SCHOOL OF ECONOMICS AND MANAGEMENT

Another Lunatic Line? An Economic Evaluation of the Standard Gauge Railway in Kenya

By

Hedda Most

January 2021

Bachelor Programme in Economics

School of Economics and Management

Department of Economics

Lund University

Supervised by Gunes Gokmen

Abstract

This research paper seeks to examine the emerging role of transportation infrastructure on economic activity in low-income countries due to its vital part in facilitating deeper integration of markets, hence fostering inclusive growth and prosperity for the poor. Although its economic causal framework suggests ambiguous results, transportation investments account for the largest share of the World Bank's lending to developing countries. Through an inductive approach, the paper assesses the new Standard Gauge Railway's effects on income in Kenya by looking at regional disparities induced by the railway. This empirical research adopts the Differences-in-Differences estimation and identifies an increase in income attributed to both the construction of the railway and the railway being in operation. When exploring heterogeneity in the data, further significant results disclose that young men residing in rural areas are identified as main economic beneficiaries of the railway. Moreover, additional results on potential channels of the income relationship advance the analysis and provide a versatile picture of the new railway's role in the economy. Linkages to the heterogeneity analysis help explain the results and demonstrate that people living along the tracks have altered their means of gaining an income due to the railway. Together these results provide important insights into the vital role transportation plays in accelerating income levels and tackling the broader challenge of alleviating global poverty.

Keywords: Transportation Infrastructure, Railways, Kenya, Differences-in-Differences, Income, Economic Growth

Acknowledgements

I would like to express my gratitude to my supervisor Gunes Gokmen for his valuable and constructive suggestions throughout the process of conducting this paper. His willingness to give his time so generously has been highly appreciated, and his guidance has served as useful critiques to successfully carry out this research.

Table of Contents

List of Tables

List of Abbreviations

BRI	Belt and Road Initiative
DD	Differences-in-Differences
FSD	Financial Sector Deepening Kenya
KeNADA	Kenyan National Bureau of Statistics' National Data Archive
KNBS	Kenyan National Bureau of Statistics
SGR	Standard Gauge Railway
NASSEP V	National Sample Survey and Evaluation Program

1. Introduction

Economic growth is considered the most powerful tool for reducing poverty and improving living standards in low-income countries due to its nature of generating cycles of prosperity and opportunities that advance human development (DfID, 2021). Reaching sustained levels of economic growth, Solow (1956) emphasises the pivotal roles of capital accumulation, technological progress, and population growth. He suggests that these three components accelerate growth rates in developing countries and induce a shift in the economy, which alters most dimensions of people’s lives. A key aspect of increased growth levels related to physical capital is the improvement of transportation infrastructure, which has a vital part to play in fostering inclusive growth and boosting prosperity for the poor (World Bank, 2021). Many low-income countries are currently investing billions of dollars in improving transport infrastructure, demonstrated by the fact that about 15 % of the World Bank’s lending goes into transport projects – exceeding the Bank’s lending share to any other sector (World Bank, 2020). Hence, the studying of transport infrastructure’s role in emerging economies plays a fundamental part in understanding how to carry out successful strategies to mitigate poverty.

A considerable amount of literature has been published recognizing the critical role of transportation infrastructure in generating economic activity. However, the exact causal economic effects of transport infrastructure are disputed, especially in emerging economies. On the one hand, Banerjee, Duflo, and Qian (2020) argue that adequate transport infrastructure is considered a prerequisite for achieving sustained growth since it enables communication and facilitates trade creating deeper integrating markets. On the other hand, the authors also recognize a possible scenario where human and physical capital exit remote areas subject to improved transportation possibilities due to cities' well-known agglomeration effects, hence risking these areas' impoverishment. Contrasting this potential scenario, Baum-Snow, Brandt, Henderson, Turner, and Zhang (2017) find proof that railways decentralize industrial production and its labour force out of the cities possibly enhanced by China’s substantial historical reliance on railways for transporting freight. Thus, less land-intensive activities that profit from agglomeration effects, such as tradable services, benefit from this decentralisation. The dubious economic causal framework of transportation infrastructure stresses the need for further research and analysis.

Most studies in the field of railway infrastructure's economic effects have approached the subject from a macroeconomic perspective, thus neglecting critical heterogeneity. Moreover, few studies have investigated the regional disparities induced by railways in a sub-Saharan setting. Understanding how efforts to acquire sustained growth levels in low-income countries requires comprehensive assessment and grasping these efforts' impact on all societal levels to tackle the broader challenge of alleviating poverty. Hence, there is a great need to fill the gap of uncertainty regarding the railway's economic role in emerging economies.

This paper seeks to examine the emerging role of transportation infrastructure on economic activity in a developing context. It is done through an inductive approach to the subject in which I assess the economic effects of the new Standard Gauge Railway (SGR) in Kenya. This approach to empirical research adopts the Differences-in-Differences (DD) method. As stated by Angrist and Pischke (2009), the DD method is based on a counterfactual framework, where changes in the trend of the dependent variable are solely attributed to the existence of an event affecting only a part of the population (i.e., the treatment group), while the trend of the non-affected population (i.e., the control group) remains unchanged. The microdata used is aggregated to the county level to allow for comparison of income between the counties that construct/have a station along the railway (i.e., the treatment group), and the rest of Kenya's counties (i.e., the control group).

This study's unique objective of investigating the Standard Gauge Railway's economic effects in Kenya through a quantitative research design proves its significance and demonstrates its novelty. The research question is composed as "How has the Standard Gauge Railway affected the economic activity along its tracks in Kenya?" The hypotheses arising from the research question proceed as follows: H_1: the income level has risen in counties that have a station, as a result of constructing the railway, and H_2: the income level has risen in counties that have a station, as a result of the railway being in operation.

The main results suggest a positive and significant relationship between income and the railway. It shows that the railway construction induced an increase in income, potentially resulting from greater employment opportunities. Simultaneously, the occurrence of an existing and operating railway also induced an increase in income, potentially stemming from further integrated markets and improved communication and mobilization. Thus, the relevance of assessing the economic influences of railways is clearly supported. When performing a heterogeneity analysis comparing gender, cluster type, and age, the results appoint the new railway's primary beneficiaries as men below 45 years old residing in rural areas. However, significant results of increased income for women can also be depicted from the results implying a generally increased level of economic activity attributed to the railway. Moreover, I find that people residing in urban areas have earned a lower income due to the railway being in operation, suggesting a potential scenario where the economic activity decreased in urban areas due to loss of employment accompanying the completed construction.

Furthermore, the analysis' aims to identify the drivers of the outcomes depict that a potential shift in livelihood sectors has taken place due to the railway being in operation, which in turn could explain the rise in income levels. I find that the possibility of earning money from farming and wage employment decreased, while increased for self-employment and casual work in the areas close to the railway. This implies a shift in individuals' weighting of the different sectors they rely on gaining income from. Comparing these results with the results extracted from the heterogeneity analysis, further linkages are displayed and imply a shift of livelihood activities to more formal ways of earning an income, especially in rural areas possibly explaining the main findings of higher income induced by the railway.

Nevertheless, the limitations of this study are many. I want to stress the study's inability to encompass the long-term effects of the SGR due to the recent date of the construction, implying that all effects have not yet been displayed. Moreover, the results should be interpreted in the country context and setting of Kenya and its demographics. Finally, it is beyond this study's scope to examine the SGR's effects on trade flows. Further studies, which take these concerns into account, will need to be undertaken to illustrate the aftermath of the SGR in Kenya.

The rest of the thesis has been organized in the following way: first, chapter two gives a brief overview of the recent history of railways in Kenya, followed by a review of the theoretical framework and previous research in chapter three. The fourth chapter is concerned with the methodology and data of the research paper. Chapter five presents and explains the findings of the research, while chapter six concludes the thesis.

2. Historical Context

Research on Kenya's demographic history by Jedwab, Kerby, and Moradi (2017) describes a country with widespread poverty during the late 19th century, considering that except for Nairobi (founded in 1899), in 1901, there were only four cities, all of which were located along the coastline. They suggest that high trading costs constrained economic activity and trade between these coastal cities and neighbouring land-locked countries. Accordingly, the British colonial power constructed a new railway between 1896 and 1901, connecting the big shipping port in Mombasa to Uganda, hence the name Uganda railway. According to their research, the British had obvious geopolitical incentives for constructing the railway since it unified their colonies and connected them to the big port in Mombasa, thereby increasing their exports and production volumes. However, not all of that time's current leaders were wildly convinced of the new railway's utility. One notably displeased voice of the British parliamentarians was Henry Labouchere, who wrote a satirical poem to express his concern regarding the railway:

> What it will cost no words can express,
> What is its object no brain can suppose,
> Where it will start from no one can guess,
> Where it is going to nobody knows,
> What is the use of it none can conjecture,
> What it will carry there’s none can define,
> And in spite of George Curzon’s superior lecture,
> It clearly is naught but a lunatic line (Spicer, 2010, p.17).

In retrospect, the railway proved to induce several positive outcomes. Mombasa's port was modernised and became East Africa's gateway (Jedwab, Kerby & Moradi, 2017). Moreover, the function of the railway being in operation did not just generate increased economic activity. Hillbom and Green (2019) present empirical evidence that the railway's actual construction, more specifically the labour force, created maintenance clusters along the route, which grew to become towns, one of which became Kenya's capital Nairobi. As they recognise the railway's critical role in creating clusters, they also describe the railway's fall into disrepair after Kenya's independence in 1963 due to mismanagement and lack of maintenance. In the same vein, Jedwab, Kerby, and Moradi (2017) find that although the Uganda railway was built to satisfy colonial trading routes, instead of benefitting the already existing trade routes African communities had built up pre-colonial times, the railway's decreased number of operations abated Kenya's economic development vastly. Once again, the purposes of the railway to reduce transport costs and generate economic activity receded.

Following Kenya's independence in 1963, the first decades experienced rapid growth rates and praised the country's potential becoming one of Africa's success stories (World Bank, 2021). However, as Gil-Alana and Mudida (2017) described, the political instability entrenched in the post-independence era the last 50 years caused lacking inclusivity of political institutions that generated an unsolid foundation for the country's economic institutions. Accordingly, this weakened the accompanying economic growth. To address the low growth rate and to transform Kenya into an industrialised middle-income country by 2030, the Kenyan government launched a development program named Kenya Vision 2030 (vision2030, 2021). The construction of the Standard Gauge Railway (SGR) portrayed a flagship project of this vision and was built to run parallel to the dilapidated Uganda railway, thus benefiting Kenya and East Africa's economy in general (Taylor, 2020).

The Standard Gauge Railway was initiated under China’s Belt and Road Initiative (BRI) and is viewed as a vital contemporary element in a wider expansion of Chinese involvement in Africa according to Taylor (2020). The SGR project is Kenya’s biggest infrastructure investment since independence, and the project relies on a Chinese loan amounting to roughly US$3.2 billion, 6% of Kenya’s GDP (Kacungira, 2017). Simultaneously, it is reported that jobs were created for Kenyan citizens directly through the construction and maintenance of the SGR. Furthermore, Taylor (2020) reported the line's aspirations to expand to the landlocked neighbouring countries such as Uganda, Rwanda, and South Sudan, thereby connecting them to the big port in Mombasa and the under-construction port in Kisumu (also Kenya).

According to Taylor (2020) the process of building the railway has been constructed during two phases: Phase 1, routing from Mombasa to Nairobi, where the construction work began in December 2014, and the railway became operational in May 2017; and Phase 2, routing from Nairobi to Malaba, where the construction work began in 2016, and the railway became operational for passengers in 2019. Furthermore, he reports a farther division of the route, including three sub-phases: Phase 2a: Nairobi-Naivasha, Phase 2b: Naivasha-Kisumu, Phase 2c: Kisumu-Malaba. Phase 2 of the SGR opened up passengers' transportation in June 2019

(Salaudeen, 2019). Phase 1 of the SGR operates two different passenger trains: The Intercity and the County Trains. The Intercity travels between Nairobi and Mombasa without stop while the County Train stops at six intermediate stations along the route (Athi River, Emali, Kibwezi, Mtito Andei, Voi, Miasenyi, and Mariakani) (Taylor, 2020).

The expected positive outcomes of the SGR are many, but the project has also faced drawbacks and criticism. On a positive note, Taylor (2020) writes that the SGR expects to reduce the journey time for passenger trains routing from Nairobi to Mombasa from ten to four hours, while the cargo trains are expected to finish the route in eight hours. Moreover, the hope is to generate greater exchange within East Africa through vastly lower transportation and production costs. On the other hand, he argues that many drawbacks of the railway have been revealed, including obstacles to transport the cargo from the port onto the SGR due to malfunctioning cranes, hence hindering the cargo from being loaded onto the SGR. Besides, he reports that concerns have been raised that businesses relying on cargo and passenger traffic transported by road will be undermined by the new SGR while simultaneously, informing that road transport of cargo is often still preferred over the SGR due to logistical barriers. However, no concrete evidence on this matter has been provided yet. Lastly, the BRI project in itself has been widely debated and questioned for its actual purpose. On the one hand, it is viewed as creating a new global growth engine connecting the world further (Cohan, 2017). However, on the other hand, sceptics have interpreted the project as a plan for a Sino-centric international trade network and a disguised investment of buying political influence (Wo-lap Lam, 2016).

Amin (1974) advocates for a strategy of self-reliance for developing countries, in opposition to relying on investments from major economies, such as China. He suggests that this type of dependency on foreign capital investments causes structural distortions or disarticulation of the country's economy. Building on similar arguments, Shivji (2009) highlights a scenario where Africa exhibits a disarticulation between production structure and consumption structure. He refers to the term 'structural disarticulation' as a situation where the produced quantities are not consumed, and the consumed quantities are not produced. Building on both these theories, Taylor (2020) criticises the SGR project by pointing out the SGR's inability to transport balanced amounts of imported and exported cargo, causing a trade deficit. He emphasises the railway's ability to accommodate for importing Chinese goods but not for exporting Kenyan goods.

As outlined throughout this chapter, the SGR project is highly contentious and has raised many concerns. It is mainly questioned for its validity to transport cargo and the skewed power balance between China and Kenya due to the latter's significant debt. However, as emphasised before, this paper focuses its research on the short-term economic effects on the income of the people living along the tracks of the SGR.

3. Theoretical Framework and Previous Research

3.1. Theoretical Framework

This section aims to provide a conceptual framework in which the Standard Gauge Railway's influences on the Kenyan economy can be assessed. The main theoretical basis to address this question includes Solow's (1956) theory of economic growth, focusing on capital accumulation and technological progress, which reasonably applies to transportation infrastructure. Moreover, it includes the framework developed by Banerjee, Duflo, and Quian (2020) that assesses the causal effects of improved transportation infrastructure on regional Chinese growth rates since this allows for exploring certain heterogeneity in my analysis.

In the literature of economic growth, Solow's (1956) traditional neoclassical growth theory constitutes the field's core. He argues that output growth stems from three main components: population and labour force growth, capital accumulation, and technological progress. Moreover, he suggests that economies reach convergence on output per worker, conditioned that countries have the following identical parameters: savings rate, rate of population growth, depreciation, and productivity growth (technology). This suggestion illustrates that, when assuming identical parameters between several countries, all countries will eventually end up at the same development level, even without international development assistance. Even though some economists argue that the model does not include necessary aspects (i.e., human capital, spillovers between firms) and is more fitting for a high-income rather than a low-income country, it remains an important reference point in the literature of growth (Todaro & Smith, 2011).

Solow's (1956) theory of capital accumulation and technological progress applies to transportation infrastructure. He refers to capital accumulation as the increase in a country's stock or real capital and includes all new investments in land, physical equipment, and human resources. He suggests that future expanded output levels can be achieved by investing in the present physical capital stock. In consonance with this theory, infrastructure investments can be seen to improve the productivity of already existing resources within a country. Technological progress, on the other hand, Solow (1956), refers to as the increased application of new scientific knowledge that generates improved ways of accomplishing traditional tasks concerning physical and human capital. This economic growth component is often viewed as an essential source. More specifically, he argues that capital-augmenting technological progress suggests more productive use of existing capital goods – e.g., building new railways for high-speed trains. Investing in these two critical sources of economic growth yields improvement of existing physical capital and resources, and the combination of the two serves as the prime catalyst for economic growth in emerging economies (Todaro & Smith, 2011).

This paper's conceptual framework could also be likened to the framework used by Banerjee, Duflo, and Qian (2020), where labour is assumed to be immobile and capital less mobile than goods, resulting in that even remote areas persist in holding onto a part of their capital. Besides, they assume more limited mobility of capital in remote areas, indicating that improved transport infrastructure will affect GDP per capita in two contending directions. On the one hand, longer distances are costly, making exports less profitable. On the other hand, assuming that remote areas retain more capital per head than better-connected areas, this ratio boosts labour productivity in those areas. This results in relatively small differences in GDP levels between remote and better-connected areas indicating non-existent, or close to, differences in growth rates.

3.2. Previous Research

The importance of methodically approaching the critical question how low-income countries can grow and prosper has recently been accentuated by the Nobel Laureates Banerjee, Duflo, and Kremer, whose work has altered the field and focus of development economics (Olken, 2020). The emphasis that focuses should be shifted to evaluating development challenges through micro empirical studies illustrates a gap in the literature. Concurrently, the literature on the railway's role in the economy has highlighted several causal effects calling for ambiguous conclusions.

Banerjee, Duflo, and Qian (2020) research the causal effects of improved transportation infrastructure on China's economic growth. They point out transportation infrastructure as a key element for sustained economic growth relying on the logical need for accessibility to markets and ideas to benefit from them. However, their results suggest that the distance between regions and transportation networks has not affected income growth during the twenty years studied (1986 - 2006), even though regions closer to historical transportation networks have higher GDP per capita. Although, they emphasise the possibility that increased transport infrastructure benefits can be found in the economy as a whole, rather than regional localisation of the gains due to factor immobility. More closely explained, the lack of factor mobility prevents the gains from being concentrated in relatively better-connected areas.

A different approach has been adopted by Baum-Snow et al. (2017), who look at the contributions of transportation infrastructure on the decentralisation of population and industries from central cities to suburban regions in China since 1990. They find strong evidence for a decrease in population density in central cities derived from the emergence of radial highways. Though no proof is found that radial railways influenced population allocation, they find proof that the railways decentralised industrial production and its labour force out of the cities, possibly enhanced by China's historical reliance on railways for transporting freight. This result implies potential productivity gains from freeing up space for activities requiring land abundance, such as tradable services, who naturally benefit from cities' agglomeration effects. Moreover, they conclude that the railway infrastructure affected employment distribution in terms of sector and location by providing data on declining rates of manufacturing employment in central cities induced by the railway.

Storeygard's (2016) results add an essential dimension to the research of linkages between railway networks and cities by discovering that higher transportation costs reduce cities' size in sub-Saharan Africa. Storeygard points out that formal manufacturing activity to a large extent is concentrated in the largest cities of sub-Saharan Africa, restricting most hinterland cities from exporting manufacturing goods due to high transport costs. He further argues that transport costs per se have a crucial effect on cities' income by comparing the income levels for periphery cities with lower transport costs to the closest port with cities farther away or poorer road connections when oil prices rise. The results suggest that cities further away from a port see their transport costs increase more than cities closer, indicating a fall in income of the city.

Other researchers who have approached the emergence of cities through empirics in history find evidence of the railway's vital role in the economy. Jedwab, Kerby, and Moradi (2017) argue that Kenyan cities emerged due to the colonial Uganda railway in the late 19th century. They suggest that the new cities formed as a result of the railway and served as engines of the economy through various functions: trading stations, which expanded the private industrial and service sectors, and administrative seats, which increased the demand for labour. Furthermore, they suggest that the early emergence of railway and settler cities in Kenya served as a mechanism to coordinate spatial investments in subsequent periods. Their results imply that the colonial sunk investments, such as the railway, and the ability to solve spatial coordination failures during the colonial era, accounts partially for path dependence in Kenya. In other words, they find that Kenya's early railway investment still accounts for the economic outcomes of modern times.

The studies presented thus far provide evidence that railways have had historically and present influence on low-income and developing countries' economy. However, most relevant previous literature has approached the long-term effects of railways through macroeconomic analysis, thus neglecting essential heterogeneity. Instead, I aim to assess the micro-level since this approach allows for a more thorough evaluation of the effects. Furthermore, I use an inductive approach in which I aim to apply the empirics to the theory. Understanding how efforts to acquire sustained growth levels in low-income countries requires comprehensive assessment and grasping these efforts' impact on all societal levels to tackle the broader challenge of alleviating poverty.

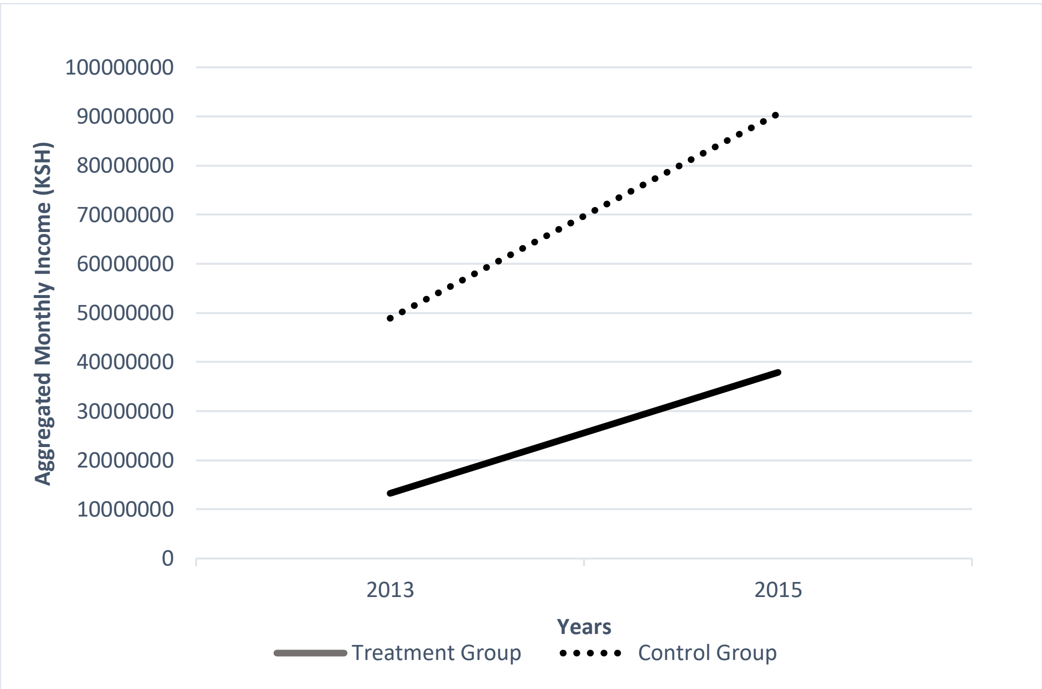
4. The Empirical Strategy

4.1. Differences-in-Differences

This chapter aims to explain the Differences-in-Differences model (DD) and evaluate its applicability for researching the economic impact, in terms of income, of the SGR in Kenya. Moreover, the chapter outlines the appropriate extensions of the model to fit the context of the research, as well as the model’s limitations. It also provides a detailed examination of the microdata used in the analysis and its validity for assessing the research question.

The DD estimation is suitable for the paper's applied research question "How has the Standard Gauge Railway affected the economic activity along its tracks in Kenya?", since it in line with Dougherty (2016) identifies differences in pre-and post-treatment outcomes. He explains that the DD strategy builds on a structure in which changes in the outcome variable, in this case, income, can be explained by an intervention only affecting a part of the population, i.e., the treatment group, as opposed to the unaffected part of the population, i.e., the control group, where the trend of the outcome remains unaffected. The model's simplicity is emphasised and its potential to circumvent many of the endogeneity problems that may arise when making comparisons between heterogeneous individuals. Similarly, he depicts the DD estimator's underlying assumption that both the treated and controlled group's average outcomes would have parallel trends in the absence of the treatment. This assumption is demonstrated in figure 1 but only evidently fulfilled before the second treatment due to non-existent data before the first treatment.

Figure 1: Parallel Trends Assumption



The research paper’s attempt to identify the railway’s impact on economic growth entails looking at two different treatments: (1) the construction of the railway and (2) the railway being in operation. The treatment and control group's design is constructed as follows: The treatment group consists of the seven counties that have a station along the SGR, i.e., Mombasa, Kilifi, Kwale, Taita-Taveta, Makueni, Machakos, and Nairobi. The control group consists of the counties that do not have a station, i.e., the remaining 40 counties of Kenya when evaluating treatment (2) of having a station, and an amount of 37 counties when evaluating treatment (1) of building a station (see further discussion on this in section 4.2 Data).

Given that both treatments are binary, Stock and Watson (2020) denote the treatments' effects as $E(Y|X = 1) – E(Y|X = 0)$. Moreover, they denote the DD estimator as following:

$$\hat{\beta}_1^{differences-in-differences} = (\bar{Y}^{treatment,post} - \bar{Y}^{treatment,pre}) - (\bar{Y}^{control,post} - \bar{Y}^{control,pre})$$
$$= \Delta\bar{Y}^{treatment} - \Delta\bar{Y}^{control}$$

where $\Delta\bar{Y}^{treatment}$ is the average change of income between pre- and post-construction of the railway (period 0 or 1 respectively) for the treatment group and $\Delta\bar{Y}^{control}$ is the average change in income for the control group (period 0 or 1 respectively). They emphasize the condition that given that the treatment is randomly assigned, the $\hat{\beta}_1^{differences-in-differences}$ is an unbiased and consistent estimator of the causal effect of the railway construction in Kenya. This allows for evaluating the research question of how the SGR has impacted economic activity in Kenya. As a reminder, the hypotheses are: H_1: the income level has risen in counties that have a station, as a result of constructing the railway, and H_2: the income level has risen in counties that have a station, as a result of the railway being in operation.

Further, the key to measuring how the SGR has impacted the economic activity can be assessed using OLS multiple regression analysis as pointed out by Stock and Watson (2020). In turn, the DD estimator is the OLS estimator of β_1 in the regression:

$$\Delta Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1)$$

They argue that by focusing on the change in Y over a set period of time, the DD estimator eliminates the initial influences of Y's values that may vary between the treatment and control group. Moreover, Stock and Watson (2020) expresses the regressions, when two time periods and two treatments are looked at, as follows:

$$Y_{i,t} = \beta_o + \beta_1 * D_i^{construction} + \beta_2 * D_t^{post} + \beta_3 * D_i^{construction} \times D_t^{post} + \varepsilon_{it} \quad (2)$$

$$Y_{i,t} = \beta_o + \beta_1 * D_i^{railway} + \beta_2 * D_t^{post} + \beta_3 * D_i^{railway} \times D_t^{post} + \varepsilon_{it} \quad (3)$$

where β_0 reflects the level of income in absence of the railway. Further, $D^{construction}$ and $D^{railway}$ illustrates a dichotomy reflecting whether the county is in the treatment group (or in the surrogate treatment group if the observation is in the pre-treatment period) (i = 1) or in the control group (i = 0). It estimates the baseline differences between the groups before the intervention was applied to the control group and takes care of the inherent differences between the treatment and control groups. D^{post} is the binary indicator equalling 0 in the pre-treatment period (t = 1), and 1 in the post-treatment period (t = 2). Its estimations reflect the pure effect of the passage of time in the absence of the actual intervention. Lastly, Stock and Watson (2020) point out the coefficient of interest in the analysis as the interaction term between time and period depicted as $D^{construction} \times D^{post}$ and $D^{railway} \times D^{post}$ in the equations. The interaction term estimates the actual treatment effect of the i^{th} county in the cross section in period t (t = 1, 2), and tells us whether the expected mean change in income from before to after is different in the two groups. It takes on the value 1 for the counties exposed to the treatment after implementing the railway, otherwise 0 (Stock & Watson, 2020).

4.1.1. Extensions of the Model

Altering the DD model to fit the paper's empirical research question and the accompanying context requires certain extensions of the model. Grasping Kenya's disparities across the country in terms of demography, it is crucial to acknowledge cross-country heterogeneity. Even though the treatment and control groups consist of Kenyan counties, their populations are likely to vary in several characteristics. Therefore, the model is expanded further to include additional regressors to control for differing characteristics before the intervention of the railway and isolate the effect of the railway (Stock & Watson, 2020). The additional regressors take the form of control variables and are chosen carefully not to possess a risk of being outcomes of the actual treatment, i.e., bad controls (Angrist & Pischke, 2009). The variables I control for are gender, cluster type (urban or rural), and age, and they are denoted as X'_{it} in regression (4). By including controls X'_{it} in the model, I aim to increase comparability between treatment and control groups over time and reduce standard errors (Stock & Watson, 2020).

Moreover, fixed effects for county and year are included in the equation. As concluded by Angrist and Pischke (2009), including county fixed effects prevents omitted variable bias arising from omitted factors as, e.g., cultural differences in transportation habits that vary across counties but are held constant over time within a county. Furthermore, the differences between treatment and control groups are supposed to be captured by the county-fixed effects, which in practice plays the same role as unobserved individual effects. When performing the actual estimations, the variables $D^{construction}$ and $D^{railway}(D_i^{treatment}$ in regression (4)). respectively are dropped, since these variables serve no purpose as county fixed effects already account for time invariant specifics over time. Moreover, the year effects control for the variables that are constant across counties but evolve over time (Stock & Watson, 2020). These effects are captured by the variable D^{post} and can thereby be interpreted as a surrogate for time fixed effects. The combined inclusion of county fixed effects and year fixed effects in regression (4) mitigate the threat of omitted variables arising from unobserved variables that either do not change over time or do not vary across counties (Stock & Watson, 2020). Finally, the model is tweaked enough to suit the research question and represented in regression (4).

$$Y_{i,t} = \beta_o + \beta_1 * D_i^{treatment} + \beta_2 * D_t^{post} + \beta_3 * D_i^{treatment} \times D_t^{post} + X'_{it} + \varepsilon_{it} \qquad (4)$$

4.1.2. Limitations of the Model

The DD estimation is often chosen when estimating the causal effects of policy implementation or event. Despite its multifaceted usability, the technique nevertheless contains drawbacks. The DD estimate may understate the standard deviation of the estimated effects severely because of serial correlation, resulting in over-estimation of significance levels. Bertrand, Duflo and Mullainathan (2004) also highlight the high risk of concluding faulty inferences when using the DD estimation over several time periods. The standard errors are clustered at the county level to address this issue since this allows for correlation within a county over time, as opposed to the usual heteroskedasticity-robust standard errors, and thus corrects for heteroscedasticity and serial correlation in the error terms (Stock & Watson, 2020).

The assumption of parallel trends of the treatment and control group prior to the intervention can be found implausible if the pre-treatment characteristics differ and are unbalanced between the two separate groups, as emphasised by Alberto Abadie (2005). Thus, he argues for the importance of that both the treatment and control group follow similar lines of characteristics.

To address this issue, as mentioned earlier, the model is expanded further to include additional regressors to control for differing characteristics before the railway intervention.

4.2. Data

To estimate the effects of the SGR on economic activity, I use microdata. The microdata originates from the FinAccess Household Survey which is a series of annual surveys jointly conducted by the Central Bank of Kenya (CBK), the Kenya National Bureau of Statistics (KNBS) and the Financial Sector Deepening Kenya (FSD Kenya, 2021). The FinAccess household survey is a series of surveys conducted every two to three years to determine the level of financial inclusion and measure the key drivers and catalysts of financial services in Kenya (FSD Kenya, 2021). The surveys chosen to evaluate the paper's research question are from 2013, 2015, and 2018 to examine two different effects: (1) the construction of the railway and (2) the railway being in operation. The surveys can be found in the Kenyan National Bureau of Statistics' National Data Archive (KeNADA) and the Harvard Dataverse.

All three surveys are based on KNBS's national sampling frame - National Sample Survey and Evaluation Program (NASSEP V) and the observations represent a randomly selected respondent per household (minimum age of 16 years old), according to The Financial Sector Deepening Kenya (2013; 2016; 2018). Although, when the 2013 survey commenced, the NASSEP V was incomplete, excluding three counties in North Eastern Kenya (Garissa, Mandera, and Wajir) due to lingering security issues. Thus, when examining the first treatment effect (the construction of the railway), these three counties' observations are excluded from the 2015 dataset describing the post-effects. The same is done for the dataset from 2018 when estimating both treatments' general effect. In other words, the sampling frame of the first treatment covers 44 out of Kenya's 47 counties, while the sampling frame of the second treatment covers all 47 counties.

The used data set comprises a small subset of variables from the FinAccess household surveys' data set. It was composed to fit a differences-in-differences estimation and thus constructed into a repeated cross-sectional data set (panel data). The premise of using cross-sectional data is that if the individuals are randomly drawn from the population, the individuals from period 1 (pre-treatment) can be used as surrogates for the individuals of period 2 (post-treatment), including both treatment and control group (Stock & Watson, 2020). Moreover, the microdata has been aggregated to the county level for the estimation. The aggregation addresses the potential problem that the grouped error terms as the observation unit is more detailed than the variation level. The problem can also be addressed by clustering the standard errors discussed in detail in section 4.1.1 about the model's extensions (Bertrand, Duflo & Mullainathan, 2004).

Descriptive statistics of the variable of interest, monthly income, can be found in Table 1 and illustrate individuals' mean income in the treatment and control groups over the observed time periods. The table also provides a framework for average income wages in Kenya when interpreting the results on income change in chapter 5. Moreover, the DD estimator measures the change in income over the course of the implementation, meaning that the estimator removes the initial influences that might vary between the treatment and control groups prior to the intervention (Stock & Watson, 2020). Thus, the absolute values of the income mean depicted in column (2) do not bias the estimator as long as both groups' values follow parallel trends.

Table 1: Descriptive Statistics on Income
Variable: monthly individual income (KSh)

Income	(1) Observations		(2) Mean		(3) Std. Dev.		(4) Min		(5) Max	
	Treatment	Control	Treat.	Cont.	Treat.	Cont.	Treat.	Cont.	Treat.	Cont.
Year 2013	1,248	4,932	10629.25	7681.203	24173.05	16601.68	50	30	450000	400000
Year 2015	1,839	6,826	26591.08	13263.49	355816.7	88151.57	0	0	1.52e+07	7000000
Year 2018	1,676	6,377	12115.39	9020.749	22446.51	17285.51	0	0	300000	400000

Information provided in the FinAccess Household surveys' technical reports explains how the data collection has been conducted without threatening internal validity. The internal threat of failure to follow the treatment protocol arises when members of the treatment group fail to receive the treatment and the control groups actually receive the treatment (Stock & Watson, 2020). This is not an issue of concern in this analysis since the railway is immobile and thus cannot accidentally impose the treatment to a county included in the control group. However, as mentioned by Taylor (2020), the SGR is built in two phases, which routes through different counties in Kenya, implying a possibility of failure to follow the treatment protocol since this paper only examines the effect of Phase 1 (Nairobi-Mombasa) and not Phase 2 (Nairobi-Malaba). This is addressed by looking at data post the opening of Phase 1 but pre-opening of Phase 2, thus implicitly assuming that the construction effects are similar for the treatment and control groups.

5. Results

To estimate the differences in income induced by the SGR in Kenya, the differences-in-differences method (DD) was used. The estimated DD regressors obtained from the OLS regression are presented as the general change in income post construction of the railway and the railway being in operation between the treatment and control groups. Construction of the railway began in 2014 and was finished roughly three years later in 2017 (Taylor, 2020). Thus, one could assume two different effects: firstly, the effect of constructing the railway, and secondly, the effect of the railway being in operation. The first section in this chapter motivates the extensions of the model followed by the second section that presents the results of treatment (1): the construction of the railway, while the third section moves on to discuss the results of treatment (2): the railway being in operation. Additionally, the total effect of both treatments is presented. Further, I compare the SGR's effects on gender, cluster type, and age. This analysis allows for identifying critical heterogeneity in the income data between the sub-groups. Lastly, I attempt to explain the relationship between income and the railway by disclosing potential channels of the relationship and linkages to the results of the heterogeneity analysis. Through these results, the Standard Gauge Railway's impact on income, and more generally, on economic activity can be concluded.

5.1. Motivation of the Extended Model

As previously mentioned in chapter 4, Bertrand, Duflo and Mullainathan (2004), highlight the DD estimation's prevalence of understating standard errors due to serial correlation. The OLS estimation used in the analysis assumes uncorrelation in the error terms over time, giving wrongly estimated results when looking at income levels between counties. By clustering the standard errors on the county level, I allow for within county correlation over time, and by that, correct for heteroscedasticity and serial correlation in the error terms. The effects of not clustering the standard errors are illustrated in the regressor showing the treatment effect; $D^{construction} \times D^{post}$ in Table 2, column (1), and in regressor $D^{railway} \times D^{post}$ Table 3, column (1). The only significant result (Table 3, column (1)) insinuates that the existence of a station in the county would have induced a 10233.95 KSh (€76) lower income compared to the counties that do not have a station. This result raises concerns and calls for further adjustment of the model.

In column (2) of both Table 2 and 3, demographic control variables are added to the regression to account for potential differences prior to the treatment between treatment and control group and isolate the railway effect. As expected, the additional regressors increase the R-squared value. However, similar levels of significance are found again, calling for further adjustment of the regression. To precise the estimate's correctness, fixed effects for county and year are evaluated and included in the regression as shown in column (3) in both Table 2 and 3. Since the variable for county fixed effects captures time-invariant specifics over time, the variables $D^{construction}$ and $D^{railway}$, that indicate whether the county will have/build a station or has/has built a station and thus are invariant over time, are dropped.

Furthermore, year fixed effects are captured by the variable D^{post} and can thereby be interpreted as a surrogate for time fixed effects. Lastly, the standard errors are clustered at the county level since this, as mentioned by Stock and Watson (2020), allows for correlation within a county over time, thus correcting for heteroscedasticity and serial correlation in the error terms. Illustrated in Table 2, the standard error of the interaction term decreases from 6016.691 in column (2) to 5.22e-09 in column (3), and from 4987.344 in column (2) to 3.46e-08 in column (3) in Table 3, when clustering the standard errors at the county level. Finally, both county and year fixed effects, clustered standard errors, and demographic control variables are taken into account in the regression and depicted in column (4) in both tables.

5.2. The Construction of the Railway

This section presents the construction effects of the railway and attempts to explain the results by relating to the context and the previous literature. The regressor $D^{construction} \times D^{post}$ in column (4), Table 2, illustrates the treatment effect of constructing the railway. It shows that the people living in counties that constructed a railway station had a significantly higher income of 4298.655 KSh (€32), which corresponds to 16% of the treatment group's average income in 2015 (depicted in Table 1). This result is not surprising given that the railway's construction and maintenance were expected to generate Kenyans' employment opportunity. It also substantiates the idea of Jedwab, Kerby, and Moradi (2017), that the construction of the railway solely generated economic activity. They argue that the Uganda railway construction gave rise to new cities, which served as engines of the economy. Even though these results yield no information regarding the rise of cities in Kenya due to the study's short time frame, it implies that the construction of a railway, now and 100 years ago, has positively impacted the economies along the routes. Moreover, this result implies that the recruiting took place locally relative to the railway line. Stretching the results' interpretation further resonates that the construction works also induced indirect effects of the increased employment (of presumable men), through increased demand generating income for e.g., street vendors and women working at the markets.

Table 2: The Construction of the Railway

Dependent variable: monthly income (KSh) years 2013 and 2015

Regressors	(1)	(2)	(3)	(4)
$D^{construction}$	2984.043	-594.7251		
	[4499.544]	[4748.407]		
D^{post}	5773.393*	4709.734	1704.184***	1469.079***
	[2704.303]	[2789.525]	(1.37e-08)	(200.8462)
$D^{construction} \times D^{post}$	10188.44	10903.69	4545.031***	4298.655***
	[5868.181]	[6016.691]	(5.22e-09)	(475.1914)
Gender		-8230.555**		-8469.803**
		[2486.43]		(3190.915)
Cluster type		-10607.78***		-11051.77**
		[2607.021]		(4214.089)
Age		153.9088*		156.2242
		[76.3524]		(82.80334)
Observations	14,272	13,944	14,272	13,944
R-squared	0.0017	0.0039	0.0055	0.0077
County Fixed effects	NO	NO	YES	YES
Year fixed effects	NO	NO	YES	YES
Clustered standard errors	NO	NO	YES	YES

Standard errors in brackets and clustered standard errors in parenthesis, * p < 0.05, ** p < 0.01, *** p <0.00

5.3. The Railway being in Operation

This section evaluates the economic effects of the railway actually being in operation and hence facilitating trade and communication to markets in Kenya. Column (4) in Table 3 depicts this effect and shows that the railway induced a significantly higher monthly income of 2434.381 KSh (€18), correspondent to 20% of the average income of the treatment group in 2018 (depicted in Table 1), suggesting a positive relationship between income and the railway. To bear in mind, the effect is measured only a year after the railway's opening, meaning that the long-term effects have not yet been unveiled. Furthermore, when interpreting the results from the railway treatment (i.e., between 2015 and 2018), one should consider that year 2017 was fraught with several factors that disrupted Kenya's income development as depicted by the negative value of coefficient D^{post} in Table 3. The coefficient implies that Kenyans, no matter geographical location, generally earned a significantly lower income of 7813.079 KSh (€58) during this time period. This general negative trend in income in Kenya can be attributable to three main drivers; firstly, widespread drought affected agricultural output negatively; secondly, economic activity was constricted by a continued decline to credit in the private sector; thirdly, the private sector activity was weakened due to uncertainty originating from Kenya’s general election in 2017 (World Bank, 2017). However, despite the dropping levels of income in Kenya, the positive coefficient of $\boldsymbol{D^{railway}} \times \boldsymbol{D^{post}}$ uncovers positive regional effects on income along the railway induced by the railway's influence.

Table 3: The Railway being in Operation

Dependent variable: monthly income (KSh) years 2015 and 2018

Regressors	(1)	(2)	(3)	(4)
$\boldsymbol{D^{railway}}$	13327.59***	10727.35**		
	[3451.689]	[3507.698]		
$\boldsymbol{D^{post}}$	-4242.741	-4413.316	-8251.892***	-7813.079***
	[2288.054]	[2290.665]	(6.80e-08)	(182.9908)
$\boldsymbol{D^{railway}} \times \boldsymbol{D^{post}}$	-10232.95*	-10312.94*	3750.736***	2434.381***
	[4991.908]	[4987.344]	(3.46e-08)	(516.7637)
Gender		-7684.798***		-7868.838**
		[2070.209]		(2661.631)
Cluster type		-9767.899***		-9979.245**
		[2148.025]		(3363.367)
Age		155.8707		160.4808*
		[61.3055]		(66.85174)
Observations	16,718	16,718	16,718	16,718
R-squared	0.0015	0.0039	0.0055	0.0078
County Fixed effects	NO	NO	YES	YES
Year fixed effects	NO	NO	YES	YES
Clustered standard errors	NO	NO	YES	YES

Standard errors in brackets and clustered standard errors in parenthesis, * p < 0.05, ** p < 0.01, *** p <0.001

5.4. The General Effect of the Railway

The railway's overall effect is highlighted in Table 4. It shows that the combined effect of both the railway construction and the operating railway induced a significantly higher monthly income of 3355.406 KSh (€25), which corresponds to almost 28% of the treatment group's average income in 2018 (depicted in Table 1). Moreover, a negative trend in income is depicted in coefficient D^{post} in column (1), reflecting the pure income reduction of 961.1387 KSh (€7) for both groups in the absence of both the construction and the operating phase of the railway. This direction of abating incomes is reasonably related to Kenya's general negative trend in income in 2017, as mentioned above. Considering this, the total effect ($D^{railway} \times D^{post}$) is not expected to sum up to a total of both treatment's coefficients.

Table 4: Total Effect of the Railway

Dependent variable: monthly income (KSh) years 2013 and 2018

Regressors	**(1)**
$\boldsymbol{D^{post}}$	-961.1387***
	(41.00284)
$\boldsymbol{D^{railway} \times D^{post}}$	3355.406***
	(69.35625)
Gender	-4110.227***
	(425.3291)
Cluster type	-4860.853***
	(542.9923)
Age	-1.625799
	(10.94211)
Observations	13,483
R-squared	0.0731
County Fixed effects	YES
Year fixed effects	YES
Clustered standard errors	YES

Clustered standard errors in parenthesis, * p < 0.05, ** p < 0.01, *** p <0.001

5.5. Heterogeneity Analysis

Disintegrating the analysis further to explore heterogeneity in the data, additional significant results are found in Table 5. The sub-groups of the population are depicted in the various columns of Table 5. I find that both the constructions of the railway, and the railway being in operation induced a positive income gap between counties that constructed/had a station and those that did/had not, for both women (1) and men (2). Furthermore, a significant result in column (2) shows a substantial difference in income of 5855.046 KSh (€44) when comparing the effect induced by the construction on men living in counties with stations on the railway with the control group, suggesting that mostly the men living nearby the stations were hired to construct the railway. Although men generally earn a higher income than women, the heavy physical work required in the construction sector could also explain the unequal distribution of income illustrated in the table. However, the corresponding results for women's income (depicted in column (1)) suggest a multiplier effect increasing the economic activity for women residing along the railway as well.

When comparing urban (3) and rural (4) areas, a significantly higher income for rural areas is found for both treatments. This higher level of income for rural areas is possibly driven by the fact that the marginal effects of new transport infrastructure are diminishing, meaning that the effect's substantiality of having an additional transport should depend on whether the area already has a wide range of transports or not. It is also plausible that the railway has facilitated a decentralisation effect, as seen in China by Baum-Snow et al. (2017), where the industrial production is moved out of the central cities and thus potentially benefits people living in rural areas along the railway through increased economic activity. On the contrary, the lower income induced by the railway in urban areas also confirms the results of Baum-Snow et al. (2017), that railways caused a decline in manufacturing employment for urban residents. Another potential inference suggests that as the railway constructions were finished, urban people living in counties that constructed a station saw a decrease in economic activity possibly due to loss of employment that accompanied the completed construction.

Finally, an arbitrary boundary is drawn at 45 years old when comparing age depicted in columns (5) and (6). It shows that a significantly higher income induced by the railway can be found for people aged 45 and younger. These results imply that the younger population was hired to construct the railway since it logically entailed heavy physical labour. Even after the construction was completed, people aged 45 and younger have a higher income due to the railway, possibly explained by the fact that younger people are more mobile and use transportation to a relatively larger extent to spur their economic activity. Since the results concerning people older than 45 are insignificant, there is nothing to be concluded rather than a skewed distribution of age in the sample. Summing-up, when exploring heterogeneity in the data, significant results disclose that young men residing in rural areas are identified as main economic beneficiaries of constructing the railway and the railway being in operation.

Table 5: Heterogeneity Analysis

Dependent variable: monthly income (KSh), sub-groups of the total population

Interaction term	(1) Women	(2) Men	(3) Urban	(4) Rural	(5) Age (>45)	(6) Age (≤45)
$D^{construction} \times D^{post}$	2914.168***	5855.046***	2197.124	4254.943***	164.6577	4467.794***
	(753.084)	(999.603)	(2287.751)	(72.6996)	(3385.037)	(7.801577)
$D^{railway} \times D^{post}$	2139.854***	2603.924*	-6152.609***	5957.534***	-185.7461	3717.198***
	(525.9625)	(1226.077)	(878.6867)	(49.83735)	(1442.502)	(48.5631)
Observations construction 2013 & 2015	8,375	5,569	5,671	8,273	3,704	10,568
Observations railway 2015 & 2018	9,942	6,776	7,062	9,656	4,604	12,114
R- squared construction	0.0111	0.0162	0.0139	0.0553	0.0166	0.0522
R-squared railway	0.0115	0.0163	0.0143	0.0483	0.0166	0.0524
County Fixed effects			YES			
Year fixed effects			YES			
Clustered standard errors			YES			

Clustered standard errors at county level in parenthesis, * $p < 0.05$, ** $p < 0.01$, *** $p <0.001$

5.6. Channels of the Results

Turning to the final part of the analysis, I look at the potential channels and drivers of the positive relationship between income and the railway being in operation (the second treatment 2015-2018). The regression follows the same format (including demographic control variables, county and time fixed effects, and clustered standard errors) as for previous models, but instead of the income as the dependent variable, potential sources and drivers of the positive relationship between the railway treatment and income is evaluated. The dependent variables are transport expenses, savings and investment expenses, and income sources (different livelihood sectors). The data for this analysis can be found in the columns of Table 6.

Column (1) depicts the monthly transport expenses of individuals. The positive coefficient of 511.4753 KSh (€3.8) indicates that people living along the railway have higher transport expenses than the rest of Kenya's population. The positive outcome could indicate increased mobility due to the new railway connecting people to the markets to a larger extent. Transport infrastructure should logically catalyse increased moveability and thus, higher transport expenses. However, no straightforward linkages can be drawn between income and transport expenses since the transport expenses are not measured relatively income.

Moreover, a significantly positive relationship between monthly savings/investment rate of a county's individuals, and the railway can be established. The coefficient of column (2) illustrates that the railway induced a higher savings/investment rate of 3669.335 KSh (€27). This result can be compared with the main result of Table 3, which shows that the railway generated higher income. Subsequently, a higher income would logically entail a higher savings rate and greater opportunity for investing your money. However, these types of conclusions cannot be drawn with complete confidence this shortly after the shock of constructing a new railway but calls for further research in people's financial behaviour linked to better commuting.

Column (3) depicts the income source that generated the most money, where alternative (a), (b), (c), and (d) are binary dependent variables indicating whether the individual gained the most income from farming (crops or keeping livestock), casual work, employment (working for a wage), or self-employment/running own business. Thus, the coefficients should be interpreted as weights of the probability of gaining the most income from the mentioned source. The railway induced a shift in the weighting of how people earn money where the probability of earning money from casual work and self-employment increased by 11.5% and 6.7%, respectively, while the probability of earning money from farming and employment decreased by 28.5% and 2.5%, respectively. Linking these results to the heterogeneity analysis, potential conclusions can be drawn.

Firstly, the reduced possibility of gaining income from wage employment depicted in column (c) could explain the negative coefficient of income for people living in urban areas (Table 4, column (3)), assuming that formal employment rates are generally higher in urban areas compared to rural areas. A generally reduced rate of wage employment logically derives a lower income. It is also in line with the previously articulated thought that people living in urban areas saw a decrease in economic activity due to the loss of employment that accompanied completed construction and that the expected positive effects of the railway being in operation have not uncovered in these areas yet. This result could also be compared to the result of Baum-Snow et al. (2017) who found that the railway induced declining rates of manufacturing employment in central cities. However, without further data on specific sectors, this inference cannot be taken with complete confidence.

Secondly, the reduced probability of earning most income from farming (column (a)) and the higher income of people living in rural areas in counties with a station along the railway come off as contradicting, at first sight, assuming relatively higher farming rates in rural areas. Reasonably, the probability that people engage in farming would increase due to the greater access to markets the railway induced. The declining probability rate refutes this conclusion but encourages further investigation when more extensive data on employment sectors is available. The results also depart from the framework of Banerjee, Duflo, and Qian (2020) who assumed small differences in GDP levels between remote and better-connected areas due to a contending direction of effects induced by improved transportation infrastructure. Thus, assuming that rural areas retain more capital per head than urban areas, the railway boosts labour productivity in those areas, hence possibly explaining increased income. Nevertheless, it could also indicate decentralisation of industrial production out of the cities to rural areas, in consonance with Baum-Snow et al. (2017), implying a sectoral shift in rural areas to more formal ways of earning money.

Finally, the higher probability of earning money from being self-employed and performing casual work, due to the railway as depicted in column (d) and (b) respectively confirms the increased levels of economic activity along the tracks. The higher self-employment rate suggests an entrepreneurial spirit and a more comprehensive range of possibilities to earn an income, while the increased activity from casual work could illustrate the short-term effects of increased economic activity induced by the railway.

Table 6: Channels of the Results

Dependent variables depicted in the columns

	(1) Transport expenses	(2) Savings/ Investment	(3) Most Important Income Source			
			(a) Farming	(b) Casual Worker	(c) Employed	(d) Self-Employed
$D^{railway} \times D^{post}$	511.4753*** (7.313293)	3669.335*** (26.44576)	-0.2854702*** (0.0024381)	0.1153519*** (0.001523)	-0.0251281*** (0.0010856)	0.0666782*** (0.0013771)
R-squared	0.0563	0.0376	0.2072	0.0474	0.0640	0.0521
Observations	17,039	16,938	17,334	17,334	17,334	17,334
County Fixed effects	YES					
Year fixed effects	YES					
Clustered standard errors	YES					

Clustered standard errors at county level in parenthesis, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6. Conclusion

The main objective of the thesis was to evaluate the Standard Gauge Railway's effects on the economic activity along the railway tracks. The paper's research design targeted how the shock of constructing a new railway and the railway being in operation affected the regional income of individuals residing in counties with a station on the SGR line. The results make two key contributions that should be emphasised. Firstly, the counties' income that constructed a station was significantly higher, potentially related to the employment opportunities arising from the construction. Secondly, a significantly higher income was disclosed in the counties where a station was located, plausibly due to the increased accessibility and transportation opportunity stemming from the operating railway. These findings have significant implications for the understanding of the role that adequate transportation infrastructure plays in emerging economies.

Furthermore, the analysis exploring heterogeneity in the data suggests critical differences within the population where the most noteworthy results unveil an increased income in rural areas, but a decrease in income in urban areas due to the railway. Aiming to explain the latter result, the analysis uncovering the results' drivers suggests a linkage between the reduced income in urban areas and the reduced possibility of earning income through formal employment. Putting these results in an empirical context, the completion of the railway and the plausible accompanying decline in economic activity could also serve as an explanation. Moreover, the crucial question of the railway's differing influences on gender acknowledges corresponding positive results for both women and men, even though a larger increase in income for men can be depicted. The findings shed new light on the importance of exploring heterogeneity when determining the causal effects and extend our knowledge of railways' contrasting economic implications.

In attempts to explain the potential channels of the concluded positive linkage between income and the railway, a potential sectoral shift in how people earn incomes uncovered. The reduced probability of earning money from farming and the higher income of people living in rural areas located alongside the railway contrasts at first sight but could also hint of an industrialized change where people now earn money through more formal activities. Moreover, the higher probability of engaging in casual work or self-employment advocates for generally increased levels of economic activity attributed to the railway. These results endorse the identified linkage between an increased income generated by the railway and prove useful in expanding our understanding of how railways play a critical part in people's livelihoods.

Several questions remain to be answered. There is abundant room for further progress in determining the Standard Gauge Railway's long-term effects in Kenya. Neither do the results yield knowledge of the SGR's role as transportation of cargo or its economic implications on different employment sectors. These questions are left for future research. However, in spite of its limitations, the study certainly demonstrates the vital role of adequate transportation in the economy of low-income countries. And more importantly, the Standard Gauge Railway does not prove to be particularly lunatic after all.

7. References

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators, *Review of Economic Studies*, vol. 72, no. 1, pp.1-19

Amin, S. (1974). Accumulation and Development: A Theoretical Model, *Review of African Political Economy*, vol. 1, no. 1, pp.9-26

Angrist, J. D., & Pischke, J. S. (2009). Mostly Harmless Econometrics: An empiricist's companion, Princeton: Princeton University Press.

Banerjee, A., Duflo, E., & Qian, N. (2020). On the Road: Access to Transportation Infrastructure and Economic Growth in China, *Journal of Development Economics*, vol. 145

Baum-Snow, N., Brandt, L., Henderson, V. J., Turner, M. A, & Zhang, Q. (2017). Roads, Railroads, and Decentralization of Chinese Cities, *The Review of Economics and Statistics*, vol. 99, no. 3, pp.345-448

Bertrand, M., Duflo, E. & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates?, *The Quarterly Journal of Economics*, vol. 119, no. 1, pp.249- 275

Cohan, W. (2017). What Is One Belt One Road? A Surplus Recycling Mechanism Approach [pdf], Available at http://dx.doi.org/10.2139/ssrn.2997650 [Accessed 10 January 2021]

Department for International Development. (2021). Growth Building Jobs and Prosperity in Developing Countries [pdf], Available at https://www.oecd.org/derec/unitedkingdom/40700982.pdf [Accessed 2 January 2021]

Dougherty, C. (2016). Introduction to econometrics, 5th edn, Oxford: Oxford University Press

Financial Sector Deepening Kenya. (2013). 2013 FinAccess Technical Note [pfd], Available at: https://dataverse.harvard.edu/ [Accessed 17 December 2020]

Financial Sector Deepening Kenya. (2016). 2016 FinAccess Technical Note [pfd], Available at: https://dataverse.harvard.edu/ [Accessed 17 December 2020]

Financial Sector Deepening Kenya. (2019). 2019 FinAccess Technical Note [pfd], Available at: https://statistics.knbs.or.ke/ [Accessed 17 December 2020]

Financial Sector Deepening Kenya. (2021). FinAccess, Available online: https://www.fsdkenya.org/ [Accessed 2 December 2020]

Gil-Alana, L., & Mudida, R. (2018). The growth rate series in Kenya: Evidence of non-linearities and factors behind the slow growth, *International Journal of Finance & Economics*, vol. 23, no. 2, pp.111-121

Hillbom, E., & Green, E. (2019). An Economic History of Development in sub-Saharan Africa, 1st edn, [e-book] Cham: Springer International Publishing, Available through: LUSEM University Library website http://www.lusem.lu.se/library [Accessed 28 November 2020]

Jedwab, R., Kerby, E., & Moradi, A. (2017). History, Path Dependence and Development: Evidence from Colonial Railroads, Settlers and Cities in Kenya, *The Economic Journal*, vol. 127, no. 603, pp.1467-1494

Kacungira, N. (2017). Will Kenya get money for value from its new railway, 7 June, Available Online: https://www.bbc.com/news/world-africa-40171095 [Accessed 10 January 2021]

Kenya Vision 2030. (2021). Vision 2030, Available Online: http://vision2030.go.ke/ [Accessed 27 November 2020]

Kenya's GDP Growth Slumps in 2017, but can Rebound over the Medium-Term. (2017). World Bank, Available online: https://www.worldbank.org/en/news/press-release/2017/12/06/kenyas-gdp-growth-slumps-in-2017-but-can-rebound-over-the-medium-term [Accessed 14 December 2020]

Olken, A. (2020). Banerjee**,** Duflo**,** Kremer**,** and the Rise of Modern Development Economics, *Scandinavian Journal of Economics*, vol.122, no.3, pp.853-878

Salaudeen, A. (2019). Kenya launches the second phase of its billion-dollar Chinese railway project, CNN, 16 October, Available Online: https://edition.cnn.com/2019/10/16/africa/kenya-launches-chinese-railway-line/index.html [Accessed 29 November 2020]

Shivji, I. (2009). Accumulation in an African periphery: A theoretical framework, Dar es Salaam: Mkuki na Nyota Publishers

Solow, R. M. (1956). A contribution to the theory of economic growth, *Quarterly Journal of Economics*, vol. 70, no. 1, pp.65-94

Spicer, P. (2010). The Temptress: The Scandalous Life of Alice, Countess de Janzé. London: Simon and Schuster

Stock, J. H., & Watson, M. W., (2020). Introduction to Econometrics, 4th edn, Harlow: Pearson Education Limited

Storeygard, A. (2016). Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa, *Review of Economic Studies,* vol. 83, no. 3 (296), pp.1263-1295

Taylor, I. (2020). Kenya's New Lunatic Express: The Standard Gauge Railway, *African Studies Quarterly*, vol. 19, no. 3-4, pp.29-52

Todaro, M. P., & Smith, S. C., (2011). Economic Development, 11th edn, Essex: Pearson

Wo-lap Lam, W. (2016) Getting Lost in 'One Belt, One Road', EJ Insight, 12 April, Available Online: https://www.ejinsight.com/eji/article/id/1281194/20160412-getting-lost-one-belt-one-road [Accessed 22 December]

World Bank. (2020). Annual Report 2020 Lending Data [pdf], Available at: worldbank.org [Accessed 2 November 2020]

World Bank. (2021). Transport. Available online: https://www.worldbank.org/en/topic/transport/overview [Accessed 28 December 2020]

World Bank. (2021). GDP growth (annual %) (World Development Indicators). Available online: https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=KE [Accessed 20 December 2020]