July, 2022
Wageningen University - Social Sciences

TAKING THE TRAIN TO WORK

FEMALE EMPLOYMENT AND EDUCATION EFFECTS OF RAILWAY DEVELOPMENT IN NIGERIA

WAGENINGEN
UNIVERSITY & RESEARCH

MSc Thesis Agricultural Economics and Rural Policy
Supervisor: dr. Esther Gehrke
MSc International Development Studies:
Economics of Development

BY JURRIAN NANNES - 1041806

Abstract

Transportation infrastructure enhancement is often associated with economic development. However, studies show that the presence of positive economic effects usually depends on the context of the infrastructure project and its economic demand. The nationwide Lagos-Kano railway project, which is constructed in Nigeria this decade and is expected to cost USD 11.1 billion, is an infrastructure project which policymakers assume to be "critical to development". The railway network is deemed to become an essential factor for effective labour force distribution in unemployment-struck Nigeria. In this thesis, I investigated short-term female education and employment effects of the firstly operating Lagos-Kano railway network segment running between Abuja and Kaduna. Data was acquired from the Demographic Health Survey program's 2018 Nigeria wave out of which four subgroups were created: rural and urban women younger than 30 years old, and rural and urban women aged 30 and older. Treatment subgroup women were selected based on their geographical distance to the nearest new train station. Control subgroups were selected and reweighted using entropy balancing. The thesis' main findings show decreased female labour participation and decreased agricultural employment for young rural women. In contrast, their husbands show increased employment, increased non-agricultural employment and decreased agricultural employment. Furthermore, married young urban women in treatment areas were more likely to be enrolled in higher education. These findings are in line with Goldin's (1995) U-shaped female labour force function theory. It is concluded that the railway development did generally not improve female employment and education, but did increase employment rates and non-agricultural labour of their husbands.

Acknowledgements

Spending practically thirteen months on an MSc thesis is only possible with the help and support of amazing friends, family and supervisors. First, I would like to thank my supervisor Esther Gehrke for her patience, help, being permanently available for confusing questions and her quick and direct responses to emails. Second, I would like to thank and wish luck to all Gaia-during-lockdown students, especially the ones who moved the fun to Italy with me. Third, everyone who joined ping-pong breaks with me: Luuk Staal, Jord Feijtel, Kimbery Morren, Jasper Kleinsmann, Loris Monasso, Microsoft Word-guru Karlijn Arbouw, and many, many others. Fourth, my family who, whatever struggle we needed to go through, always supported me in every way possible. Fifth, special thanks to my dad, for support during the thesis work but especially for everything else, who wasn't given the time to see the result but who will be close with me for everything to come.

Content

List of tables

1. Introduction

An old Chinese saying goes, "if you want to be rich, build a road". Without exaggeration, one could say that throughout history this mantra has been put into Chinese practice. The People's Republic of China (PRC) experienced strong economic growth between the 1970s and the 2000s and the immense scale of transportation infrastructure development projects across the country is believed to have been a crucial driver (Chen, 2018; Sahoo, Dash & Nataraj, 2010). After PRC emerged from an economy driven by agriculture to the industrialised modern economy it is today, the ruling Communist Party shifted its focus on infrastructure investment projects beyond PRC's borders. This resulted in what has in 2013 been announced as the 'One Belt, One Road'-initiative: an intercontinental infrastructure network that connects Beijing to Southeast Asia, Europe and Africa, intended to enforce international trade and cooperation with the countries lying in between. According to President Xi Jinping, PRC and the investment-receiving countries benefit not only from increased trade: he believes strengthening infrastructure induces peace and will therefore alleviate poverty in participating countries (Chen, 2018).

In the literature, transportation infrastructure enhancement is often positively associated with indicators of economic development. Studies show that infrastructure development causes economic growth (Pradhan & Bagchi, 2013), increases interregional trade and income levels (Donaldson, 2018), improves trade environments (Keller & Shiue, 2008; Michaels, 2008), improves labour allocation (Asher & Novosad, 2016), increases nonfarm employment (Lei, Desai & Vanneman, 2019), reduces income inequalities (Zou et al., 2008), and increases literacy (Chaudhary & Fenske, 2020) and education enrolment (Aggarwal, 2018). In contrast, other studies show poor effects on economic growth (Banerjee, Duflo & Qian) and on agricultural outcome, employment and income (Asher & Novosad, 2020). Economic effects of infrastructure development are shown to be context-dependent (Okoye, Pongou & Yokossi, 2019) and only present if combined with other socio-economic interventions (Mohmand, Wang & Saeed, 2017; Ng et al., 2019). Furthermore, theory states that infrastructure development is only economically beneficial when there is certain demand to connect different regions (Fishlow, 1965). In their report on the essence of infrastructure development, the World Bank (1994) points out to policymakers that transportation infrastructure could bring major benefits in economic growth and poverty alleviation, "*but only when it provides services that respond to effective demand and does so efficiently*" (World Bank, 1994, p. 2).

Nigeria is one of the countries that received Chinese investments as part of the Belt & Road Initiative (BRI) (as the One Belt, One Road initiative is called now). In Nigeria, the main project is a country-wide railway network from the coastal metropolis Lagos to Kano, the country's second-largest city in the far north. The entire project is estimated to cost USD 11.1 billion and USD 5.3 billion has so far been borrowed from PRC's EXIM Bank for constructing the first segments (The China Africa Project, 2021). In 2016, the first segment of the nation-wide railway network started its operations between capital Abuja and Kaduna. Construction was executed by China Civil Engineering Construction Corporation (CCECC) (Railway Gazette International, 2016). Despite the immense investment size, rising debts with foreign banks and the tightening dependency on Chinese construction companies and workers, Nigerian authorities claim the investment is worthy as the railway network is going to be "critical to development" for Nigeria, according to its minister of transportation (BBC News, 2017). With its rapid urbanisation and population growth, modernising and expanding Nigeria's infrastructure network seems like a logical decision for stimulating economic development (Aliyu & Amadu, 2017). Well-functioning infrastructure could mobilise the large group of (young) unemployed Nigerians in northern states. Chen (2018) claimed two years after the first operations started, that the Abuja-Kaduna segment seems to satisfy the demands of commuters and migrants "*seeking employment and economic opportunities*", according to anecdotal evidence (Chen, 2018, p. 4). This would be a welcome effect, given that Nigeria's unemployment rate in 2016 reached 14% (National Bureau of Statistics, 2020).

In this thesis, I investigate the short-term effects of the Abuja-Kaduna railway segment on female labour participation and education enrolment of women living nearby the new train stations. Most available datasets regard women, so focusing this thesis on women leaves me better possibilities of finding sufficient data. The thesis will contribute to the existing literature on socio-economic effects of infrastructure development in developing countries. It will provide insights into how modern infrastructure projects affect female labour and education enrolment in a modern-day Sub-Saharan African (SSA) economy. Existing literature on socio-economic effects of transportation infrastructure for individuals mainly concerns colonial railways (e.g., Donaldson, 2012; Chaudhary & Fenske, 2020; Okoye, Pongou & Yokossi, 2019) or paved roads (e.g., Adukia, 2017; Asher et al., 2016; Asher & Novosad, 2020; Banerjee, Duflo & Qian, 2020; Chaudhary & Fenske, 2020; Donaldson, 2012). Research on 21st-century infrastructure projects in SSA is rare, but, considering the rising interest in its investments and the uncertainty of economic effects of infrastructure, very relevant.

The main question that is handled in this thesis is whether women who live nearby a new railway segment are (1) more likely to be employed, (2) more likely to be employed in non-agricultural sectors, (3) less likely to be employed in agriculture, and (4) more likely to be enrolled in secondary or higher education. It is hypothesised that women living nearby new train stations benefit from the railway's increased mobility in terms of increased employment and increased education enrolment and show a shift from agricultural labour to non-agricultural labour. The train is assumed to be a safer and more efficient way of commuting to and from the cities of Abuja and Kaduna, where formal labour demand is expected to be higher and secondary or higher education opportunities more present. This could directly lead to increased access to more diverse labour markets (in non-agricultural sectors), and increased net returns on education when costs of education (time and transportation) are reduced and job prospects improved. Also, indirectly, the train stations may attract economic activity to the new stations, which may lead to increased labour demand for women living nearby.

890 women are selected from the Demographic Health Survey (DHS) Program's 2018 Nigeria wave who live 'nearby' the new train stations. These women are distributed over four treatment subgroups: rural women younger than 30 years, rural women aged 30 and older, urban women younger than 30 and urban women aged 30 and older. Entropy balancing is used as a matching method to define control groups and to obtain appropriate analytic weights for regressions. The main results show diverse effects for the different subgroups, but generally, it could be concluded that the railway development is not associated with increased female labour force participation (FLFP) and education enrolment. To provide a more complete picture of how the railway development changed the livelihoods of treated women, additional regressions are executed to investigate employment effects of husbands, fertility effects, marriage effects and immigration effects.

This thesis continues as follows. First, background information regarding the context of the infrastructure development and the investigated region is given. Second, a theoretical framework is provided along with the hypotheses for the four subgroups. Third, the empirical method is explained. This is followed by, fourth, results of the regressions and additional regressions, fifth, a discussion section and, sixth, a conclusion.

2. Background

2.1 The Belt & Road Initiative and its African debts

Since its economy opened for foreign trade throughout the 1970s and 1980s, PRC became a strongly growing economic force in world trade (Yang, 2006). After the 1997 Asian economic crisis, PRC adopted its "go global" policy and joined the World Trade Organisation (Davies, 2009; Kaplinsky & Morris, 2016; WorldBank, n.d.; Yang, 2006). PRC's Overseas Foreign Direct Investments (OFDIs) soared in the years

following, resulting in a doubling of the amount of capital outflow in 2008 compared to 2007, while global FDI fell by 20% during the economic crisis. Reasons for this unprecedented increase of OFDI from a 'developing country' could be found in (1) increasing wages causing domestic companies to replace their manufacturing labour to low-wage-countries, (2) state-owned companies (SOCs) losing monopolistic positions and looking to diversify abroad, (3) large domestic companies acquiring global brands and (4), most importantly, when PRC needed natural resources to fuel its economic growth, offering Official Development Assistance (ODA) to developing countries to increase the willingness of having Chinese firms investing in natural resource extraction projects (Davies, 2009; Tvaronavičienė, 2019).

Chinese OFDI reached a new high tide in 2013 when president Xi Jinping announced the 'One Belt, One Road' initiative; a now USD 1 trillion worth investment project intended to boost cooperation and trade between countries on the Eurasian continent and primarily focused on infrastructure development and mining (Frankopan, 2018). As the name of the initiative changed to 'Belt & Road Initiative', so did the scope. Today, BRI investments also occur in Latin America and SSA, amounting to 80 countries worldwide according to PRC's Belt and Road Portal (n.d.). Along with the announcement of the BRI arose a variety of vehicles through which the Chinese government sends capital overseas, such as the BRICS Development Bank, Asia Infrastructure Investment Bank, and the Silk Road Fund (Abdusalam et al., 2021). According to the OECD, PRC's outward FDI went from nearly USD 65 billion in 2012 to USD 216 billion in 2016, whereas it dropped to a somewhat constant yearly USD 130 billion between 2017 and 2020 (OECD, 2021).

BRI investment projects benefit from PRC's abundancy of manufacturing resources (steel, cement and machinery) and knowledge and capabilities that were no longer necessary within the national borders after the country's spectacular transition from a manufacturing-driven developing country to a service-driven world power between the 1990s and 2000s (Frankopan, 2018). At first glance, BRI investments intended to improve infrastructure and energy supply systems supported by PRC's abundant resources, knowledge and capabilities seem like a welcome development for the global south. However, the broader public opinion regarding these investments is, to put it cautiously, negative (e.g., Davidson, 2018; The Economist, 2018). Important to remember is that most of these investments are not grants but loans against competitive interest rates, and that most BRI investment-receiving countries are vulnerable states with relatively corrupt governments (Balding, 2018; Were, 2018; Hillman, 2019). Many argue that PRC forces developing countries into debts that are fiscally unsustainable, resulting in forced takeovers of equity or land from these countries. This happened in Sri Lanka in 2017, when Chinese SOCs gained ownership over the Hambantota Port after the Sri Lankan government appeared to be unable to repay its over one billion US dollar debt (e.g., Fernholz, 2018 & Frankopan, 2018). Similar incidents occurred with the port of Djibouti and infrastructure projects in Pakistan (Brautigam, 2020)[1].

Disregarding the nature of PRC's intentions, fact is that SSA's Chinese debt-to-GDP ratio is at an all-time high and expected to be growing in the coming years as receiving governments struggle to repay (IMF, 2021). In the 1980s and 1990s, SSA countries were heavily indebted which led to the necessity of structural adjustment programmes that obliged creditors to write off debts and let these countries start over with little liabilities. Consequently, in 2012, foreign debt-to-GDP ratios in SSA were only 30% - far below the IMF warning threshold of 40% (Were, 2018). Nevertheless, increased demand for credit and the sudden supply boost of Chinese credit swept the debt ratios back to 58% in 2020 (IMF, 2021). The annual flows of Chinese investments in Africa increased from USD 75 million in 2003 to USD 2.7 billion in 2019 (China Africa Research Initiative, 2021). Whether or not Chinese development banks and SOCs comply with "debt-trap diplomacy", the situation is worrying for many SSA countries (Strohecker, 2020). First, it is

[1] Recent reports somewhat downplay PRC's alleged foul strategies by explaining that the "takeovers" of properties in these three countries are rather a (very) long-term lease contract to Chinese firms to generate income to repay foreign debts (Brautigam, 2020; Carmody, 2020; Singh, 2020). According to these authors, the problematic situations that the recipient countries are in, are mainly due to poor strategies and coordination of the Chinese development financing systems, rather than malicious strategies. Furthermore, others claim that Chinese OFDI in SSA is still modest compared to European OFDI in SSA (Herbling & Li, 2020)

uncertain whether drawn credit will be used for the benefit of the average resident of the receiving country, given the SSA countries' high corruption rankings (Transparency International, 2019; Were, 2018). Second, with incredibly high debts and market interest rates, a large proportion of governmental income will be dedicated to debt repayments, leaving little for government spending on domestic programmes (Coulibaly, Gandhi & Senbet, 2019). Third, both Chinese investing companies and SSA governments are not transparent regarding the amount of credits that flows into the receiving country, making it difficult to provide any risk indication of the credit worthiness of the country (Were, 2018).

2.2 The Abuja-Kaduna railway line

According to Nigeria's Debt Management Office, the country is USD 3.1 billion in debt with Chinese parties, which accounts for "*only*" 4% of Nigeria's public debt and 11.3% of its external debt (D.M.O.N., 2021), even though other media report higher amounts. The majority of the funds come from China EXIM Bank and are used for transportation infrastructure investments, such as the metro line in Abuja, the modernization of the Lagos-Ibadan railway section, the expansion of airports around the country and the Abuja-Kaduna railway line which eventually will be connected to the Lagos-Ibadan line and the coastal network to connect the entire country by a 2,733 km railway network. The entire project is anticipated to cost USD 11.1 billion (Adepoju, 2021), and current loans outstanding have been approved to raise by USD 5.3 billion in 2021 (Global Construction Review, 2021; The China Africa Project, 2021).

The Abuja-Kaduna railway line - which is central in this thesis - is an entirely new segment within the Nigerian railway network as it did not exist in the colonial network. It connects the quickly growing capital Abuja to the trading centre of the North, Kaduna. It is constructed by China Civil Engineering Construction Corporation (CCECC) and USD 500 million of the estimated cost of USD 876 million is funded by China EXIM Bank (Railway Gazette International, 2016; Railway Technology, 2020). Construction started in 2013 and finished in late 2014, whereas the segment was inaugurated in July 2016 by President Muhammadu Buhari (Railway Gazette International, 2016). The segment is 187km long, it consists of nine stations in total and the latest numbers show that approximately 3,700 passengers use the railway daily while its cargo trains transport 800 tonnes of goods per trip in one-and-a-half hours (Ndonim, 2019; Railway Technology, 2020).

Next to connecting Nigeria's capital to another city by train for the first time, the Abuja-Kaduna segment also functions as an alternative for the highway between Abuja and Kaduna. This highway is infamous for its abductions and killings which the local authorities have many difficulties solving (This Day Live, 2020). This "most dangerous highway of Nigeria" is therefore avoided by many travellers which makes travelling between these cities very time-consuming (BBC News Pidgin, 2020). The new train line brings travellers more safely from Abuja to Kaduna within two-and-a-half hours (with several stops in between) for USD 4.50 in 2018, which is USD 1.50 cheaper than taking the car alone (Africa-China Reporting Project, 2018).

2.3 The region of the new railway

The railway connects two Nigerian states: Federal Capital Territory Abuja (FCT-Abuja) and Kaduna State. FCT-Abuja experienced great economic growth in the previous decades since Abuja became the new national capital in 1991. While the country's GDP increased by 6.7% annually between 2009 and 2014, FCT-Abuja's GDP grew by a yearly 11% growth in the same period (Oxford Business Group, 2020). FCT-Abuja's population increased by 8.3% in the years before 2017, while its satellite cities (quickly growing slums around the centre) experienced population growths of over 20% annually. Despite its economic growth in the past decade, residents of Abuja and FCT-Abuja are not necessarily better off. Because of the

relatively high property prices in the city, only 20% of FCT-Abuja's population lives in Abuja's city centre. The vast majority of the state's population lives in satellite cities or townships around Abuja. Furthermore, what is now the city of Abuja used to be the ancestral land of different ethnic groups, who were forcibly removed from their lands and placed in impoverished peripheral urban areas such as Kubwa, Gwagwalada and Karu (Mohamed, 2018). In Kubwa, the biggest satellite city of Abuja, education opportunities and houses are scarce, and unemployment rates are soaring (61 % in 2012 amongst displaced residents) compared to the rest of the country (Ismail, 2014; Sylvester, 2014). Currently, FCT-Abuja suffers an unemployment rate of 40% (Statista, 2020).

The city of Kaduna also suffers from great unemployment among young, educated adults who have difficulties finding jobs in the city, according to an article published by Financial Times (Munshi, 2021). According to this article, 300,000 Nigerians per month enter the labour force while only 20% of these young Nigerians find jobs. This is particularly evident in Kaduna, a university city where the freshly graduated youth struggles to find jobs in an economy that cannot keep up with the supply of highly qualified employees. Consequently, Kaduna state has an unemployment rate of 44 % (Statista, 2020).

2.4 Female employment and education enrolment in Nigeria

The Nigerian female employment rate was 50% in 2016, but used to be around a stable 55% in the years before 2011 back to the 1990s (World Bank, 2021b). An additional 22% of the female labour force works in agriculture, while 47% of the women working in agriculture do not get paid[2], according to the DHS study used for this thesis (National Population Commission (NPC) [Nigeria] and ICF, 2019). Since 2013, there is a rising trend in the proportion of both men and women working in agriculture. Iweagu et al. (2015) found that marital status, religion and poverty rates are determinants of Nigerian female labour participation. In rural areas, married women and poor women are more likely to be employed, whereas religious women - primarily from Muslim households - are less likely to be employed. Also, Anyanwu et al. (2021) found an inverse relationship between female employment and economic growth in Nigeria. Conversely, the DHS study observed that female employment increases with household wealth, while male employment decreases with household wealth (NPC and ICF, 2019)[3].

Around 50% of all female respondents of the 2018 DHS study have been enrolled in secondary education or higher (NPC and ICF, 2019). In rural areas, 51% of the women have no education (meaning not finished primary education), whereas in urban areas this is 16%. Ibrahim, Khan and bin Ramli (2020) and Okafor (2012) found that in Northern Nigerian states, poverty is the main factor for low female enrolment in education, along with religion, child labour, early marriage and uneducated parents. Anugwom (2009) reports that in rural Nigeria, there is a negative net return to human capital investments, as women who graduate from tertiary education often cannot find jobs in their areas. The DHS study reports that the median age of women getting married for the first time is 19.1 and increases with wealth (NPC and ICF, 2019). 25% of the women have never been married and 31% of married women reported that their husbands have more than one wife. Women without education have 6.7 children on average, while educated women have between 3.4 to 5.8 children on average, depending on their education level. Women have their first birth at the median age of 19 in rural areas and 22.3 in urban areas.

[2] For non-agricultural workers, this is 10%.

[3] This is particularly strange given the observation that Iweagu et al. (2015) omit agricultural workers for calculating the female labour force participation, while the DHS survey includes agricultural workers.

3. Theoretical framework

3.1 Employment

Developing countries tend to have relatively big parts of their population working in agricultural sectors (McMilan & Rodrik, 2011). According to ILO estimates, more than 36% of Nigeria's population worked in agriculture in 2019[4]. Nevertheless, non-agricultural labour sectors are associated with much higher productivity (Gollin, Lagakos & Waugh, 2014; World Bank, 2021a). Explanations for this agricultural productivity gap - i.e., the "misallocation" of labour amongst sectors - could be found in labour market imperfections, as reasoned from papers by Lewis (1954), and Harris and Todado (1970) (Lei, Desai & Vanneman, 2019). One of these imperfections is transportation costs, which are shown to be positively related to a region's agricultural employment rate (Gollin & Rogerson, 2014; McMillan & Rodrik, 2011). Reduced transportation costs are therefore associated with increased employment in non-agricultural sectors (Bryan, Chowdhury & Mobarak, 2014; Gollin & Rogerson, 2014), reduced agricultural work (Aggarwal, 2018) and a shift from agricultural (unpaid) work to non-agricultural (paid) work, called "labour sector reallocation" (Asher & Novosad, 2016). Women also enjoy increased non-agricultural employment due to infrastructure development, but only when perspectives on women's household roles change amongst husbands and employers (Lei, Desai & Vanneman, 2019)[5].

FLFP is determined by labour demand and labour supply conditions. Labour demand conditions are generally determined by characteristics of the labour market, such as job availability and production quantity, and gender discrimination and segregation (Spierings, 2014). In developing countries, age of marriage, fertility, education opportunities, work experience, and obligations to households and family businesses or farms are factors that largely determine labour supply conditions (Brinton, Lee and Parish, 1995). Also, the working behaviour of the husband and the rest of the household plays a role. Women in countries with a clear "male breadwinner/woman caregiver model" (most developing countries) are only able to work away from the household during the daytime if the household is taken care of by others (Brinton & Lee, 2016). Therefore, increased labour demand or improved labour supply conditions do not linearly lead to increased FLFP.

Economic theory endorses the non-linear relation between economic prosperity and FLFP. Women in low economic tiers with subsistence employment are more likely to work than women in slightly higher economic tiers, who can rely on their husband's income and focus on household tasks and childcare (Kapsos, Bourmpoula & Silberman, 2014). On a larger time and geographic scale, studies show U-shaped relationships between FLFP and economic development (Goldin, 1995; Idowu & Owoeye, 2019). The theory states that undeveloped (informal) economies allow women to combine low-productive work with household caretaking (Goldin, 1995). In industrialising economies, FLFP drops due to a lack of educational attainment, knowledge and skills related to the economy's new industries. FLFP increases again in more industrialised economies due to an increase in women's educational attainment, lower fertility and changed perceptions of a woman's role in the household and labour market.

3.2 Education

Female educational attainment is determined by prospects of net returns to education and factors such as early marriage, early childbearing and gender stigmas in families (Adukia, Asher & Novosad, 2020; Lloyd & Mensch, 2008). How human capital investments are affected by changing labour market opportunities is theoretically explained by Becker (1954) (Adukia, Asher & Novosad, 2020). There is a trade-off between

[4] As comparison, in the EU and the US this is 5% and 1% respectively, and the world's average is 26%

[5] To the best of my knowledge, this paper is the only serious study that empirically investigated FLFP alterations as a result of transportation infrastructure development in developing countries

long-run net returns to education and short-run return to labour. It is believed that remote (rural) areas have lower unskilled labour wages and skilled labour wages than more urbanized areas. Once remote regions are connected to external markets in urban areas via infrastructure development, both unskilled wages and returns to education are likely to go up. Higher unskilled wages are expected because real wages are expected to increase with local goods prices due to infrastructure construction. Net returns to education increase due to the higher-skilled wages that can be found in external markets and lower transportation costs. Higher unskilled wages cause opportunity costs of schooling, which decreases the demand for education. But increased net returns to education increase demand for education, which leaves us with an undecisive answer to the question of whether infrastructure increases schooling within this framework. The answer would depend on the size of the opportunity costs and the difference between skilled and unskilled wages (Adukia, Asher & Novosad, 2020). This context-dependency is why empirical studies found diverging education effects of infrastructure development (Adukia, Asher & Novosdad, 2020; Aggarwal, 2018; Chaudhary & Fenske, 2020).

3.3 Hypotheses

The hypotheses are designed along the possible livelihood decisions of women in the dataset. For the rural subgroups, agricultural labour participation and non-agricultural labour participation are investigated, but these variables are left out for the urban subgroups as there are too few women present in urban subgroups who work in agriculture (see Table 1 in section 4.2). The same argument holds for education effects: education enrolment will not be investigated for the 30-and-older subgroups, since there are too few women in this age group enrolled in education in the samples. Even though theory and empirical studies do not unanimously point out the expected effects of infrastructure development, with this thesis' hypotheses I assume positive employment and education effects. For the hypotheses, expected effects are equal for all subgroups, even though it is likely that the outcome variables vary amongst different subgroups. For example, older women might be more likely to work since they are more likely to have older children who need less caretaking, but also less likely to work as they are more likely to live in households with traditional gender roles. Reduced transportation costs and increased safety of travelling are expected to improve the allocation of labour (in non-farm sectors) and willingness to study. This leads to the following hypotheses:

H1A: Rural women under 30 years old in treatment areas are more likely to have (paid) employment

H1B: Rural women under 30 years old in treatment areas are more likely to work in non-agricultural sectors and less likely to work in agricultural sectors

H1C: Rural women under 30 years old in treatment areas are more likely to be enrolled in secondary education or higher

H2A: Rural women 30 years and older in treatment areas are more likely to have (paid) employment

H2B: Rural women 30 years and older in treatment areas are more likely to work in non-agricultural sectors and less likely to work in agricultural sectors

H3A: Urban women under 30 years old in treatment areas are more likely to have (paid) employment

H3B: Urban women under 30 years old in treatment areas are more likely to be enrolled in secondary education or higher

H4: Urban women 30 years and older in treatment areas are more likely to have (paid) employment

4. Method

In this section, the method that has been used to test the hypotheses is described. This method section continues with a description of the treatment sample selection procedure, the formulation of dependent variables, an explanation of the used empirical method and strategy, a description of the formation of covariates used for the empirical strategy and the descriptive statistics of these covariates.

4.1 Sample selection

The DHS Program frequently collects data on health and population in developing countries. In these surveys, a range of questions regarding socioeconomic issues, population, personal health and children's health are targeted primarily at women. Furthermore, geographic, economic and geospatial data per cluster is given, including GPS locations. The seventh survey wave in Nigeria was conducted between August and December 2018 and 40,427 households were interviewed spread across 1,389 clusters covering all 36 states and FCT-Abuja. 41,821 women between the age of 15-49 and 13,311 husbands between the age of 15-59 were interviewed, covering almost all eligible women and men of the selected households: response rates were 99.2% and 99.3% respectively (NPC & ICF, 2019).

The current research investigates employment and education effects on women living 'close' to the new stations along the Abuja-Kaduna railway. By using several sources combined with Google Maps, the following eight train stations that form the Abuja-Kaduna segment were pinpointed: Abuja (Idu), Kubwa, Gami, Jere, Rijana, Dutse, Kakau, and Kaduna (Rigasa) (Malumfashi, 2020; NRC, n.d.). Thanks to cluster GPS data the linear distance between the centre of the DHS cluster to the nearest station can be calculated. Rural and urban clusters cover an area with a radius of 10km and 2km respectively. In this research, a linear 21km distance between the train station and the cluster centre is used as a cut-off value that determines whether a cluster is sufficiently 'close' to the station to be selected as a treatment cluster. In similar research, Jedwab and Moradi (2016) used cell dummies capturing rail connectivity with 0-10, 10-20, 20-30, 30-40 km radii and Okoye, Pongou and Yokossi (2019) used a 20km cut-off value for similar research on railway effects in Nigeria. Considering this, 21km has been chosen as a reasonable cut-off value to balance between keeping a sufficiently large sample size and not losing any treatment effects. One exception to this cut-off value has been made in the region around the capital Abuja. Here, maintaining a 21km distance to a station for assigning a household to be 'treated' will cause a loss of treatment effects, since travelling 21km would also bring the traveller to the city centre of Abuja, where it is likely that the traveller finds her place of destination or someplace similar. Travelling 21km to a station would not gain the traveller any advantages in this case. Therefore, I chose to maintain a 7km radius in the area that covers Abuja's slums and satellite cities but is sufficiently far from the city centre that a train connection would be beneficial for the traveller. This exemption holds for treatment clusters that are close to the stations Abuja and Kubwa. Based on the treatment selection radii, 282 rural women spread across seven clusters are selected to form the treatment rural sample, and 608 women spread across 17 clusters make up the treatment urban sample. Rural households and urban households are treated separately in this research, since urban and rural women are hardly comparable (see section 4.5).

4.2 Descriptive statistics of outcome variables

In this section, descriptive statistics of outcome variables are given per subgroup. In the tables below, the upper six variables relate directly to the above-mentioned hypotheses. The lower six variables are included additionally to the six main outcome variables, to provide a more inclusive picture of the changes in

livelihoods of the considered women. In chapter 5, an explanation is given of why these additional outcome variables are included.

Table 1. Outcome variables descriptive statistics of treatment subgroups

	Rural younger than 30		Rural 30 and older		Urban younger than 30		Urban 30 and older	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Employed [a]	.422	.495	.741	.44	.404	.492	.803	.399
Employed and paid in cash [b]	.361	.482	.603	.491	.352	.478	.785	.411
Employed in agri-sector	.078	.269	.267	.444	0	0	.014	.118
Employed in non-agri-sector	.343	.476	.474	.501	.404	.492	.789	.409
Enrolled in secondary education or higher	.199	.4	0	0	.355	.479	0	0
Neither working nor enrolled in education	.398	.491	.259	.44	.324	.469	.197	.399
Migrated from other state in previous 4 years [c]	.09	.288	.017	.131	.164	.37	.113	.317
Currently married	.639	.482	.948	.222	.352	.478	.838	.369
Husband has work [d]	.981	.137	.927	.261	.991	.097	.962	.192
Husband works in agri-sector	.349	.479	.518	.502	.028	.167	.068	.251
Husband works in non-agri sector	.632	.485	.409	.494	.962	.191	.895	.308
Number of births in previous 5 years	.976	1.003	.922	.926	.469	.772	.623	.777
Observations	166		116		324		284	

[a] Dummy variable that gives a value of 1 if the respondent worked in the previous twelve months

[b] Dummy variable that gives a value of 1 if the respondent worked in the previous twelve months and earned cash payments

[c] Dummy variable that gives a value of 1 if the respondent migrated from another state in the previous four years

[d] Dummy variable that gives a value of 1 if the respondent's husband worked in the previous twelve months

4.3 Matching

In this research, treatment assignment is assumed to be non-random, since the location of railway placement is most likely determined based on the economic potential of an area (Banerjee, Duflo & Qian 2020). Investigating treatment effects when treatment assignment is endogenous causes a risk of selection bias. Also, the only usable dataset provides post-treatment data, which makes selecting a control group on unobserved time-invariant characteristics impossible. The DHS dataset includes a large set of post-treatment observables. Therefore, I chose to select control groups by using matching. With matching, one assumes conditional independence or unconfoundedness, meaning that treatment assignment is determined by observables, and that these observables are not affected by treatment while potential treatment outcomes are independent of treatment assignment conditional on observables X_i (Khandker, Koolwal & Samad, 2009; Rosenbaum & Rubin, 1984). In other words, potential outcomes Y_i^0 and Y_i^1 are, before treatment, independent from treatment assignment T_i (whether subject i lives close enough to a new station) conditional on observables X_i, meaning that once I controlled for observables X_i, T_i can be assumed to be random: $Y_i^0, Y_i^1 \perp T_i \mid X_i$. This implies, in the context of the thesis, that railway placement T_i should be determined by observables X_i drawn from the available datasets, while these observables are not affected by the railway placement, and the potential outcomes in employment and higher education enrolment are independent of the railway placement, conditional on these available observables. Treatment and control groups should become equal on several observables that possibly determined treatment assignment and potential outcomes, while treatment assignment cannot be confounded by other factors than these selected observables.

Matching on observables that determined treatment placement will be difficult, as it is not precisely known which factors lead to choosing the specific location of the railroad placement. However, if I can match on economic characteristics that possibly relate to factors determining infrastructure placement decisions, and on characteristics that relate to employment and education decisions, the selection bias will be overcome as much as possible, and the treatment can be perceived to be random. Railroads connect populous places and markets and are placed based on the profits anticipated by construction and operating companies, which relates to the demand for transportation in the area. Treatment placement could therefore be indirectly determined by observables, such as population growth, pre-treatment travel time to populous places and wealth of the area. These economic factors could be drawn from the DHS dataset and will be used to match treatment clusters to control clusters, along with other observables that may be related to outcome variables.

The matching method will be performed by using nine cluster-level observables. Cluster-level (rather than individual-level) covariates are used, since cluster-level data is assumed to be more related to economic factors that determined railway placement than characteristics of individual women. The intention here is to select control clusters that have had similar chances of receiving a railway line. The selection and justification of these observables are explained in section 4.5. Using these observables, the matching method will distil control units out of the DHS dataset's 16,376 untreated urban and 23,958 untreated rural households. The selection will take place on a cluster level, which forms step 1 of the empirical strategy. Step 2 consists of reweighting the selected control units with a balancing method. These empirical steps are explained in the following section.

4.4 Empirical strategy: Entropy balancing

To omit selection biases, it is necessary to control for as many observables X_i that may influence outcomes Y_i^0, Y_i^1 and treatment assignment T_i as possible. Matching on more than two covariates causes a risk of the curse of dimensionality, which makes a conventional method such as Propensity Score Matching (PSM)

unusable (Hainmüller, 2012). By using a large set of covariates with relatively few treatment subjects, PSM has difficulties balancing all the covariates to generate a comparative control group, causing the on-support untreated sample to have a very large size with very small weights (Hainmüller, 2012).

Entropy balancing is proposed by Hainmüller (2012) as an alternative balancing method to adjust the possible inaccurate covariate balancing generated by common matching practices. The method is designed for observational studies with binary treatments. It allows the researcher to obtain high degrees of covariate balancing while the unit weights vary across every control observation but stay as close as possible to base weights to retain information. In this research, the STATA-command *ebalance*, written by Hainmüller and Xu (2013), has been used for both the selection of the control group in step 1 (on cluster-level) and the unit reweighting for the regression analysis in step 2 (on individual level).

As mentioned, the empirical strategy of this thesis consists of two steps. First, a control group is selected based on the value of the weights that are generated by using the *ebalance* command for the cluster-covariates. This means that the majority of DHS clusters are dropped in the first round when the entropy-balanced weights do not cross a certain threshold (i.e., when clusters diverge too much from treatment clusters in the balancing covariates). This threshold lies at weights that are larger than 0.02 and is chosen, rather arbitrarily, to ensure a sufficiently large control sample. However, the control sample will remain substantially larger than the treatment sample, as the units are assigned new weights in the second round, which will result in a sum of weights lower than the control sample size. The number of selected rural and urban control clusters is 46 and 99 respectively, yielding a total rural and urban sample size of 1,647 and 3,550 respectively. I chose to include this step to reduce the size of the control group and only use the clusters that are most similar to the treatment clusters.

Second, the selected control-units will be reweighted again by using the *ebalance* command, but now individual-level covariates are added to the cluster-level covariates, to ensure that the control group is well-balanced in individual characteristics and a representative comparison unity for the treated sample. The list of covariates used for the second round of entropy-balancing, therefore, includes both individual- and cluster-level covariates to capture both the observables that (likely) determined treatment placement and covariates that need to be balanced for comparable samples as they are likely to affect outcome variables. The weights calculated in the second step are used as sampling weights for the regression of treatment status T_i on each dependent variable Y_i.

4.5 Covariates

In this section, descriptive statistics on selected matching covariates X_i are given. Step 1 of the empirical strategy is matching treatment clusters to control clusters. For this, covariates are selected of which the statistics of the treatment groups are shown in Table 2. Step 2 of the empirical strategy is reweighting (*ebalance)* observables to create control groups that are comparable to the treatment subgroups. Treatment groups' statistics on covariates used for reweighing are shown in Tables 3 and 4. Balance tables to show the differences between control and treatment groups pre- and post-reweighing are given in Appendix tables 1 and 2.

Table 2. Cluster descriptive statistics of all rural and urban treatment clusters

	Rural		Urban	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Cluster's travel time to 'populous place' in hours [a]	.576	.332	.055	.147
Linear distance to station [b]	11.582	5.607	7.451	4.719
Log of cluster's gross cell production [c]	7.402	.013	7.41	.016
Cluster's population growth between 2010 and 2015	.308	.161	.312	.195
Proportion of women with primary education [d]	.624	.287	.883	.184
Proportion of women who migrated more than 4 years ago [e]	.138	.087	.252	.116
Cluster BUILT population in 2014 [f]	.023	.015	.536	.219
Slope [g]	.832	.202	1.077	.605
Proportion of Islamic women [h]	.514	.378	.584	.381
Cluster average wealth index [i]	-9.11	35.623	119.005	59.834
Observations	282		608	

[a] The average time in hours required to reach an urban centre with "high density", based on infrastructure data from 2015.

[b] Linear distance from the centre of the cluster to the closest new station along the Abuja-Kaduna line. Naturally, this variable has not been used for matching, as this variable defined the treatment sample.

[c] The natural logarithm of the average Purchasing Power Parity in 2005 US Dollars within the cluster. The GCP shows the economic activity of the area and is related to the GDP.

[d] Proportion of eligible women that mentioned that they have never followed any form of education. It is assumed that this covariate has not been affected by the treatment, as the minimum age of a respondent is 15 years old, and it is unlikely that these uneducated women decided to follow primary education in the previous three years.

[e] This covariate has been calculated by counting the number of women in a cluster who live for more than 4 years in the same location, but have migrated to the cluster from another state in the past. The threshold is more than 4 years, because the construction of the railway started in 2013.

[f] Indicating how "urban" a certain cluster is with a value between 0.00 (very rural) to 1.00 (very urban) in the year 2014. The index is based on the cluster's population density and location and is obtained from the Global Human Settlement Layer (GHSL) project that produces global spatial information on human presence with satellite imagery (Pesaresi et al., 2015). Human presence includes not only physical presence of people, but also infrastructure, buildings, farms and other aspects of human influence on the cluster. The larger this human footprint, the higher the Built Population index.

[g] The average slope of a cluster, which relates to the convenience of constructing a railway through the area.

[h] The division of Islamic and Christian households in the dataset is approximately 50-50, which is why this Islam-related discrete value functions as a religion-covariate. This variable indicates the proportion of Islamic households in the cluster. Islamic women in Nigeria are generally less well-off (NPC & ICF, 2019), so using this covariate means including additional information in the balancing methods.

[i] The wealth index has been created by the DHS Program as an indication of a household's wealth relative to the average of the country. Generally, it is based on a household's assets (car, motorbike, television, etc.), materials used for house construction, and water access and sanitation facilities. Since these components do not change suddenly, a household's wealth index is assumed to change with some delay after an occurrence such as suddenly increased mobility. However, it is unrealistic to assume that the wealth index does not change at all within the four years between railway construction and the moment of surveying. It is, on the other hand, also unlikely that within four years increased mobility affected labour activity. Especially the components indicating the constructed characteristics of the house (type of floor and roof, sanitation facilities, etc.) are unlikely to be influenced by the increased mobility and, in turn, influence the household member's decisions regarding work, school or moving away, all within four years.

Table 3. Individual descriptive statistics of rural treatment women

	Younger than 30		30 and older	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Age	21.602	4.449	37.164	5.736
Primary education	.717	.452	.491	.502
Number of children older than 5 [a]	.687	1.100	4.431	2.243
Islamic	.566	.497	.44	.498
Wealth index	-6.69	64.08	-12.573	66.437
Migrated from another state more than 4 years ago	.084	.279	.216	.413
Observations	166		116	

[a] This individual covariate provides insights into fertility, a commonly used indicator of the economic condition of a region. Only the children older than 5 have been counted, as younger children could theoretically have been born as a result of the treatment.

Table 4. Individual descriptive statistics of urban treatment women

	Younger than 30		30 and older	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Age	21.219	4.42	38.651	5.87
Primary education	.923	.267	.838	.369
Number of children older than 5	.324	.808	3.475	2.666
Islamic	.645	.479	.514	.501
Wealth index	117.945	70.973	120.215	80.349
Migrated from other state more than 4 years ago	.157	.365	.359	.481
Observations	324		284	

5. Results

For clarity reasons, the results are handled per subgroup. This section is structured as follows: regression results for the different subgroups related to the hypotheses are given first. Then, additional regressions on immigration, husband employment, marriage and fertility are given.

5.1 Rural women under 30 years old

Table 5a regards the first livelihood choice (study, work or stay home) for younger women living in rural areas. It shows that the placement of the railway line did not affect secondary or higher education enrolment (p=.872) but did affect general employment negatively and increased the probability that women neither work nor study. Living in the treatment area is namely associated with an average 10.5% lower probability of being employed (p=.016) and an average 7.31% higher probability of neither working nor studying (p=.09) compared to the reweighted control group. Even though the latter effect is not statistically significant at conventional significance levels, given the low sample sizes, the p-value suggests at least some effect, but it cannot be declared with much certainty. The results indicate that there is no evidence to support H1A of increased female employment due to the railway placement – the opposite has been found. Also, there is no evidence to support H1C of increased education enrolment for this subgroup.

Regarding H1B, table 5b shows that the phenomenon of female labour sector reallocation from agricultural to non-agricultural work is not found for younger women living in rural areas. Agricultural employment in treatment areas is on average 12.7% lower (p=.000), whereas non-agricultural employment is not different from the reweighted control group (p=.599). This clarifies the results in table 5a of decreased total employment: agricultural employment decreased, and non-agricultural employment is unchanged. Furthermore, paid employment did not change (p=.763), indicating that paid employment and agricultural work do not behave in sync which suggests that agricultural labour is generally unpaid.

Table 5a. Education and employment effects of rural women under 30

	(1) Enrolled in higher education	(2) Has employment	(3) No education enrolment and not employed
Treatment	0.00567	-0.105*	0.0731
	(0.872)	(0.016)	(0.090)
_cons	0.193***	0.527***	0.325***
	(0.000)	(0.000)	(0.000)
N	879	879	879

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5b. Employment quality effects of rural women under 30

	(1) Works in agri-sector	(2) Works in non-agri-sector	(3) Gets paid for employment
Treatment	-0.127***	0.0219	0.0128
	(0.000)	(0.599)	(0.763)
_cons	0.206***	0.321***	0.349***
	(0.000)	(0.000)	(0.000)
N	879	879	879

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.2 Rural women of 30 years old and over

With hypotheses H2A and H2B, increased (paid) employment and labour sector reallocation from agricultural to non-agricultural, respectively, were expected for rural women of 30 and over. Table 6 shows no statistically significant effects within conventional significance levels in terms of labour quality and quantity. However, the signs and p-values in columns 2 and 3 suggest uncertain but possibly interesting movements of reduced agricultural work and increased non-agricultural work with comparable magnitudes (p=.099 & p=.142 resp.). Given the small sample size, the high p-values should not be a reason to reject labour force sector reallocation instantly. Logically, the general employment participation rate does not change (p=.967), but, surprisingly, paid employment does also not change (p=.457), suggesting a contradiction for the above-mentioned suggestion that agricultural work is unpaid and non-agricultural work is paid, if the assertions of labour force sector reallocation are true.

Table 6. Employment effects of rural women of 30 years and over

	(1) Has employment	(2) Works in agri-sector	(3) Works in non-agri-sector	(4) Gets paid for employment
Treatment	-0.00190	-0.0773	0.0754	0.0379
	(0.967)	(0.099)	(0.142)	(0.457)
_cons	0.743***	0.344***	0.399***	0.566***
	(0.000)	(0.000)	(0.000)	(0.000)
N	768	768	768	768

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.3 Urban women

For the support of hypotheses H3A and H3B, which expect increased employment and education enrolment for young urban women, no evidence has been found. There is no education enrolment effect (p=.659), no alteration of the labour force participation (p=.934) and no changes in quality of labour in terms of payment (p=.530) as a result of the treatment placement (Table 7). For urban women aged 30 years and over, also no effects are found. P-values in Table 8 show no effects regarding employment rates (p=.919) or paid-employment rates (p=.216), leaving no evidence for support of H4.

Table 7. Education and employment effects of urban women under 30

	(1) Enrolled in higher education	(2) Has employment	(3) Gets paid for employment	(4) No education enrolment and not employed
Treatment	0.0142	-0.00270	0.0198	0.0104
	(0.659)	(0.934)	(0.530)	(0.743)
_cons	0.341***	0.407***	0.332***	0.314***
	(0.000)	(0.000)	(0.000)	(0.000)
N	1903	1903	1903	1903

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8. Employment effects of urban women of 30 years and older

	(1) Has employment	(2) Gets paid for employment
Treatment	-0.00299	0.0381
	(0.919)	(0.216)
_cons	0.806***	0.747***
	(0.000)	(0.000)
N	1647	1647

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.4 Immigration

To provide more clarity on the results and the underlying mechanisms causing them, immigration effects will be analysed additionally. Transportation infrastructure development could either increase migration to the area - as increased economic activity due to increased connectivity lure people from rural areas (e.g., Diaz, Behr & Ng, 2016) - or decrease migration to this area, as households can live more remote from economic active areas and commute instead of live there (e.g., Xu & Sun, 2021). Unfortunately, with the available data, there is no way of knowing the number of respondents who migrated away from the treatment or control clusters: the DHS data only provides the previous state of residence, not the previous cluster of residence which is required to know whether a respondent used to live in the treatment area.

The dependent variable that is used to determine migration effects is constructed such that it gives a value of 1 if the respondent moved from another state to the current state of residency within the previous four years, and 0 if not. This means that if this variable is regressed on treatment, a comparison will be made between the fraction of women who migrated from different states to the treatment clusters versus the fraction of women who migrated from different states to the control clusters. The outcome is shown in Table 9.

Table 9 shows divergent outcomes in migration effects. Younger rural and younger urban women do not seem to be affected by the treatment in their decision to migrate to the treatment area. However,

rural women aged 30 and over in treatment areas are on average 3.39% less likely to have been migrated to the current place of residency in the previous four years (p=.029), while urban women aged 30 and over in treatment areas are associated with an average 4.57% higher probability to have been migrated (p=.048). Note that this does not mean that rural respondents over 30 are migrating away from the treatment area: this variable only counts the number of women who migrated *to* the area.

Table 9. Immigration effects per subgroup

	(1) Rural under 30	(2) Rural 30 and older	(3) Urban under 30	(4) Urban 30 and over
Treatment	-0.0198	-0.0339*	0.0172	0.0457*
	(0.440)	(0.029)	(0.514)	(0.048)
_cons	0.110***	0.0511***	0.146***	0.0670***
	(0.000)	(0.000)	(0.000)	(0.000)
N	879	768	1903	1647

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Now, it is worth investigating whether the found treatment effects differ among migrated and unmigrated treatment respondents. For this, it has been decided to compare the main results (employment and education) with the outcomes from the same regressions executed only with respondents who did not migrate in the previous four years. The reason for this is that the fraction of respondents who migrated in the previous four years is low: between 1.9% and 17% depending on the treatment subgroup. Performing regressions with these fractions of the treatment samples will not provide insightful results, as these samples would be very small. Comparing main results with results when only non-migrants are considered is deemed to provide sufficient information regarding possible different effects with different migration statuses. Appendix tables 3 to 6 show that there are hardly any differences in outcomes if only the non-migrants are included. The only different outcome is that non-migrated urban women over 30 in treatment areas appear to be more likely to have paid employment (p=.042), while the general employment rate remained unchanged. Table 8 in section 5.3 showed less strong effects in paid employment when also migrated older urban women are considered. This is the only case out of the four subgroups in which non-migrants differ from whole subgroups, and there is no direct reason why employment quality only changes for older urban women if non-migrated women are considered.

5.5 Marriage, husbands and children

Given the above-mentioned results, it is possible that women are less or not more likely to work since their husbands can use new ways of transportation to find more productive and higher-income work elsewhere, resulting in decreased necessity for their wives to work for low- or no-wage jobs[6]. However, marital status could be influenced by the treatment. Therefore, it is first necessary to see whether there has been a marriage effect for any of the subgroups. Subgroups with respondents older than 30 are included - even though it is unlikely that the treatment will affect their marital status at this age[7] - to check whether there appear unanticipated differences between control and treatment groups that are unlikely to be caused by the treatment.

[6] Which was the assumption composed from the theory on lower employment rates amongst women in higher economic tiers (Kapsos, Bourmpoula & Silberman, 2014)

[7] The DHS data shows that more than 80% of women between 25 and 29 are married, whereas the proportion of married women per age group does not exceed 87%. Therefore, it is somewhat unlikely that after the age of 30 a respondent marries for the first time and that this decision was affected by the treatment.

Table 10 shows that the treatment did not affect marriage decisions; there is no support for changing average marital statuses within all subgroups. To investigate whether the main outcomes differ for different marital statuses, the main regressions are performed again by marital status. Results are provided in the appendix (Appendix tables 7 till 10 and Appendix balance table 11 to compare treatment and reweighted control groups for only married women). Here, it appears that there are no differences in the measured effects if regressions for married and unmarried women are done separately. The only slight difference is that young married women living in urban areas are slightly more likely to be enrolled in education than their control group (p=.054)[8], whereas this effect does not exist within the same-aged unmarried group in urban areas. This result will be further discussed in 6.4.

Table 10. Marriage effects per subgroup

	(1) Rural under 30	(2) Rural 30 and older	(3) Urban under 30	(4) Urban 30 and over
Treatment	0.0325	0.0204	-0.00974	0.00353
	(0.445)	(0.384)	(0.761)	(0.904)
_cons	0.606***	0.928***	0.362***	0.835***
	(0.000)	(0.000)	(0.000)	(0.000)
N	879	768	1903	1647

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Now, employment effects for husbands in all four subgroups are analysed. Table 11 shows that for the treatment subgroups rural under-30, urban under-30 and urban over-30, their husbands are statistically significantly more likely to be employed (p=.023, p=.022 and p=.035 resp.). Their husbands are respectively 4.5%, 4.1% and 4.6% more likely to have employment on average. Subsequently, it is investigated whether husbands experience differences in labour in the agricultural and non-agricultural sectors. For every subgroup, the changes in employment rates in both sectors are analysed in tables 12a and 12b[9]. Table 12a shows that only the husbands of young rural wives are statistically significantly less likely to be working in the agriculture sector. On average, they are 18.6% less likely to be employed in the agricultural sector (p=.001). There are no statistically significant differences in the composition of the agricultural labour force for the three other subgroups.

Looking at table 12b, it appears that the treated young-rural husbands are more likely to be working in non-agricultural sectors (p=.000 with a magnitude of .231). The older-rural and young-urban subgroups also show weak positive effects in non-agricultural labour participation that, despite the relatively high p-values, are worth mentioning (p=.163 and p=.103 resp.).

Table 11. Employment effects of husbands

	(1) Rural under 30	(2) Rural 30 and older	(3) Urban under 30	(4) Urban 30 and over
Treatment	0.0451*	0.0229	0.0409*	0.0460*
	(0.023)	(0.428)	(0.022)	(0.035)
_cons	0.936***	0.904***	0.950***	0.916***
	(0.000)	(0.000)	(0.000)	(0.000)
N	527	703	703	1405

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[8] Note that the urban-under-30 treatment group and the reweighted control group differ in a few covariates (see Appendix table 11). Because of this, the urban-under-30 control group has been reweighted with only married women. The regression shows that the effect appeared to be similar as described above: married urban women under 30 are on average 4% more likely to be enrolled in higher education (p=.032, see Appendix table 12)

[9] As agricultural work rate amongst urban husbands lies slightly higher than for women (which is in the case of young urban women non-existent), the decision has been made to check for labour force reallocation amongst urban men as well.

Table 12a. Agricultural employment effects of husbands

	(1) Rural under 30	(2) Rural 30 and older	(3) Urban under 30	(4) Urban 30 and over
Treatment	-0.186***	-0.0495	0.000376	0.0129
	(0.001)	(0.353)	(0.983)	(0.496)
_cons	0.535***	0.568***	0.0279***	0.0546***
	(0.000)	(0.000)	(0.000)	(0.000)
N	527	703	703	1405

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12b. Non-agricultural employment effects of husbands

	(1) Rural under 30	(2) Rural 30 and older	(3) Urban under 30	(4) Urban 30 and over
Treatment	0.231***	0.0724	0.0406	0.0331
	(0.000)	(0.163)	(0.103)	(0.236)
_cons	0.401***	0.337***	0.922***	0.861***
	(0.000)	(0.000)	(0.000)	(0.000)
N	527	703	703	1405

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In addition to investigating marital status and effects on husbands, young women's fertility changes will be investigated as well. Treated rural women tend to stay home more, which could be related to increased fertility. Married urban women are more likely to be enrolled in secondary education or higher, which could be a result of reduced fertility. Nevertheless, Table 13 shows no statistically significant alterations in fertility amongst most subgroups. One possible exemption is urban women over 30 years old who bore fewer children in the previous five years (p=.058).

Table 13. Changes in women's number of births in the previous five years

	(1) Rural under 30	(2) Rural 30 and older	(3) Urban under 30	(4) Urban 30 and over
Treatment	0.119	0.0685	-0.0267	-0.113
	(0.172)	(0.465)	(0.604)	(0.058)
_cons	0.857***	0.854***	0.496***	0.736***
	(0.000)	(0.000)	(0.000)	(0.000)
N	879	768	1903	1647

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6. Discussion

The Discussion section will be structured as follows. First, the found results are interpreted, explained and clarified to make sense of the underlying socioeconomic mechanisms that steered them. Second, the found results are reflected in female labour participation theories. Third, the current study's limitations are discussed and suggestions for future studies and policies are given.

6.1 Labour force reallocation in rural areas

One of this study's main hypotheses is the presence of rural female labour sector reallocation from agriculture to more productive and higher income-generating non-agricultural sectors because of increased mobility to remote (and more productive) labour markets. On the contrary, rural women under 30 were

negatively affected by the railway placement with regards to employment, and this effect runs primarily through the significantly lower participation in agricultural work since non-agricultural work participation hardly changed. Paid employment did not change, which is in line with the assumption that rural agricultural work is more likely to be unpaid. Leaving agricultural work will therefore not influence the paid employment rate. Surprisingly, these women did not reallocate their time to education, as enrolment in secondary and higher education also remained unchanged. Considering this, young rural women appear to spend less time on work and study combined compared to their control counterparts, but it remains unclear how these women do allocate their time.

Interestingly, results on rural women of 30 and older do suggest (with relatively high p-values[10]) a labour force sector reallocation from agricultural to non-agricultural labour. Simultaneously, the total employment rate of these women remained unchanged. There seems to be a difference between younger and older rural women that causes different responses to the railway placement. One determining difference could be the respondent's household position (mother or daughter): 97.4% of the rural women of 30 and over are either household head or wife of the household head, whereas from the rural women under 30, 62.7% are household head or wife while 30.1% are daughters. Perhaps daughters respond differently to the treatment than mothers. However, when only married women are included in the regressions, the results remain the same. If it is assumed that mothers are predominantly married and daughters are not, household position does not drive this difference.

Another explanation could be the fact that young Nigerians suffer way more from job insecurity, unemployment and low returns on education than older Nigerians (National Bureau of Statistics, 2018). Across the entire DHS dataset, (rural) women under 30 are less likely to be employed than (rural) women of 30 and over: the employment rate is 56% for younger rural women versus 81% for older rural women. When corrected for education enrolment, 66.5% of the younger rural women are either working or studying, versus the 81% of older rural women. Moreover, when working, 29.8% of the younger rural respondents reported not getting paid for work, while this is 20.0% for the older group who are similarly distributed amongst the labour sectors. As mentioned in chapter 2, unemployment and job insecurity amongst the young is an increasing problem in Nigeria. Nevertheless, if the railway placement reduced total employment for younger rural women, and only a labour force sector reallocation instead of total employment increase occurred for older rural women, it can be concluded that the railway placement did not benefit the already unemployed. Increased mobility did not lead to increased employment rates, perhaps simply because job opportunities in more remote places are also absent. A train line, logically, cannot increase employment rates in remote labour markets if these labour markets have an abundant supply.

On the other hand, the treated rural women's husbands appear to experience somewhat opposite effects: husbands of the younger group show a clear and statistically strong labour sector reallocation from agriculture to non-agricultural jobs, as well as a total employment increase. Findings of decreased (agricultural) employment for wives and increased (non-agricultural) employment for husbands suggest that railway placement affects household structures in terms of work division. If a husband's new income is higher and more stable than the wife's disappeared income, it might even benefit the household as a whole financially. However, studies suggest that not earning individual income decreases young women's empowerment, happiness and their children's welfare (Anderson & Eswaran, 2009; Koolwal & Van de Walle, 2013; Van den Broeck & Maertens, 2017). The question then remains whether this financial "benefit" for the household is also tangible for the household's wives individually. Lower employment rates for young rural wives could be a consequence of increased (non-agricultural) employment of their husbands, as the wife's relative income decreases and the necessity for the wife to work diminishes once her husband finds a job with a stable income. Older rural women might experience this mechanism to a lesser extent, as they are more likely to have older children and therefore a lesser necessity to stay home and take care of the

[10] P-values that showed a decrease in agricultural work and increase in non-agricultural work were relatively high for conventional standards, but given the small sample sizes, the effects could not be ignored.

household. Theories on the employment behaviour of married women when their economies develop vis-a-vis the findings of this thesis will be further discussed in 6.4.

To summarise, the railway placement negatively affected the total FLFP of rural women under 30 through the withdrawal from agricultural work, while their husbands switched from agricultural to non-agricultural work and gained total labour participation. Explanations for this could be found in the reduced necessity for women to work when husbands yield higher and more stable incomes. Simultaneously, older rural women appear to experience a labour force reallocation, while their husbands' employment statuses remain unchanged. Possible explanations for this differing effect could be the stronger position of older Nigerians in labour markets, due to more experience and more willingness to take on lower qualified jobs (according to Akinwotu (2021) in The Guardian), or lower necessity to stay home and take care of older children.

6.2 Labour effects in urban areas

The new railway line reduces the costs and risks of travelling between the main cities of Kaduna and Abuja, and was therefore expected to contribute to the efficiency of labour allocation and the increase of employment rates. Women in urban treatment areas were expected to reach labour markets more easily, or benefit from increased economic activity lured to the area around the new train station. However, it appears that the new railway did not change the quantity and quality of employment for urban women. A reason for this could be the uncertainty whether urban labour markets have sufficient growth potential for the new railway line to affect the number of people working in these labour markets. Easing labour supply by improving infrastructure only affects employment rates if there is sufficient (potential) labour demand. As explained in 6.1 and chapter 2, Nigerian labour markets lack demand to a great extent, possibly leaving little possibility for the improved infrastructure to affect employment rates. Besides, the new railway stations did not ignite additional economic activity in the slums and satellite cities that benefited these women.

On the other hand, husbands show increased employment rates, much like their rural counterparts. Apparently, some part of the hypothesis that improved transportation infrastructure from slums or satellite cities to city centres generates increased employment is correct, but only holds for men instead of women in this context. An explanation could be that in these urban areas traditional household structures uphold, and when new labour opportunities emerge through infrastructure development, it is decided that husbands seize these opportunities and wives stay at home. Allotey et al. (2022) confirm that within urban areas in Kaduna State, fathers are still primarily seen as 'providers' and mothers as 'caregivers'. This, again, suggests that infrastructure development financially benefits the household as a whole, but fails to improve the individual lives and empowerment of women.

6.3 Secondary and higher education enrolment

Increased education enrolment due to the treatment was expected through two effects. Directly, through lower transportation costs and, therefore, lower costs of education. Indirectly, through increased expected returns to education when the railway provides improved access to more diverse labour markets. It was expected that the direct mechanism is more present amongst rural women, while the indirect mechanism is more present amongst urban women. That is, because of the assumption that urban women's travelling time to secondary education institutions before treatment is shorter than for women in rural areas.

Nevertheless, results show that in almost all cases, the treatment does not affect secondary or higher education enrolment. There are several possible explanations for finding this. First, the new railway line may not provide any noteworthy benefits over pre-existing ways of transportation to education institutions. Benefits of the new railway are insufficient if the costs of transportation remain too high,

education institutions are already easily reachable by foot or bus, or if travelling time remains too long for those living further away. Second, opportunity costs of studying are increasing as low-skilled wages increase due to the enhanced infrastructure (Adukia, Asher & Novosad, 2020). However, it is unlikely that the railway increased low-skilled wages in the treatment clusters within three years of operations. Third, perceptions on returns to education remain too low despite the treatment, considering the problematic situation of the labour market for young Nigerians discussed in 6.1. Saturated labour markets leave little motivation for young women to follow up their education, if they know that finding a job after graduation is unlikely. However, female educational attainment in urban Nigeria has been significantly growing between 2000 and 2015 (Graetz et al., 2018). Considering this, it is more reasonable to state that the treatment did not provide *additional* motivation to continue studying for urban women. However, when only married women are included, the train line appears to have some positive effect on the education enrolment rate. Apparently, increased mobility is only a determining factor for the education decision if a woman is married. This finding will be further discussed in section 6.4. Fourth, the decision of whether to continue or resume studying might be irrelevant to many young rural women who are already married and have children at the age at which this decision takes place. Young (Nigerian) women who get married and start to have children are more likely to drop out of school (Envuladu et al., 2016; Kyari and Ayodele, 2014; Lloyd & Mensch, 2008), and increased mobility might not convince them to resume school. In fact, rural women in Nigeria are on average 18 years old with their first birth and the median age of first marriage is 17.2. If the majority of rural women marry and bear children around the age of secondary education graduation, a new railway connection that reduces transportation costs is unlikely to suddenly motivate these women to resume or continue school.

To summarise, it appears that education enrolment is only affected by railway placement in the case of married urban women. For rural women, it is likely that the railway does not provide convincing benefits over existing ways of transportation, or effects will only be observable in the long run as rural women marry and give birth at such a young age that infrastructure development does not affect current women's motivation to continue studying. For urban women, increased connectivity to labour markets after graduation probably does not provide additional motivation to study when unemployment among higher education graduates is a known problem.

6.4 The U-shaped FLFP function and implications

In this section, the found results will be discussed in light of Goldin's (1995) U-shape FLFP function. To draw this theory into the context of the thesis, the assumption that railway infrastructure development is a determining component of 'economic growth' should be maintained. Goldin's (1995) FLFP theory states that the relation between FLFP and economic development is U-shaped: high in early and late stages of economic development, and low in between. This implies that women in less developed economies - positioned on the declining left side of the U - move rightwards to the bottom of the U when their economies industrialise, their relative income from agricultural or other informal work declines and negative stigmas towards working wives withhold them from searching (manufacturing) jobs in formal sectors (Goldin, 1995). Meanwhile, their husbands take advantage of industrialisation and find jobs in formal sectors, leaving little necessity for their wives to work due to the household's income effect.

First, this thesis's findings on rural young women are consistent with this theory. Women in this subgroup appear to be less participating in agricultural labour and less overall employed, while their husbands exchange agricultural labour for non-agricultural labour and show increased overall employment. They do not show increased fertility, which indicates that their tendency to stay home more often is not related to the necessity to take care of young children. Second, older rural women showed hints of labour sector reallocation from agricultural to non-agricultural labour and were not affected in overall employment rates. In the light of the U-shaped FLFP function theory, this might be surprising. Older women are likely

to be withheld from exchanging agricultural labour for labour in other sectors due to traditional gender role stigmas against female labour in formal sectors (Goldin, 1995), which are more likely to be present in older households. However, as older women have older children on average, they might be less occupied with caretaking which allows them to work in other sectors. This assumption is in line with the above-mentioned suggestion that stigmas against women working outside the house are disappearing, but that childcare tasks still keep them home[11]. It can therefore be concluded that the mechanisms explained by Goldin's (1995) theory are in this setting more applicable to younger rural women.

Third, young urban women - whose employment rate remained unchanged despite the treatment - find themselves in industrialising economies where jobs require skills and knowledge that are often not possessed by under- or uneducated women (Goldin, 1995). Goldin states that available manufacturing jobs are reserved for men, but that female labour participation rises once women "*gain education in the secondary school level*" (Goldin, 1995, p.6). Results discussed in 5.5 show that married urban women in treatment areas seem to be more likely to be enrolled in secondary or higher education and have fewer young children. Increased female education enrolment in industrialising economies can be reconciled with Goldin's (1995) theory: as mentioned, the theory states that FLFP increases once they complete secondary education and find "white collar" (service sector) jobs. Increasing education enrolment due to economic development is therefore an omen for increased FLFP. The finding implies that increased mobility is a more impactful factor in the decision to continue or resume education for married women than for unmarried women. Time and safety constraints might be factors that limit the willingness to continue schooling for wives with families at home, whereas unmarried women might be limited to continue schooling by other factors such as money constraints or negative stigmas towards studying women held by parents (when they still live at their parental house). Treated young urban women did not show altered fertility. In studies, reduced fertility is associated with increased education enrolment through increased opportunity costs of having children (Easterlin & Crimmins, 1985) and changing family values due to education (Axinn & Barber, 2001). The effect of increased education enrolment could therefore be seen as independent of family planning effects.

Acknowledging that the found results to some extent support the U-shaped FLFP function theory implies that support for the emancipation hypothesis (Shorter, 1973) or the constancy hypothesis (Bose, Feldberg & Sokoloff, 1987) is lacking. First, for the emancipation hypothesis (also called modernization theory) no support has been found. This hypothesis states that FLFP rates increase with economic development. Second, the constancy theory implies that during the industrialisation phase women continue working, but it is not accounted for in more conservative studies due to the under-numeration of women's work activity in less formal sectors. This theory cannot be supported, as in the current research informal or unpaid jobs have been accounted for, or at least were intended to be included by the DHS survey. Therefore, the U-shaped FLFP function theory is, considering this current research, the theory that best fits the findings.

6.5 Limitations

The limitations of this thesis can be divided into the limitations of the used method, the limitations of the available dataset, and decisions made based on time and resource constraints. First, regarding the method, one of the main limitations of this thesis is the extent to which the treatment group is comparable to the reweighted control group. By using entropy balancing as a matching and reweighting strategy to ensure the treatment and control group are comparable in the observable variables, I assumed that the variables that determined treatment placement were observable. However, as mentioned, the exact way how the location of placement of the railway line was determined remains unclear. Most likely, Nigerian authorities and

[11] This is, in my opinion, to a lesser extent a stigma and much more a matter of choice who is staying home to take care of children. The fact that mostly the wife stays home is rather a result of centuries old customs.

Chinese contractors decided to start building the country-wide railway connection in Abuja because of its status as a rapidly growing city in economic, business and geographic terms, while it did not have railway connections yet. Since the region of Abuja is the region that attracts most foreign investments and outperforms all other Nigerian regions in terms of urbanisation and population growth, Abuja should be handled as a case apart and it is difficult to compare economic development components in Abuja with regions elsewhere (Oxford Business Group, 2020). The assumption of conditional independence states that the employment and education outcomes should be independent of railway placement conditional on all covariates, but some essential covariates might have been missed. Even though it was intended to match and balance to the best of the available abilities, it remains uncertain whether all relevant covariates are conditioned on, as it was only possible to condition on observables included in the DHS dataset. Therefore, there might be a reason to worry that a selection bias is not entirely omitted.

Second, following this limitation is the fact that the research of effects covers a relatively short time frame between the initiation of the treatment and the moment of conduction of the surveys, with relatively small sample sizes. Using a short timeframe has a risk of not capturing the entire true effect. But having small sample sizes reduces the statistical power, which reduces the probability that the found statistically significant effects are true (Button et al., 2013). Therefore, the found results are handled and discussed with care, by acknowledging that the results are only a suggestion of causality instead of claiming that the effects are causally determined by the treatment only. Third, the allocation of respondents to treatment groups, which has been determined by whether a respondent lives within 21km from a new railway station, is arguably arbitrary. In rural Nigeria, where existing roads are unpredictably scattered around the area, there could be critical differences between the linear distance and the travel time (and safety). In practice, respondents living further away than 21km could benefit more from the train connection, if they are better connected to the nearest station than respondents living closer in distance but with poorer connection to the nearest station. "Treated" respondents are then in essence not truly treated, and vice versa. Allocating respondents to treatment groups based on the travel time to the new station would generate more representative treatment groups. Similarly, creating subgroups based on age with a cut-off value of 30 years could be perceived as questionable and random. Different socioeconomic mechanisms could be at play for women within the same subgroup but with an age difference of more than ten years.

The limited availability of data and the necessity of conditioning on a set of covariates made me decide to execute entropy balancing as a matching strategy. However, matching strategies for non-experimental economic data are commonly used in combination with other statistical methods or with other matching methods to check for robustness (e.g., in Girma & Gardebroek, 2015; Sparrow, Suryahadi, & Widayanti, 2008). In comparable studies regarding the effects of infrastructure development, multiple time periods data is preferred (e.g., in Aggarwal, 2018; Lei, Desai & Vanneman, 2019). The DHS dataset was chosen to be used because of its sufficiently large size which allowed me to include a reasonable number of respondents in the treatment group and its wide variety of socioeconomic data (contrary to, for example, LSMS data which provided very few respondents in treatment areas). Unfortunately, however, older DHS Nigeria waves did not sufficiently correspond with the 2018 wave in terms of the number of respondents per treatment subgroup, making it impossible to compare pre- and post-treatment groups. Propensity score matching could not be used, as the large set of covariates brings the risk of overspecification. Nevertheless, the large set of covariates is used for matching and balancing, as the combination of covariates was assumed to contribute to the treatment allocation and these covariates are likely to influence outcome variables. Also, using only entropy balancing for the statistical analyses is not very common and did not allow me to check for robustness. Furthermore, the DHS dataset did not provide data on some key economic variables that could have provided more information to understand the underlying mechanisms of the found effects. Examples of these variables would be household income and expenditures, personal income, migration rate of respondents leaving the area, and ownership and quantity of land used for agriculture.

In future studies, it might be worth exploring suggestions that were made based on the findings of this thesis. First, it might for example be interesting to investigate whether the labour force sector

reallocation for husbands and older rural women is associated with changing individual or household income, expenditures, health care insurance, and ownership of agricultural land. Second, it might be interesting to investigate to what extent areas that now experience improved mobility develop increased economic activity. One of the motivations for constructing the new nationwide railway line was to reduce the urbanisation pressure on cities like Lagos (Olurounbi, 2021). So, it could be worth investigating whether the urbanisation rate dropped and the economic activity that used to be attracted to cities like Lagos moved to less populous places around the new Abuja-Kaduna train line stations. Third, it might be interesting to further investigate the mechanisms at play that were relatable to Goldin's (1995) U-shaped FLFP function. For example, whether urban women do experience increased FLFP if only the educated urban women are considered, which specific labour sector do rural husbands enter (white-collar or blue-collar), and what types of stigmas cause rural women to stay home in case of industrialisation in this decade. Fourth, for some treatment groups, effects on paid employment and non-agricultural employment did not behave equally. This indicates that deeper investigation into the effects of increased mobility on labour quality aspects is required to better understand mechanisms that drive female income.

7. Conclusion

This thesis aimed to investigate female education and employment effects of railway development in the context of the new Abuja-Kaduna railway line in Nigeria. It was hypothesised that (1) in rural areas nearby the infrastructure development project, female labour force sector reallocation from agricultural to non-agricultural work occurred, (2) urban women enjoy increased employment, and (3) both rural and urban women show increased secondary or higher education enrolment. To test these hypotheses, DHS data from the 2018 wave was analysed using entropy balancing. Findings show weak support for FLFP reallocation from agricultural to non-agricultural work among rural women over 30 years old, but rural women under 30 showed decreased overall and agricultural employment whereas non-agricultural employment remained unchanged. On the contrary, husbands of the rural under-30 group did experience labour force reallocation and growing overall employment rates. These results are in line with Goldin's (1995) theory of a U-shaped FLFP function. This theory states that in informal economies that industrialise, female labour decreases due to the household's income effect when husbands find stable work and due to stigmas against female labour in the growing manufacturing sector. This thesis shows that when the new infrastructure allows husbands to replace agricultural for non-agricultural work, young rural women stop working, but do not reallocate their time to schooling or bearing more children. Paid work and non-agricultural work for these women remain unchanged. These results suggest that treated young rural women abandon home-based agricultural or other informal work when their husband's new work and income allow them to.

Furthermore, in urban areas increased mobility and connectivity did not lead to increased female employment. Reasons for this could be found in the tight and saturated labour markets throughout urban Nigeria. However, the results of increased employment amongst urban husbands (and younger rural husbands) indicate that male employment is more sensitive to infrastructure improvement and that the distribution of male and female work within households is still bounded by traditional gender roles. If new labour opportunities arise, husbands might traditionally be more likely to take them than wives. Interestingly, married urban women are more enrolled in secondary or higher education and less fertile, which is, again, in line with the theory of a U-shaped FLFP function (Goldin, 1995): in industrialised economies that modernise, female labour increases after secondary education increases. However, considering Nigeria's growing unemployment rates amongst graduates and others entering the labour force, the prediction of growing female employment on the right tail of the U-shaped FLFP function is not guaranteed in today's urban Nigeria.

In conclusion, the new Abuja-Kaduna railway line did not cause evidently positive female employment or education effects - with a few possible exceptions. With overflowing labour markets across

the country, infrastructure enhancement alone does not lead to women finding non-existent job opportunities. However, it did result in male labour sector reallocation and increased overall male employment. It is therefore plausible that the railway segment actually has been an impactful intervention for labour allocation in the considered areas, as claimed by Chen (2018) and Nigeria's infrastructure minister (BBC News, 2017). It remains questionable how the effects relate to women's empowerment when only men benefit directly from this project. On the one hand, it could have created even wider opportunity and empowerment gaps between men and women. On the other hand, the improved infrastructure could be a component of general economic growth which is not proportionally related to female labour participation, but develops indeed in the shape of a U. In that case, the different employment effects for men and women could be seen as part of a process with a longer-term outcome of optimised female employment in the areas. A process that current generations of women, however, will not benefit from.

References

Abdulsalam, A., Xu, H., Ameer, W., Abdo, A. B., & Xia, J. (2021). Exploration of the Impact of China's Outward Foreign Direct Investment (FDI) on Economic Growth in Asia and North Africa along the Belt and Road (B&R) Initiative. *Sustainability, 13*(4), 1623.

Adepoju, P. (2021, January 19). Nigeria's China-built railway has to avoid debt pitfalls. *Quartz*. Retrieved from https://qz.com/africa/1958964/nigerias-china-built-railway-has-to-avoid-debt-pitfalls/ on 26-07-2021

Adukia, A., Asher, S. & Novosad, P. (2020). Educational Investment Response to Economic Opportunity: Evidence from Indian Road Construction. *American Economic Journal: Applied Economics 12*(1): 348-376.

Africa-China Reporting Project. (2018, January 9). Chinese-built Abuja-Kaduna railway: Opening fresh vistas in Nigeria's new economy. Retrieved from https://africachinareporting.co.za/2018/01/chinese-built-abuja-kaduna-railway-opening-fresh-vistas-in-nigerias-new-economy/ on 10-8-2021

Aggarwal, S. (2018). Do rural roads create pathways out of poverty? Evidence from India. *Journal of Development Economics, 133*, 375-395.

Akinwotu, E. (2021, June 14). Young, qualified and barely scraping by – inside Nigeria's economic crisis. *The Guardian*. Retrieved from https://www.theguardian.com/world/2021/jun/14/young-qualified-and-barely-scraping-by-inside-nigerias-economic-crisis on 2-11-2021

Allotey, D., Flax, V. L., Ipadeola, A., Kwasu, S., Bentley, M. E., Worku, B., ... & Martin, S. L. (2022). Maternal and paternal involvement in complementary feeding in Kaduna State, Nigeria: The continuum of gender roles in urban and rural settings. *Maternal & Child Nutrition*, e13325.

Anderson, S., & Eswaran, M. (2009). What determines female autonomy? Evidence from Bangladesh. *Journal of Development Economics, 90*(2), 179-191.

Anugwom, E. E. (2009). Women, education and work in Nigeria. Educational Research and Reviews, 4(4), 127-134.

Anyanwu, S. O., Adesanya, B. M., Adediji, A. M., & Adesanya, A. E. (2021). Female Labour Force Participation and Economic Growth Nexus: Evidence from Nigerian Economy.

Asher, S., & Novosad, P. (2016). Market access and structural transformation: Evidence from rural roads in India. *Manuscript: Department of Economics,* University of Oxford.

Asher, S., & Novosad, P. (2020). Rural roads and local economic development. *American economic review, 110*(3), 797-823.

Axinn, W. G., & Barber, J. S. (2001). Mass education and fertility transition. *American Sociological Review*, 481-505.

Balding, C. (2018). Why democracies are turning against Belt and Road. *Foreign Affairs, 24*.

Banerjee, A., Duflo, E., & Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics, 145,* 102442.

BBC News. (2017, December 7). Getting Nigeria's railways back on track with China's help. Retrieved from https://www.bbc.com/news/world-africa-42172955 on 26-07-2021

BBC News Pidgin. (2020, December 24). Christmas: Kaduna to Abuja road na di most dangerous for Nigeria? Retrieved from https://www.bbc.com/pidgin/tori-55434555 on 26-07-2021

Bose, C. E., Feldberg, R., & Sokoloff, N. J. (Eds.). (1987). *Hidden aspects of women's work*. New York: Praeger.

Brautigam, D. (2020). A critical look at Chinese 'debt-trap diplomacy': The rise of a meme. *Area Development and Policy, 5*(1), 1-14.

Brinton, M. C., & Lee, D. J. (2016). Gender-role ideology, labor market institutions, and post-industrial fertility. Population and Development Review, 405-433.

Brinton, M. C., Lee,Y. and Parish, W.L. (1995). "Married Women's Employment in Rapidly Industrializing Societies: Examples from East Asia." American Journal ofSociology 100(5): 1099–130

Bryan, G., Chowdhury, S., & Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. Econometrica, 82(5), 1671-1748.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature reviews neuroscience*, *14*(5), 365-376.

Carmody, P. (2020). Dependence not debt-trap diplomacy. *Area Development and Policy, 5*(1), 23-31.

Chen, Y. (2018). China's Role in Nigerian Railway Development and Implications for Security and Development. *United States Institute of Peace*.

Clemente, J. (2019, October 18). China Is The World's Largest Oil & Gas Importer. Forbes. Retrieved from https://www.forbes.com/sites/judeclemente/2019/10/17/china-is-the-worlds-largest-oil--gas-importer/ on 6-7-2021

Coulibaly, B. S., Gandhi, D., & Senbet, L. W. (2019). Is sub-Saharan Africa facing another systemic sovereign debt crisis? *Brookings India*

Davidson, H. (2018, May 15). Warning sounded over China's "debtbook diplomacy." *The Guardian*. Retrieved from https://www.theguardian.com/world/2018/may/15/warning-sounded-over-chinas-debtbook-diplomacy on 19-10-2021

Davies, K. (2009) "While global FDI falls, China's outward FDI doubles." *Transnational Corporations Review 1*(4)*:* 20-23.

Diaz, R., Behr, J. G., & Ng, M. (2016). Quantifying the economic and demographic impact of transportation infrastructure investments: A simulation study. *Simulation*, *92*(4), 377-393.

D.M.O.N. (2021, July 19). FACTS ABOUT CHINESE LOANS TO NIGERIA - Debt Management Office Nigeria. Dmo.Gov.Ng. Retrieved from https://www.dmo.gov.ng/facts-about-chinese-loans-to-nigeria on 26-07-2021

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review, 108*(4-5), 899-934.

Easterlin, R. A., & Crimmins, E. M. (1985). *The fertility revolution: A supply-demand analysis*. University of Chicago Press.

Economist, the. (2018, September 6). The perils of China's "debt-trap diplomacy." *The Economist*. Retrieved from https://www.economist.com/asia/2018/09/06/the-perils-of-chinas-debt-trap-diplomacy on 19-10-2021

Envuladu, E. A., Umaru, R. J., Iorapuu, N. O., Osagie, I. A., Okoh, E. O., & Zoakah, A. I. (2016). Determinants and effect of girl child marriage: a cross sectional study of school girls in Plateau State, Nigeria. *International Journal of Medicine and Biomedical Research*, *5*(3), 122-129.

Faria, J. (2021a, June 1). Proved crude oil reserves in Africa in 2021, by country. *Statista*. Retrieved from https://www.statista.com/statistics/1178147/crude-oil-reserves-in-africa-by-country/ on 6-7-2021

Faria, J. (2021b, February 22). Export volume of crude oil from Africa in 2019, by country. *Statista*. Retrieved from https://www.statista.com/statistics/1197901/export-volume-of-crude-oil-by-country-in-africa/#statisticContainer on 6-7-201

Fernholz, T. (2018, June 28). China Debt Trap: These eight countries are in danger of debt overloads from China's Belt and Road plans. *Quartz*. Retrieved from https://qz.com/1223768/china-debt-trap-these-eight-countries-are-in-danger-of-debt-overloads-from-chinas-belt-and-road-plans/ on 12-7-2021

Fogel, R. W. (1963). *Railroads and American economic growth: essays in econometric history* (Doctoral dissertation, Johns Hopkins University).

Gaddis, I., & Klasen, S. (2012). Economic Development, Structural Change and Women's Labor Force Participation A Reexamination of the Feminization U Hypothesis (No. 71). *Courant Research Centre PEG*.

Girma, J., & Gardebroek, C. (2015). The impact of contracts on organic honey producers' incomes in southwestern Ethiopia. *Forest Policy and Economics*, *50*, 259-268.

Global Construction Review. (2021, January 11). A year on, Nigeria still waiting for $5.3bn Chinese loan for railway to Kano. *GlobalConstructionReview.Com*. Retrieved from https://www.globalconstructionreview.com/nigeria-has-been-waiting-year-53bn-chinese-loan-ra/ on 9-9-2021

Goldin, C. (1995). The U-Shaped Female Labor Force Function in Economic Development and Economic History. *Schultz TP Investment in Women's Human Capital and Economic Development. University of Chicago Press* pp. 61-90.

Gollin, D., Lagakos, D., & Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics, 129*(2), 939-993.

Gollin, D., & Rogerson, R. (2014). Productivity, transport costs and subsistence agriculture. *Journal of Development Economics*, 107, 38-48.

Graetz, N., Friedman, J., Osgood-Zimmerman, A., Burstein, R., Biehl, M. H., Shields, C., ... & Hay, S. I. (2018). Mapping local variation in educational attainment across Africa. *Nature*, *555*(7694), 48-53.

Hainmüller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, *20*(1), 25-46.

Hillman, J. E. (2019, January 18). Corruption Flows Along China's Belt and Road. *Center for Strategic and International Studies*. Retrieved from https://www.csis.org/analysis/corruption-flows-along-chinas-belt-and-road on 6-7-2021

Ibrahim, Z. L., Khan, A., & bin Ramli, J. (2020). Cultural and Socio-economic Status Factors Affecting Female Education in Sokoto State, Northern Nigeria: Implication for Counselling. *Universal Journal of Educational Research, 8*(11C), 124-128.

Idowu, O. O., & Owoeye, T. (2019). Female labour force participation in African countries: An empirical analysis. *Indian Journal of Human Development*, *13*(3), 278-293.

IMF (2021). Six Charts Show the Challenges Faced by Sub-Saharan Africa. (2021, April 12). IMF. Retrieved from https://www.imf.org/en/News/Articles/2021/04/12/na041521-six-charts-show-the-challenges-faced-by-sub-saharan-africa on 22-7-2021

Ismail, M., Muhammed, Z. D., Farouk, M., Usman, K., Tanko, A. M., & Adamu, G. (2014). Urban Growth and Housing Problems in Kubwa District of Bwari Area Council, Abuja, Nigeria. *Glob. J. Res. Rev., 1,* 1-14.

Iweagu, H., Yuni, D. N., Chukwudi, N., & Andenyangtso, B. (2015). Determinants of female labour force participation in Nigeria: The rural/urban dichotomy. *Journal of Economics and Sustainable Development, 6*(10), 212-219.

Kapsos, S., Bourmpoula, E., & Silberman, A. (2014). Why is female labour force participation declining so sharply in India? (No. 994949190702676). *International Labour Organization.*

Keller, W., & Shiue, C. H. (2008). Tariffs, trains, and trade: The role of institutions versus technology in the expansion of markets.

Koolwal, G., & Van de Walle, D. (2013). Access to water, women's work, and child outcomes. *Economic Development and Cultural Change, 61*(2), 369-405.

Kyari, G. V., & Ayodele, J. (2014). The socio-economic effect of early marriage in North Western Nigeria. *Mediterranean Journal of Social Sciences*, *5*(14), 582-582.

Lei, L., Desai, S., & Vanneman, R. (2019). The impact of transportation infrastructure on women's employment in India. *Feminist economics, 25*(4), 94-125.

Lewis, W. A. (1954), Economic Development with Unlimited Supplies of Labour. *The Manchester School, 22*(2), 139–191

Li, Z., Ito, K., & Komiyama, R. (2005). Energy Demand and Supply Outlook in China for 2030 and A Northeast Asian Energy Community—The automobile strategy and nuclear power strategy of China. *Japan: The Institute of Energy Economics,* 2005 August, 31.

Liu, D. H., & Raftery, A. E. (2020). How do education and family planning accelerate fertility decline?. *Population and development review*, *46*(3), 409-441.

Lloyd, C. B., & Mensch, B. S. (2008). Marriage and childbirth as factors in dropping out from school: an analysis of DHS data from sub-Saharan Africa. *Population studies*, *62*(1), 1-13.

McMillan, M. S., & Rodrik, D. (2011). Globalization, structural change and productivity growth (No. w17143). *National Bureau of Economic Research*.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics, 90*(4), 683-701.

Mohamed, H. (2018, April 11). Nigeria: Clearing the locals to make Abuja the capital. Poverty and Development | Al Jazeera. Retrieved from https://www.aljazeera.com/features/2018/4/11/nigeria-clearing-the-locals-to-make-abuja-the-capital on 17-8-2021

Mohmand, Y. T., Wang, A., & Saeed, A. (2017). The impact of transportation infrastructure on economic growth: empirical evidence from Pakistan. *Transportation Letters*, *9*(2), 63-69.

Monks, K. C. (2017, September 15). China to build $5.8 billion hydropower plant in Nigeria. CNN. Retrieved from https://edition.cnn.com/2017/09/14/africa/nigeria-china-hydropower/index.html on 6-7-2021

Munshi, N. (2021, April 5). Nigeria's graduates live 'hand to mouth' as jobs crisis worsens. Financial Times. https://www.ft.com/content/1fb8a9d8-6891-48ac-a79a-59318e87e828

National Bureau of Statistics. (2018). Labor Force Statistics - Volume I: Unemployment and Underemployment report. Nigeria: Proshare

National Bureau of Statistics. (2020, August). Labor Force Statistics: Unemployment and Underemployment Report.

National Population Commission (NPC) [Nigeria] and ICF (2019). Nigeria Demographic and Health Survey 2018. Abuja, Nigeria, and Rockville, Maryland, USA: NPC and ICF.

Ng, C. P., Law, T. H., Jakarni, F. M., & Kulanthayan, S. (2019, April). Road infrastructure development and economic growth. In *IOP Conference Series: Materials Science and Engineering* (Vol. 512, No. 1, p. 012045). IOP Publishing.

Nnodim, O. (2019, October 29). Insecurity: Abuja-Kaduna railway passengers rise by 270%. Punch Newspapers. Retrieved from https://punchng.com/insecurity-abuja-kaduna-railway-passengers-rise-by-270/ on 26-07-2021

OECD. (2021), FDI flows (indicator). doi: 10.1787/99f6e393-en Retrieved on 7-6-2021

Okafor, G. O. (2010). Factors Affecting Women Education in Maiduguri, Borno State. *Knowledge Review, 21*(4), 61-68.

Okoye, D., Pongou, R., & Yokossi, T. (2019). New technology, better economy? The heterogeneous impact of colonial railroads in Nigeria. *Journal of Development Economics*, *140*, 320-354.

Olurounbi, R. (2021, April 28). *Nigeria is spending billions on its railways. Will it work?* The Africa Report.Com. Retrieved from https://www.theafricareport.com/83687/nigeria-increased-railway-lines-will-decrease-road-congestion-and-help-trade/ on 19-11-2021

Osili, U. O., & Long, B. T. (2008). Does female schooling reduce fertility? Evidence from Nigeria. *Journal of Development Economics*, *87*(1), 57-75.

Oxford Business Group. (2020, February 19). Rapid urbanisation in Abuja, Nigeria, drives infrastructure and real estate investment. Retrieved from https://oxfordbusinessgroup.com/overview/keys-city-rapid-urbanisation-one-continent%E2%80%99s-most-prosperous-cities-drives-investment-infrastructure on 17-8-2021

Pradhan, R. P., & Bagchi, T. P. (2013). Effect of transportation infrastructure on economic growth in India: The VECM approach. *Research in Transportation economics*, *38*(1), 139-148.

Railway Technology. (2020, June 15). Abuja-Kaduna Rail Line, Nigeria. Retrieved from https://www.railway-technology.com/projects/abuja-kaduna-rail-line/ on 26-07-2021

Sahoo, P., Dash, R. K., & Nataraj, G. (2010). Infrastructure development and economic growth in China. *Institute of Developing Economies Discussion Paper*, *261*.

Singh, A. (2020). The myth of 'debt-trap diplomacy' and realities of Chinese development finance. *Third World Quarterly, 42*(2), 239-253.

Shorter, E. (1973). Female emancipation, birth control, and fertility in European history. *The American Historical Review*, *78*(3), 605-640.

Sparrow, R., Suryahadi, A., & Widayanti, W. (2008). Public health insurance for the poor: targeting and impact of Indonesia's Asheskin Programme. *The Hague, Netherlands: Institute of Social Studies*.

Spierings, N. (2014). "The Influence of Patriarchal Norms, Institutions, and Household Composition on Women's Employment in Twenty-Eight Muslim-Majority Countries." *Feminist Economics 20*(4): 87–112.

Statista. (2021, March 19). Unemployment rate in Nigeria 2020, by state. Retrieved from https://www.statista.com/statistics/1119533/unemployment-rate-in-nigeria-by-state/ on 26-8-2021

Strohecker, K. J. B. (2020, November 20). As new debt crisis looms, Africa needs more than world is offering. U.S. Retrieved from https://www.reuters.com/article/africa-debt-idUKL1N2I51PX on 22-7-2021

Sylvester, I. T. (2014). A post resettlement appraisal of the socio-economic condition of Gbagi people in Kubwa, Federal Capital Territory (FCT) Abuja, Nigeria. *Academic research International, 5*(4), 153.

The China Africa Project. (2021, March 1). *New Railway in Nigeria to Push Up the Country's Debts to China*. Retrieved from https://chinaafricaproject.com/2021/03/01/new-railway-in-nigeria-to-push-up-the-countrys-debts-to-china/ on 4-4-2021

This Day Live (2020). Kaduna-Abuja highway: The road to death. Retrieved from https://www.thisdaylive.com/index.php/2020/11/21/kaduna-abuja-highway-the-road-to-death/ on 26-7-2021

Transparency International. (2019, November 19). Corruption Perceptions Index 2017 - News. Transparency.Org. Retrieved from https://www.transparency.org/en/news/corruption-perceptions-index-2017 on 22-7-2021

Tvaronavičienė, M. (2019). Insights into global trends of capital flows' peculiarities: emerging leadership of China. *Administratie si Management Public,* (32), 6-17.

U.S. Energy Information Administration (2018, December 31). China surpassed the United States as the world's largest crude oil importer in 2017. Retrieved from https://www.eia.gov/todayinenergy/detail.php?id=37821 on 6-7-2021

U.S. Energy Information Administration (n.d.). U.S. imports by country of origin. Retrieved from https://www.eia.gov/dnav/pet/pet_move_impcus_a2_nus_epc0_im0_mbblpd_a.htm on 6-7-2021

Van den Broeck, G., & Maertens, M. (2017). Does off-farm wage employment make women in rural Senegal happy? *Feminist Economics, 23*(4), 250-275.

World Bank. (1994) *World Development Report:Providing Infrastructure for Development;* Oxford University Press*:* New York, NY, USA, pp. 71–199.

World Bank. (2021a, January 29). Employment in agriculture (% of total employment) (modeled ILO estimate) | Data. Data.Worldbank.Org. Retrieved from https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS on 2-9-2021

World Bank. (2021b, June 15). Labor force participation rate, female (% of female population ages 15+) (modeled ILO estimate) - Nigeria | Data. Data.Worldbank.Org. Retrieved from https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS?end=2019&locations=NG&start=1990 on 6-9-2021

Were, A. (2018). Debt trap? Chinese loans and Africa's development options. *Policy insights, 66*

Xu, Z., & Sun, T. (2021). The Siphon effects of transportation infrastructure on internal migration: evidence from China's HSR network. *Applied Economics Letters*, *28*(13), 1066-1070.

Yang, Y. (2006). China's Integration into the World Economy: implications for developing countries. *Asian-Pacific Economic Literature, 20*(1), 40-56.

Zou, W., Zhang, F., Zhuang, Z., & Song, H. (2008). Transport Infrastructure, Growth, and Poverty Alleviation: Empirical Analysis of China. *Annals of Economics & Finance*, *9*(2).

Appendix

Appendix table 1. Statistical differences of covariates between rural control and treatment groups pre and post reweighting

	Younger than 30					30 and older				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variable	Control	Treatment	Difference before reweighting	Reweighted control	Difference after reweighting	Control	Treatment	Difference before reweighting	Reweighted control	Difference after reweighting
Age	21.473	21.602	0.130	21.602	-0.000	38.058	37.164	-0.894	37.164	-0.000
	(4.170)	(4.449)	(0.364)	(4.253)	(0.293)	(5.839)	(5.736)	(0.587)	(5.715)	(0.413)
Primary education	0.729	0.717	-0.012	0.717	-0.000	0.554	0.491	-0.062	0.491	-0.000
	(0.445)	(0.452)	(0.038)	(0.451)	(0.030)	(0.497)	(0.502)	(0.050)	(0.500)	(0.036)
Number of births before 5 years ago	0.557	0.687	0.130	0.687	-0.000	4.302	4.431	0.129	4.431	-0.000
	(0.990)	(1.100)	(0.087)	(1.145)	(0.076)	(2.441)	(2.243)	(0.243)	(2.476)	(0.170)
Islamic	0.478	0.566	0.088**	0.566	0.000	0.457	0.440	-0.017	0.440	-0.000
	(0.500)	(0.497)	(0.043)	(0.496)	(0.033)	(0.499)	(0.498)	(0.050)	(0.497)	(0.036)
Wealth index	-26.951	-6.690	20.261***	-6.692	0.002	-24.870	-12.573	12.296	-12.580	0.007
	(78.406)	(64.080)	(6.542)	(79.950)	(4.885)	(87.696)	(66.437)	(8.550)	(89.759)	(5.695)
Migrated from other state over 4 years ago	0.126	0.084	-0.042	0.084	-0.000	0.308	0.216	-0.093**	0.216	-0.000
	(0.332)	(0.279)	(0.028)	(0.278)	(0.019)	(0.462)	(0.413)	(0.046)	(0.412)	(0.030)
Cluster's travel time to populous place in hours	0.776	0.588	-0.188***	0.588	-0.000	0.805	0.558	-0.247***	0.558	-0.000
	(0.629)	(0.335)	(0.050)	(0.436)	(0.026)	(0.670)	(0.328)	(0.064)	(0.378)	(0.026)
Log of cluster's gross cell production	7.400	7.403	0.003	7.403	0.000	7.403	7.401	-0.002	7.401	-0.000
	(0.030)	(0.013)	(0.002)	(0.032)	(0.002)	(0.031)	(0.013)	(0.003)	(0.030)	(0.002)
Cluster's population growth between 2010 and 2015	0.255	0.296	0.041***	0.296	0.000	0.252	0.325	0.073***	0.325	0.000
	(0.170)	(0.158)	(0.014)	(0.181)	(0.011)	(0.169)	(0.165)	(0.017)	(0.183)	(0.013)
Observations	713	166	879	713	879	652	116	768	652	768

Appendix table 2. Statistical differences of covariates between urban control and treatment groups pre and post reweighting

	Younger than 30					30 and over				
	(1)	(2)	(3)	(1)	(3)	(1)	(2)	(3)	(1)	(3)
Variable	Control	Treatment	Difference before reweighting	Reweighted control	Difference after reweighting	Control	Treatment	Difference before reweighting	Reweighted control	Difference after reweighting
Age	21.419	21.219	-0.200	21.219	0.000	37.367	38.651	1.285***	38.650	0.002
	(4.392)	(4.420)	(0.268)	(4.356)	(0.201)	(5.490)	(5.870)	(0.362)	(5.777)	(0.287)
Primary education	0.895	0.923	0.028	0.923	0.000	0.738	0.838	0.100***	0.838	0.000
	(0.307)	(0.267)	(0.018)	(0.267)	(0.012)	(0.440)	(0.369)	(0.028)	(0.369)	(0.018)
Number of births before 5 years ago	0.327	0.324	-0.003	0.324	0.000	3.861	3.475	-0.386**	3.476	-0.001
	(0.795)	(0.808)	(0.049)	(0.794)	(0.037)	(2.873)	(2.666)	(0.185)	(2.638)	(0.131)
Islamic	0.759	0.645	-0.114***	0.645	-0.000	0.701	0.514	-0.187***	0.514	-0.000
	(0.428)	(0.479)	(0.027)	(0.479)	(0.022)	(0.458)	(0.501)	(0.030)	(0.500)	(0.025)
Wealth index	100.984	117.945	16.962***	117.945	0.000	104.833	120.215	15.382***	120.167	0.048
	(65.434)	(70.973)	(4.050)	(64.250)	(3.102)	(70.790)	(80.349)	(4.731)	(65.643)	(3.613)
Migrated from other state over 4 years ago	0.110	0.157	0.048**	0.157	0.000	0.205	0.359	0.154***	0.359	0.000
	(0.312)	(0.365)	(0.020)	(0.364)	(0.017)	(0.404)	(0.481)	(0.027)	(0.480)	(0.024)
Cluster's travel time to populous place in hours'	0.085	0.055	-0.030	0.055	-0.000	0.094	0.056	-0.038	0.057	-0.001
	(0.598)	(0.137)	(0.033)	(0.378)	(0.013)	(0.671)	(0.158)	(0.040)	(0.416)	(0.016)
Log of cluster's gross cell production	7.422	7.410	-0.012	7.410	-0.000	7.437	7.409	-0.028***	7.409	-0.000
	(0.148)	(0.016)	(0.008)	(0.099)	(0.003)	(0.169)	(0.016)	(0.010)	(0.097)	(0.003)
Cluster's population growth between 2010 and 2015	0.187	0.311	0.123***	0.311	0.000	0.182	0.313	0.131***	0.313	0.000
	(0.123)	(0.194)	(0.008)	(0.190)	(0.009)	(0.117)	(0.195)	(0.009)	(0.191)	(0.010)
Observations	1,579	324	1,903	1,579	1,903	1,363	284	1,647	1,363	1,647

Appendix table 3a. Employment and education effects of rural women under 30 – not migrated

	(1) Higher education enrolment	(2) Employed	(3) No education enrolment and not employed
Treatment	0.0134	-0.101*	0.0673
	(0.720)	(0.029)	(0.137)
_cons	0.192***	0.532***	0.317***
	(0.000)	(0.000)	(0.000)
N	785	785	785

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 3b. Employment quality of rural women under 30 - not migrated

	(1) Works in agri-sector	(2) Works in non-agri-sector	(3) Has paid employment
Treatment	-0.129***	0.0281	0.0166
	(0.000)	(0.524)	(0.712)
_cons	0.216***	0.316***	0.348***
	(0.000)	(0.000)	(0.000)
N	785	785	785

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 4. Employment effects of rural women of 30 years and over - not migrated

	(1) Employed	(2) Works in agri-sector	(3) Works in non-agri-sector	(4) Has paid employment
Treatment	0.00746	-0.0723	0.0798	0.0443
	(0.870)	(0.128)	(0.126)	(0.389)
_cons	0.747***	0.344***	0.403***	0.570***
	(0.000)	(0.000)	(0.000)	(0.000)
N	733	733	733	733

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 5. Employment and education effects of urban women under 30 - not migrated

	(1) Higher education enrolment	(2) Employed	(3) Has paid employment	(4) No education enrolment and not employed
Treatment	0.0208	-0.0153	0.0140	0.0115
	(0.549)	(0.660)	(0.673)	(0.729)
_cons	0.363***	0.406***	0.322***	0.302***
	(0.000)	(0.000)	(0.000)	(0.000)
N	1717	1717	1717	1717

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 6. Employment effects of urban women of 30 years and over - not migrated

	(1) Employed	(2) Has paid employment
Treatment	0.0173	0.0626*
	(0.552)	(0.042)
_cons	0.812***	0.751***
	(0.000)	(0.000)
N	1563	1563

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 7. Education enrolment and employment effects of married rural women under 30

	(1) Enrolled in higher education	(2) Has employment	(3) Works in agri-sector	(4) Works in non-agri-sector	(5) Gets paid for employment	(6) Husband works
Treatment	0.00772	-0.0946	-0.147***	0.0520	0.00555	0.0459*
	(0.420)	(0.090)	(0.000)	(0.344)	(0.921)	(0.022)
_cons	0.00171	0.623***	0.250***	0.373***	0.457***	0.935***
	(0.319)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	535	535	535	535	535	520

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 8. Education enrolment and employment effects of married urban women under 30

	(1) Enrolled in higher education	(2) Has employment	(3) Gets paid for employment	(4) Works in non-agri-sector	(5) Husband works
Treatment	0.0541	0.0322	0.0399	0.0404	0.0424*
	(0.054)	(0.563)	(0.472)	(0.468)	(0.021)
_cons	0.0336***	0.494***	0.443***	0.486***	0.948***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	726	726	726	726	687

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 9. Education enrolment and employment effects of unmarried rural women under 30

	(1) Enrolled in higher education	(2) Has employment	(3) Works in agri-sector	(4) Works in non-agri-sector	(5) Gets paid for employment
Treatment	0.0457	-0.146*	-0.104***	-0.0428	0.000897
	(0.529)	(0.020)	(0.001)	(0.464)	(0.987)
_cons	0.488***	0.380***	0.137***	0.243***	0.182***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	344	344	344	344	344

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 10. Education enrolment and employment effects of unmarried urban women under 30

	(1) Enrolled in higher education	(2) Has employment	(3) Gets paid for employment	(4) Works in non-agri-sector
Treatment	-0.0147	-0.0196	0.0115	-0.0179
	(0.728)	(0.622)	(0.756)	(0.652)
_cons	0.515***	0.358***	0.269***	0.356***
	(0.000)	(0.000)	(0.000)	(0.000)
N	1177	1177	1177	1177

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix table 11. Balance table treatment and reweighted control groups of only married women

	Rural under 30			Rural 30 and older			Urban under 30			Urban 30 and older		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Variable	Reweighted control	Treatment	Difference after reweighting	Reweighted control	Treatment	Difference after reweighting	Reweighted control	Treatment	Difference after reweighting	Reweighted control	Treatment	Difference after reweighting
Age	23.800	23.943	0.144	37.008	37.100	0.092	24.730	23.947	-0.783***	38.433	38.773	0.340
	(3.368)	(3.403)	(0.293)	(5.630)	(5.645)	(0.424)	(3.103)	(3.741)	(0.254)	(5.727)	(5.773)	(0.306)
Primary education	0.600	0.604	0.004	0.476	0.464	-0.012	0.862	0.825	-0.038	0.826	0.819	-0.007
	(0.491)	(0.491)	(0.042)	(0.500)	(0.501)	(0.038)	(0.345)	(0.382)	(0.027)	(0.379)	(0.386)	(0.020)
Number of births in previous 5 years	1.047	1.075	0.028	4.463	4.455	-0.009	0.843	0.833	-0.010	3.574	3.731	0.157
	(1.282)	(1.217)	(0.108)	(2.460)	(2.208)	(0.176)	(1.116)	(1.174)	(0.085)	(2.630)	(2.682)	(0.141)
Islamic	0.619	0.651	0.032	0.459	0.445	-0.014	0.706	0.789	0.084***	0.538	0.559	0.021
	(0.486)	(0.479)	(0.042)	(0.499)	(0.499)	(0.038)	(0.456)	(0.409)	(0.032)	(0.499)	(0.498)	(0.027)
Wealth index	-12.827	-9.911	2.916	-12.975	-15.401	-2.426	104.937	98.690	-6.248	121.263	121.905	0.642
	(79.280)	(63.745)	(6.201)	(90.378)	(66.150)	(5.956)	(63.040)	(77.572)	(5.231)	(64.990)	(81.762)	(3.930)
Migrated from other state over 4 years ago	0.089	0.085	-0.004	0.222	0.218	-0.004	0.214	0.167	-0.048	0.353	0.387	0.034
	(0.284)	(0.280)	(0.024)	(0.416)	(0.415)	(0.031)	(0.411)	(0.374)	(0.029)	(0.478)	(0.488)	(0.026)
Cluster's travel time to populous place in hours'	0.616	0.595	-0.021	0.556	0.570	0.014	0.059	0.087	0.028	0.064	0.063	-0.001
	(0.462)	(0.323)	(0.034)	(0.372)	(0.333)	(0.027)	(0.368)	(0.183)	(0.022)	(0.452)	(0.171)	(0.018)
Log of cluster's gross cell production	7.402	7.405	0.004*	7.401	7.401	0.001	7.405	7.411	0.006	7.410	7.410	0.000
	(0.033)	(0.015)	(0.002)	(0.030)	(0.013)	(0.002)	(0.106)	(0.016)	(0.006)	(0.102)	(0.016)	(0.004)
Cluster's population growth between 2010 and 2015	0.293	0.274	-0.020	0.325	0.317	-0.007	0.316	0.277	-0.038***	0.307	0.298	-0.008
	(0.186)	(0.153)	(0.015)	(0.184)	(0.166)	(0.013)	(0.190)	(0.188)	(0.014)	(0.190)	(0.193)	(0.010)
Observations	429	106	535	594	110	704	612	114	726	1,173	238	1,411

Appendix table 12. Education enrolment effects of married urban women under 30

	(1) Enrolled in higher education
Treatment	0.0401*
	(0.032)
_cons	0.0477***
	(0.000)
N	726

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$