structural change from the agriculture to the modern sector. However there exists other approaches for modelling the structural change, which is explored in more details in Appendix A

5.4.2 Causality and Omitted variables

While our research presents a clear correlation between economic growth and structural change in India, it does not show the direction of causality. It could well be that the economic growth of India is driven by increase labour in the service sector, although the relationship could also be inverse. For example, one could speculate that the increase in employment within the service sector is a consequence of increased domestic demand for services. If this is the case, the direction of causality would be opposite to that of the Lewis model. Furthermore, an economic shock as great as a global financial crisis carries implications too wide to be explained in a single paper, and likely has had a large exogenous effect on some of the variables. This applies especially to foreign direct investment and degree of openness, as those are greatly affected by foreign conditions. While data on capital formation was easily obtainable, we did find a measure that separated it by country of origin. Another omitted variable that possibly could have affected our results is the level of investments of domestic origin. We also do not control for the the rate of reinvestment, as this data was difficult to acquire.

5.4.3 Generalization

As we have established a link between employment in service and economic growth in India, it may be possible that the positive results associated with expanding the service sector could be replicated elsewhere. However, it should be noted that the development of India is somewhat unique, and has not yet been replicated in a lot of other developing countries. Nevertheless, Bangladesh and Pakistan seem to have developed large service sectors (World Bank, 2024b). Due to their geographical proximity and similarities to India, our results may be relevant to these countries as well. The conditions that have allowed for India to take this path to development are not necessarily present elsewhere.

5.4.4 Source Criticism

Most references used in the thesis are from peer-reviewed journals and books. The thesis also uses some newspaper articles, and it is important to acknowledge that even though most of the articles came from reputable newspapers the information from these is not peer-reviewed. In the thesis we chose to use the article from Arthur Lewis to understand the model. We will argue that this gives a better understanding of the Lewis model and supplement the theory, using secondary source that explains the Lewis model. The dataset used in the thesis is sourced from the development indicators from the World Bank, and even though some data from before 1990 is missing the reliability of the dataset we will argue is strong because the data came from the World Bank.

6 Conclusion

In this thesis, we have sought to understand the effects of structural change on the economic growth of India during the last three decades. To do this, we used a deductive method where we tested the strength of the Lewis model. The Lewis model predicts that a country with an unlimited supply of labour will achieve growth as workers move from agriculture to modern sector. This rests on the assumption that the profits from the modern sector are reinvested back into the economy. Although the Lewis model is typically used to describe the transition from an agricultural economy to industrial economy, we use it to examine the development of India towards a service-based one.

In the limitations discussed in Section 5.4 we address some concerns and obstacles with the data and the data process in the thesis. First, the lack and difficulty of retrieving data before 1990, which might have led to some empirical weaknesses of the thesis. Secondly, we have chosen to not address the Lewis turning point due to the lack of observations on India's wage levels. We conclude that the framework from the Lewis model can be used to explain the structural changes in India. It should still be noted that that we were not able to confirm if all they assumptions that underpin the Lewis model.

Finally, we answer the research question: Can structural changes from the agriculture sector to the service sector contribute to economic growth? By using the framework from the Lewis model in numerous multiple linear regression analysis's we were able to demonstrate a positive correlation between the independent variable service sector and dependent variable economic growth. This result was also statistically significant at 0.05 level. At the same time our analysis showed that the other independent variables, degree of openness, total employment, and foreign direct investment was not significant in the analysis.

Further research should include a broader set of data with a larger quantity of samples. A larger sample size could not only improve the data itself, but also make it easier to observe long-term trends in an economy, as well as economic growth. In the future it could be fruitful to use a similar method to observe other developing countries, in order to compare the strength of the Lewis model framework. This could reveal if other countries in South Asia follow the same pattern of development as India. Additionally, by collecting data on agricultural productivity in different regions and within individual communities, one could also attempt to see if there is a relationship between labor migration and wage rates in modern sectors.

References

Acharyya, J. (2009). Fdi, growth and the environment: Evidence from india on co2 emission during the last two decades. *Journal of Economic Development*, *34*(1), 43–58 https://EconPapers.repec.org/RePEc:jed:journl:v:34:y:2009:i:1:p:43-58.

Ahluwalia, M. S. (2002). Economic reforms in india since 1991: Has gradualism worked? *The Journal of Economic Perspectives*, *16*(3), 67–88 http://www.jstor.org/stable/3216951.

Basu, K. & Maertens, A. (2007). The pattern and causes of economic growth in India. *Oxford Review of Economic Policy*, *23*(2), 143–167, https://doi.org/10.1093/oxrep/grm012.

Bayly, C. A. (1985). State and economy in india over seven hundred years. *The Economic History Review*, *38*(4), 583–596, https://doi.org/10.1111/j.1468-0289.1985.tb00391.x https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0289.1985.tb00391.x.

Bhalla, S. S. & Bhasin, K. (2024). India eliminates extreme poverty. https://www.brookings.edu/articles/india-eliminates-extreme-poverty/?fbclid=IwZXh0bgNhZW0CMTAAAR0kkQP3aNOE8Q4Wa3d5rpKzR-0n4rEtgjMTwjiv5GYk46NbWwfHtvIeEJM_aem_AWU8uPLyxK6M1IbO9wJ8tvzExUaY6u2jjY4dAA9WNvmgdc3TCKntZ9YBbQYBemwOBmJ8DbeQ5wsDEukO_mc7hJ88. Accessed: 24/03-2024.

Bosworth, B. & Collins, S. M. (2008). Accounting for growth: Comparing china and india. *The Journal of Economic Perspectives*, *22*(1), 45–66 http://www.jstor.org/stable/27648223.

Cowen, T. (2024). Lesson from old india. https://www.nytimes.com/2014/08/24/upshot/lesson-from-old-india-when-an-economy-just-doesnt-get-better.html. Accessed: 25/04-2024.

Das, M. & N'Diaye, P. (2013). Chronicle of a decline foretold: Has china reached the lewis turning point? *IMF Working Papers*, *13*, https://doi.org/10.5089/9781475548242.001.

Ghose, A. K. (2021). Structural change and development in india. *Indian Journal of Human Development*.

Holden, S. (2016). *Makroøkonomi*. Oslo: Cappelen Damm akademisk.

IMF (2009). Balance of payments and international investment position manual. (6).

Islam, N. & Yokota, K. (2008). Lewis growth model and china's industrialization. *Asian Economic Journal*, *22*(4), 359–396 https://EconPapers.repec.org/RePEc:bla:asiaec:v:22:y:2008:i:4:p:359-396.

Joshi, S. (2004). Tertiary sector-driven growth in india: Impact on employment and poverty. *Economic and Political Weekly*, *39*(37), 4175–4178 http://www.jstor.org/stable/4415534.

Knight, J. (2021). A Tale of Two Countries and Two Stages: South Africa, China and the Lewis Model. *South African Journal of Economics*, *89*(2), 143–172, https://doi.org/10.1111/saje.12280.

Laitner, J. (2000). Structural change and economic growth. *The Review of Economic Studies*, *67*(3), 545–561 http://www.jstor.org/stable/2566966.

Laker, B. (2024). India Will Grow To Become The World's Third-Largest Economy By 2027. https://www.forbes.com/sites/benjaminlaker/2024/02/23/india-to-become-third-largest-economy-by-2027-implications-for-leaders/?fbclid=IwZXh0bgNhZW0CMTAAAR2Ej9QGBTNQj2YNKOQDzgCGEKXfj7vNRbcNEZN-HcbXKXx2FGz3ufVvHyk_aem_AWVqnPWYsBJQysolBjI6k1nR4-aD5KQ1L-014oId3VzBqgAdjslQyxe_5vrTjFhaQMMyPROFNrqMd7bqpE2-6Osr&sh=13be47d44fd5. Accessed: 22/04-2024.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, *22*(2), 139–191, https://doi.org/10.1111/j.1467-9957.1954.tb00021.x.

Mackie, A. B. (1964). The role of agriculture in economic growth and development. *Illinois Agricultural Economics*, *4*(3), 1–10 http://www.jstor.org/stable/1349016.

Mukherji, R. (2008). The political economy of india's economic reforms*. *Asian Economic Policy Review*, *3*(2), 315–331, https://doi.org/10.1111/j.1748-3131.2008.00118.x.

Roland, G. (2016). *Development economics*. Routledge.

Travelli, A. (2024). India foreign direct investments. https://www.nytimes.com/2024/01/02/business/india-economy-foreign-direct-investment.html. Accessed: 25/04-2024.

Unni, J. & Naik, R. (2011). Rural structural transformation: The case of the services sector in india. *Economic and Political Weekly*, *46*(26/27), 196–200 http://www.jstor.org/stable/23018646.

Wooldridge, J. (2019). *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning.

World Bank (2017). India poverty line. https://pip.worldbank.org/country-profiles/IND. Accessed: 25/04-2024.

World Bank (2024a). Agriculture employment. https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS?locations=IN&fbclid=IwZXh0bgNhZW0CMTAAAR05NvLCIo0w83y23tvjYH800bM1t5-bfm4TCsDR1VUk5wkJeOj3EfzRg2A_aem_AYcPr4mELP_rQsTL5z715tk98ffV9zYqJWTLs-2QdQUfIA1NhomOW-B8qMtGLJS7TLI9KaymRfzuH95gJtm041VS. Accessed: 13/05-2024.

World Bank (2024b). Service value added statistic. https://data.worldbank.org/indicator/NV.SRV.TOTL.ZS?most_recent_value_desc=true. Accessed: 13/05-2024.

World Bank (2024c). World development indicator. https://databank.worldbank.org/source/world-development-indicators. Accessed: 24/03-2024.

World Bank and OECD (2021). Trade as a share of gdp. https://ourworldindata.org/grapher/trade-as-share-of-gdp. Accessed: 13/05-2024.

Wu, Y., Guo, X., & Marinova, D. (2017). *Productivity, Innovation and China's Economic Growth*, volume 2, (pp. 213–228). ANU Press http://www.jstor.org/stable/j.ctt1trkk3v.16.

A

Alternative independent variables

The thesis showcased that structural changes in the labor market lead to a higher sustained GDP growth over a long period of time. During the analysis the relative employment in service was used for modelling the structural change from agriculture into the modern sector. This appendix explores alternative approaches for modelling the structural change in the Indian economy.

Agriculture as an independent variable

Introducing agriculture as the independent variable results in the model seen in equation with x_{AE} denoting the agriculture employment. This models how the GDP growth depends on a reduction in the agriculture sector over time.

$$\text{GDP Growth} = \beta_1 x_{\text{AE}} + \beta_2 x_{\text{Open}} + \beta_1 x_{\text{FDI}} + \beta_1 x_{\text{TE}} + \beta_0 \tag{A.1}$$

Reusing the method seen in Chapter 5 results in the regression result seen in Table A.1. Here it is clear that agriculture employment has a negative coefficient of -0.46, indicating that a one percent point increase in agriculture reduces the GDP growth rate with 0.46 percent point. In other words, the decreasing agriculture employment seen in Figure 4.3a correlates with an increase in GDP growth. The agriculture employment is significant at a 0.1 significance level.

	Estimate	**Standard error**	**P-value**	**Significance**
Intercept	-0.46	15.56	0.97	
Agriculture Employment (x_{SE})	-0.16	0.09	0.0911	*
Openness (x_{Open})	0.01	0.04	0.77	
Foreign Direct Investments (x_{FDI})	-0.06	0.53	0.90	
Total Employment (x_{TE})	0.27	0.37	0.48	

Table A.1: Multiple linear regression parameters with employment in agriculture as the independent variable.

Service and industry as an independent variable

Combining the service and industry employment into a single variable can be another approach for modelling the service sector. Applying this approach yields equation (A.2), with x_{SIA} for denoting the service sector plus industry employment.

$$\text{GDP Growth} = \beta_1 x_{\text{SIA}} + \beta_2 x_{\text{Open}} + \beta_1 x_{\text{FDI}} + \beta_1 x_{\text{TE}} + \beta_0 \tag{A.2}$$

Given the same approach as in the last section results in the regression result seen in Table A.2. It is worth highlighting that these regression coefficients are quite similar to the result observed in Section A.

This phenomenon occurs due to the combination of the service and industry employment is the inverse of the agriculture employment. Thus, the two models differ mostly in the interception coefficient and the sign of the main independent variable. Here a one percent point increase in the combined service and industry employment results in a 0.46 percent point increase in GDP growth. The aggregated independent variable is significant at a 0.1 significance level.

	Estimate	**Standard error**	**P-value**	**Significance**
Intercept	-16.9	21.86	0.44	
Service and Industry Employment (x_{ME})	0.16	0.09	0.0911	*
Openness (x_{Open})	0.01	0.04	0.77	
Foreign Direct Investments (x_{FDI})	-0.06	0.53	0.90	
Total Employment (x_{TE})	0.27	0.37	0.48	

Table A.2: Multiple linear regression parameters with the combination of service and industry employment as the independent variable.

NTNU

Norwegian University of
Science and Technology