Lighting the World: Is Electricity the Answer?

MUHAMMAD SAAD QURESHI [qureshim]

December 20th, 2021

ECN-372

Abstract

I investigate the impact of the World Bank-backed Rural Electrification Program on educational attainment and fertility outcomes in Ethiopia. Utilizing data from the Demographics and Health Surveys website, I employ a two-stage difference-in-differences model. The analysis reveals a negative overall impact of the electrification intervention on educational outcomes among treated households. Interestingly, the poorest households exhibit the least effect, in contrast to medium and high-income households. Additionally, I observe a positive treatment effect on fertility outcomes, indicating an increase in the number of children under age 5 per woman in treated households. These findings emphasize the need for further investigation and policy considerations to optimize the program's benefits and address unintended consequences.

1. Introduction

1.1 Context, Motivation, and Research Problem

Globally, 13 percent of the population lacked access to electricity in 2016, with only 10 percent of rural households in Sub-Saharan Africa having electrical access (Ritchie, 2020; World Bank, 2016). Given these statistics, it is not surprising that electrification has become a popular developmental intervention, with significant investments made worldwide. Electrification is viewed as a means of poverty alleviation and economic development, with billions of dollars spent annually on electrification programs (International Energy Agency, 2011). According to International Energy Agency, $9 billion were spent on electrification in 2009 and the figure is expected to go up to $14 billion per year by 2030 (International Energy Agency (2011)). In India, during President Modi's tenure, from 2014 to 2018, there was a more than double increase in government spending on rural electrification.[1] United Nations has also identified universal energy access as one of its main goals.

However, the evidence regarding the utility of electrification is not only scarce but also conflicting. Some studies suggest its positive impact, while others question its effectiveness for development. When evaluating the usefulness of electrification, educational outcomes are often utilized as measures of progress. There's evidence available that rural electrification may lead to better educational outcomes as students can benefit from increased study time and an improved study environment.[2] There are also reports of electrification leading to more media exposure because of access to mediums such as televisions, which can ultimately translate to increased health awareness and productivity (World Bank, 2016). On the other hand, there are concerns that electrical appliances may create distractions, hindering educational progress.

The divergent findings and theories surrounding the developmental benefits of electrification make it an intriguing research topic. This paper aims to assess the impact of the second phase of the Output-Based Aid component of the World Bank's Ethiopian Access Rural Expansion project, implemented from 2011 to 2013 in Ethiopia, on educational attainment and fertility outcomes for households in treated zones. Broadly, this research seeks to examine whether rural electrification improves educational attainment and influences fertility rates.

[1] https://economictimes.indiatimes.com/industry/energy/power/government-has-doubled-spend-on-ruralelectrification-in-4-years-rk-singh/articleshow/63992059.cms?from=mdr
[2] https://openknowledge.worldbank.org/handle/10986/6519

1.2 Literature Review and Contribution

The existing literature on electrification presents a mixed perspective, with varying conclusions regarding its advantages. For instance, Joshua Lewis and Edson Severnini conducted a study on the short- and long-term impacts of rural electrification, utilizing a panel dataset that combines county outcomes with data on power plants opened between 1930 and 1960 (Lewis et al., 2017). Employing a standard spatial equilibrium model as their conceptual framework, the authors employed a difference-in-differences empirical approach to evaluate the effects on farmland value, housing prices, and income proxies (Lewis et al., 2017). Their findings indicate that rural electrical access leads to a short-term increase in farmland and property values but does not significantly impact local incomes or nonagricultural sectors. However, in the long term, the program exhibits substantial benefits and stimulates economic activity over decades (Lewis et al., 2017).

Another relevant paper titled 'Out of the Darkness and Into the Light? Development Effects of Rural Electrification' explores the impact of rural electrification on development using a regression discontinuity specification (Burlig et al., 2016). The authors also employ high-resolution geospatial data to identify medium-term economic impacts of electrical access (Burlig et al., 2016). However, their study reveals that despite increased electricity usage, it does not translate into economic development (Burlig et al., 2016).

Similarly, Kenneth Lee et al. investigate the impact of rural electrification in Kenya through a randomized experiment that expands electric grid infrastructure (Lee et al., 2020). Their study concludes that rural electrification does not generate social surplus in Kenya when measured by standard criteria (Lee et al., 2020). Additionally, the cost of electrifying households exceeds their affordability, posing a significant challenge (Lee et al., 2020).

In this paper, I aim to contribute to the existing literature by focusing on the country of Ethiopia and examining the effects of electrification on education and fertility. The uniqueness of my research lies in the measurement of these variables. For educational attainment, I utilize the average household educational attainment score, which considers the education outcomes of all household members. Regarding fertility, I adopt a distinct approach by examining the number of children under the age of five per woman in a household. This strategy for tracking fertility has not been explored in previous studies. Furthermore, unlike most of the available literature that relies on panel datasets, I employ merged cross-sectional data, presenting an alternative approach to this issue. It is important to note that limited research exists assessing the impact of electrification on fertility, further distinguishing my study in this area.

1.3 Research Program Description and Context

The Federal Democratic Republic of Ethiopia comprises nine National Regional States, namely Afar, Amhara, Oromia, Somali, Tigray, Benishangul-Gumuz (BSG), Southern Nations, Nationalities, and Peoples Region (SNNPR), Gambella, and Harari, along with two administrative councils, Addis Ababa and Dire Dawa.[3] With a population of approximately 110.4 million, around 85 percent of Ethiopians reside in rural areas.[4] However, living conditions in these regions are generally subpar, lacking even basic amenities such as electricity. In fact, a mere 10 percent of areas in Ethiopia had access to electricity services as of 2014 (World Bank, 2016).

To address this pressing issue, the government of Ethiopia, in collaboration with external organizations like the World Bank, initiated programs aimed at improving electrical access. One notable program was launched in 2005 by the Ethiopian government, targeting the adoption of household electricity in rural towns and villages where services were already available. Subsequently, the World Bank supported this government program through the Electricity Access Rural Expansion Project (Phase 2) to enhance its sustainability. As a result of the World Bank's intervention, the Ethiopian government authorized customers to pay for connections over time, enabled by funding from the Global Partnership on Output-Based Aid grant.[4] This grant allowed the Ethiopian Electric Power Corporation to subsidize interest rates on loans for economically disadvantaged customers, thereby financing connection fees. [5]

A project completion survey report published by the World Bank in 2013 hailed the electrification program, particularly active from 2011 to 2013, as a success. According to the report, in the five study regions (Oromia, Tigray, BSG, Amhara, and SNNPR), the proportion of households benefiting from the Global Partnership on Output-Based Aid (GBOBA) and having electricity reached nearly 42 percent (World Bank, 2016). The GBOBA grant facilitated formal connections to grid-based electricity for 43,000 poor rural households, accounting for 75 percent of the country's total connections during the period from 2011 to 2013 (World Bank, 2016).

[3] Taken from the Ethiopian Embassy website. Accessed here: https://ethiopianembassy.org/overview-about-ethiopia/
[4] Taken from the Ethiopian Embassy website. Accessed here: https://ethiopianembassy.org/overview-about-ethiopia/
[4] https://openknowledge.worldbank.org/handle/10986/26317
[5] https://openknowledge.worldbank.org/handle/10986/26317

Furthermore, the survey report highlighted Oromia and Tigray as the regions with the highest share of GBOBA connections compared to the other three regions:

Table 1

Share of GPOBA Households among Households with Electricity in the Study Regions	
Region	Share of GBOBA connections (%)
Oromia	69.5
Amhara	9.3
SNNPR	36.9
Tigray	58.5
BSG	38.8
Total	41.8

Source: World Bank 2016

So, in the five regions where the program was most active and the survey was conducted, it is visible that the GBOBA connections are most concentrated in Oromia and Tigray.

1.4 Endogeneity Problem and Research Design

In this study, I am addressing the endogeneity problem associated with my research question, which focuses on the impact of an electrification program on educational attainment and fertility. One of the main challenges I face is disentangling the effects of electrification from other factors that could influence these outcomes, such as increased road access or household income. It is difficult to isolate the sole impact of electrification.

Specifically, I encounter endogeneity issues related to other potential influences on educational attainment and fertility outcomes in households located in the treated zones. For example, households in treated zones may have higher wealth levels, which could positively affect their educational attainment and fertility outcomes. Moreover, the presence of younger household members in treated zones might result in a lower overall level of education, as younger individuals might not have attained the same level of education as adults. Treated households may also have better access to urban areas, leading to improved educational opportunities. Additionally, there might be inherent differences between females living in intervention zones and those in non-intervention zones. For instance, females in intervention zones might be more involved in productive activities such as farming, thereby assigning a higher opportunity cost to having children.

These concerns are not unfounded, as at the end of the program, 90 percent of rural households still lacked electricity. Consequently, there may be fundamental differences between the regions selected for intervention and those not chosen. Thus, isolating the specific impact of providing electricity becomes challenging. To address these issues, I will employ a basic two-stage difference-in-differences model in this paper. I will examine parallel trends to ensure that pretreatment trends are similar for both the treated and control groups. Subsequently, I will analyze the treated effect in the post-intervention period once parallel trends have been established.

To mitigate endogeneity concerns, I will incorporate zonal fixed effects and include suitable control variables such as average household age, number of household members, and wealth index. The treated group consists of households located in zones where the electrification program occurred, while the control group comprises households in zones where the program did not take place. By employing these methods and controlling for relevant factors, I aim to address the challenges posed by endogeneity and obtain robust findings regarding the impact of electrification on educational attainment and fertility outcomes.

The basic regression specification that I will be using is

$$\boldsymbol{Y}_{xyz} = \boldsymbol{\beta_o} + \boldsymbol{\beta_1 Treated}_{xyz} + \boldsymbol{\beta_2 Post}_{xyz} + \boldsymbol{\beta Post}_{xyz} * \boldsymbol{Treated}_{xyz} + \boldsymbol{Controls}_{xyz}$$
$$+ \boldsymbol{ZFE} + \boldsymbol{\epsilon xyz}$$

Here, x = household in question y = year in question and z = zone in question and Z = zonal fixed effects

In the provided equation, Y represents the outcome variable, which is the focus of analysis. The variable x corresponds to a specific household, z indicates the zone in which the household is located, and y represents the year of observation. The inclusion of a treated dummy variable in the equation helps track whether a household is located in a treated zone or not. The term 'Post' indicates that the household belongs to the group interviewed in 2016, following the treatment years. The treatment effect is denoted by Beta 3, which quantifies the disparity in outcomes during the post period between individuals in the treated zone and those outside it.

This study considers three outcome variables of interest. In the initial stage, the analysis aims to determine if there is a measurable increase in electrification attributable to the program. The outcome variable 'has electricity' is employed to assess whether there is enhanced electrification in treated zones during the post period compared to control zones. Once it is established that there is additional electrification in treated zones, the study proceeds to the second stage, examining the impact of electrification on educational attainment and fertility behavior.

To identify the zones that received treatment through the program, survey reports from reputable sources such as the World Bank and African Development Bank are utilized. These organizations have published reports on the project, with the World Bank Group report utilizing data and maps to identify the exposed regions. Cross-sectional data for the pre and post periods is obtained from the Demographics and Health Surveys (DHS) website. Specifically, data for the years 2000, 2005, 2010, and 2016 are used. The survey datasets contain various variables such as household size, number of children under 5 years old, age, education level, and gender. Additionally, GPS datasets obtained from the DHS website allow for the identification of households within different zones in Ethiopia, enabling the formation of control and treatment groups.

The structure of the paper is as follows: Section 2 provides an examination of the study's theoretical framework. Section 3 describes the methodology employed and details the data used. Section 4 presents the main research findings, while Section 5 conducts robustness checks to ensure the reliability of the results. Finally, Section 6 concludes the study and offers a discussion of its implications.

2. Theoretical Framework

Economic theory suggests that electrification can impact educational attainment and fertility through different channels. There is also literature available that specifically looks at the impact of electrification on education and fertility outcome variables. The following chart incorporates these varying theories and concepts to provide a conceptual framework for this paper.

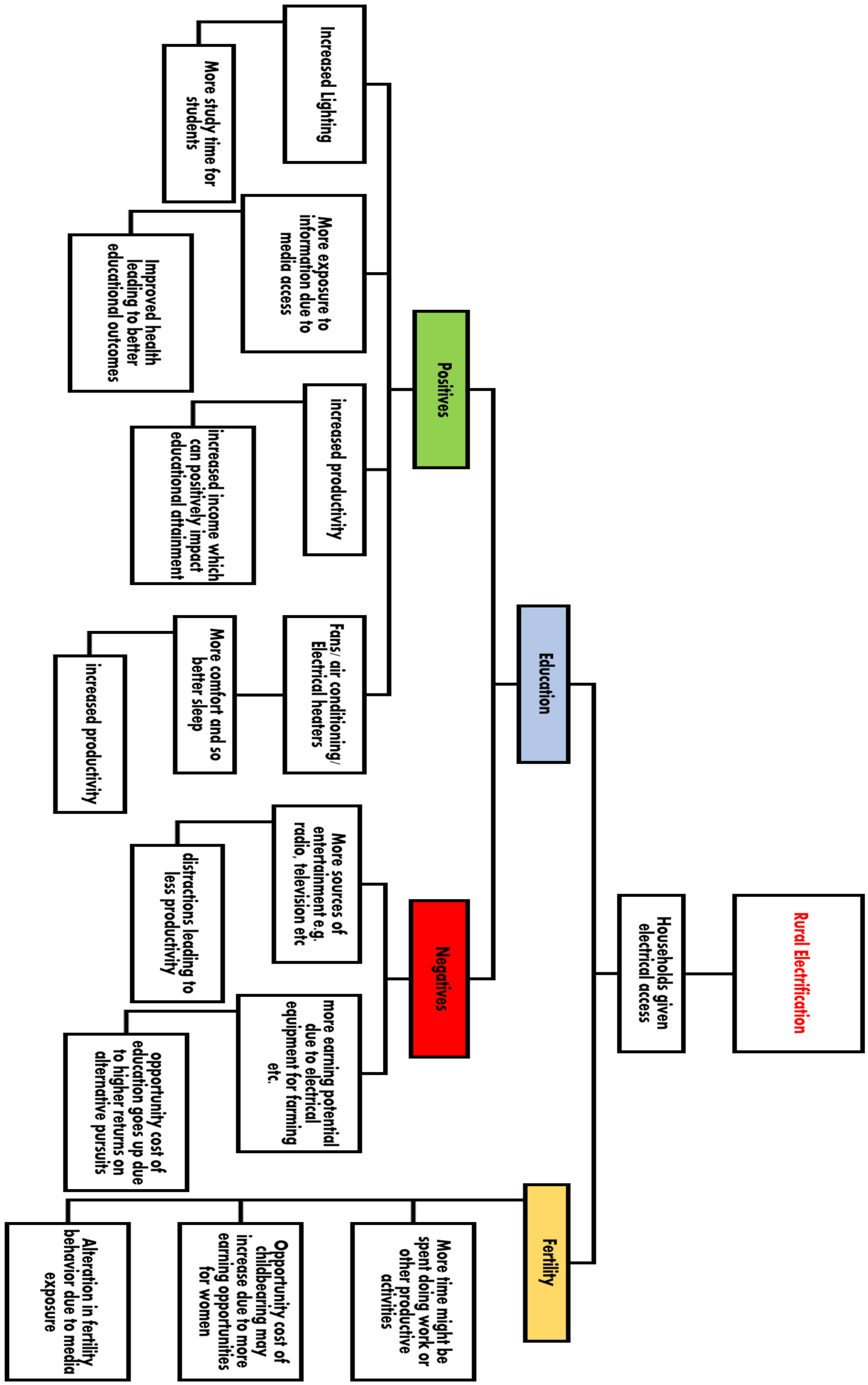
Rural Electrification
Households given electrical access
Education
Positives
Negatives
Fertility
Increased Lighting
More study time for students
More exposure to information due to media access
Improved health leading to better educational outcomes
increased productivity
increased income which can positively impact educational attainment
Fans/ air conditioning/ Electrical heaters
More comfort and so better sleep
increased productivity
More sources of entertainment e.g. radio, television etc
distractions leading to less productivity
more earning potential due to electrical equipment for farming etc.
opportunity cost of education goes up due to higher returns on alternative pursuits
More time might be spent doing work or other productive activities
Opportunity cost of childbearing may increase due to more earning opportunities for women
Alteration in fertility behavior due to media exposure

Figure 1

Theoretical considerations suggest that electricity can have both positive and negative impacts, as illustrated in Figure 1.[6] Access to electricity can bring about positive effects, such as increased lighting that allows for extended study time among students. Moreover, access to electrical appliances like televisions and radios can broaden people's exposure to information, leading to improved health outcomes by promoting awareness of nutrition and safe-sex practices. These factors ultimately contribute to better educational outcomes. Additionally, electrical access directly enhances productivity, enabling the utilization of computers and laptops for studying. Indirectly, appliances like air conditioners can enhance comfort, quality of sleep, and overall productivity. Moreover, the availability of productivity-enhancing tools, such as computers, can also boost income, providing access to better resources like cars, which have been shown to positively impact educational attainment.

Conversely, electrical access can also yield negative educational outcomes. Entertainment mediums, including televisions, can introduce distractions that hamper educational progress. Furthermore, the increased earning potential resulting from access to machinery can raise the opportunity cost of education, potentially hindering educational outcomes. Electrification can also influence fertility outcomes. For instance, increased time spent on work or productive activities due to electrification can alter fertility patterns. The higher earning opportunities resulting from electricity may increase the opportunity cost of childbearing for women, thereby influencing fertility behavior. Moreover, women's increased awareness and education regarding the benefits of family planning and contraceptive methods can further impact fertility outcomes.

Based on the aforementioned discussion, it becomes evident that electrification can affect educational and fertility outcomes in various ways. However, to assess the specific effects of the Ethiopian Electricity Access Rural Expansion Project on these outcomes in rural Ethiopia, an empirical strategy is employed.

[6] Sources used for creating figure 1: (Jimenez 2017), (Aguirre 2017), (Wagner et al., 2017), and (Fuji, 2020)

3. Data and Methodology

3.1 Data

The data utilized in this paper has been sourced from the Demographic and Health Surveys (DHS) website, as previously mentioned. To implement my empirical strategy, I have compiled a comprehensive dataset by appending cross-sectional survey data from four different years: 2000, 2005, 2010, and 2016. These surveys were obtained from the DHS website upon request. The household survey data, gathered through questionnaires, covers a wide range of variables including age, health, and educational attainment. The dataset I am specifically utilizing contains information on all household members and employs a unique household ID variable to identify each household.

In addition to the survey data, I have also incorporated GPS datasets available on the DHS website. These datasets enable the georeferencing of household clusters within the survey data, providing GPS readings accurate to within 15 meters. For my analysis, I merged the GPS data from the years 2000, 2005, 2010, and 2016 with the survey data. This matching process allows for the identification of households within their respective zones, which is crucial in distinguishing between treated and control zones. The final dataset comprises 43,506 observations, each containing demographic and health information for household members within uniquely identified households situated in different zones across Ethiopia.

3.2 Research Design

The method I am using in this paper for assessment is a two-stage difference in differences model. The difference in differences method is a quasi-experimental approach that compares changes in outcomes over time between a group that has been treated with a program or intervention and a group that has not been treated.

For the identification of treatment and control groups, a map - indicating the zones where the program was active versus where it was not active - is used.

Ethiopia GBOBA Project Areas

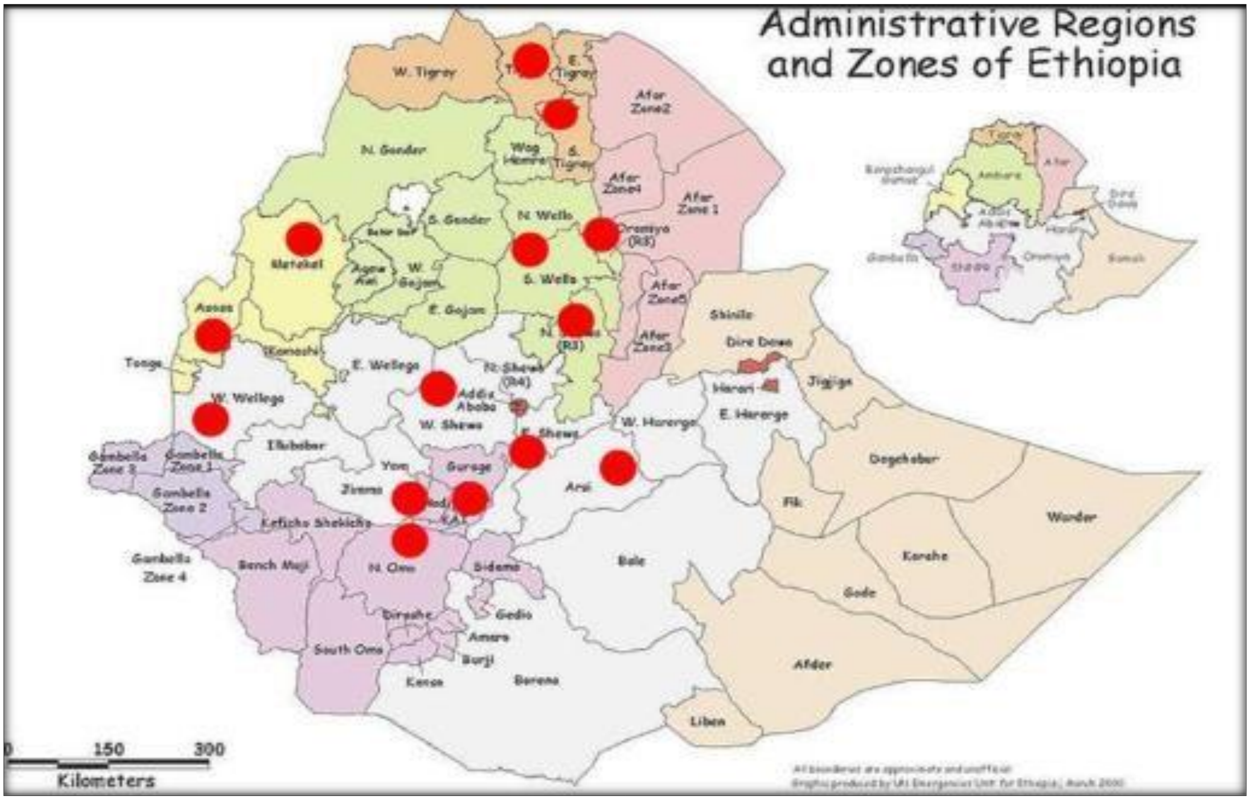

Figure 2

Source: World Bank (2016)

The red dots on the map above indicate what areas were treated. The rest of the areas I am using as controls.

Table 2 below shows the number of treated and control observations:

Table 2

	Count	percentage	cumulative percentage
Control	31568	72.56011	72.56011
Treated	11938	27.43989	100
Total	43506	100	
N	43506		

The table clearly demonstrates that the number of control observations is nearly three times larger than the number of treatment observations. This discrepancy is reasonable given the higher prevalence of households without electricity in the dataset.

The main regression specification that I am using in the paper is

$$\boldsymbol{Y}_{xyz} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 Treated}_{xyz} + \boldsymbol{\beta_2 Post}_{xyz} + \boldsymbol{\beta Post}_{xyz} * \boldsymbol{Treated}_{xyz} + \boldsymbol{Controls}_{xyz}$$
$$+ \boldsymbol{ZFE} + \boldsymbol{\epsilon xyz}$$

Here, x = household in question y = year in question and z = zone in question and Z = zonal fixed effects

In the equation, the variable "treated" serves as a binary indicator, tracking whether the household is located in a treated zone or not. Meanwhile, the "Post" variable indicates whether the household in question belongs to the group that was interviewed in 2016 or the group after the treatment years. The treatment effect is represented by Beta 3, which allows for the identification of differences in outcomes between individuals in the treated zone compared to those outside the treated zone during the post period.

The outcome variable of interest, denoted as Y in the equation, will vary depending on the specific stage and objective of the analysis. To account for variations across different zones, such as differences in weather and socioeconomic conditions, zonal fixed effects (indicated as Z in the equation) are employed. In addition, there are three control variables used in this research: Average household age, Number of household members, and Wealth index. These controls will be discussed in more detail later in this section.

There are three primary outcome variables of concern in this study. In the first stage, the analysis focuses on assessing whether there is a significant increase in electricity access in households located in treated zones during the post period compared to households in the control group. The outcome variable for this stage, denoted as Y, is the binary variable "has electricity," indicating whether the household has access to electricity or not.

After establishing the presence of increased electrification in households located in treated zones during the post period using the "has electricity" dummy variable, the analysis proceeds to the second stage. In this stage, the primary outcome variable is the average score of educational attainment, while the secondary outcome variable is the number of children aged five and under per woman. The educational attainment average score is utilized as an instrument for measuring educational attainment and is calculated by taking the mean of education values obtained by all household members in a given year. In simple mathematical terms, this can be understood as,

Educational attainment average score = household member 1 education in single years + household member 2 education in single years … household member n education in single years ÷ total number of household members (or n)

The secondary outcome 'number of children 5 years of age and under per woman' is used as to instrument for fertility. This variable is obtained by dividing the number of children in a household who are 5 years and under by the number of women in the household. The equation that describes the variable is

$$\boldsymbol{Fertility\ Outcome = \frac{Number\ of\ children\ 5\ and\ under}{Number\ of\ women\ in\ household}}$$

This outcome variable is highly suitable as it effectively captures recent birth rates and provides a reliable measure of the average number of children born per female within households. To ensure robust and accurate results, I will incorporate three essential controls in my model to minimize biases and address potential endogeneity concerns.

The first control variable, average household age, offers valuable insights by calculating the mean age of all household members. This control helps account for the potential influence of household demographics on the outcome variable.

The second control variable, the number of household members, serves as a straightforward indicator of the population residing within each household. By including this control, I can effectively adjust for variations in household size that may impact the outcome variable.

Lastly, the Wealth Index control variable plays a crucial role in addressing the potential influence of household income on the outcome variable. The Wealth Index is a composite measure that encompasses various indicators such as ownership of selected assets, housing construction materials, and access to water and sanitation facilities. By utilizing this comprehensive index, I can accurately control for disparities in household wealth across different categories.

The Wealth Index classification system divides households into five distinct categories based on their wealth levels, as detailed in Table 3. These categories include the poorest, the poorer, middle, richer, and richest households. Such stratification ensures a nuanced understanding of the wealth distribution and facilitates more precise analysis of the outcome variable in relation to household wealth.

Table 3: Wealth Index Variable description table

	Count	percentage	cumulative percentage
poorest	11529	34.86979	34.86979
poorer	7018	21.22614	56.09594
middle	6326	19.13317	75.22911
richer	5883	17.7933	93.02241
richest	2307	6.977588	100
Total	33063	100	
N	33063		

Summary statistics for the control variables, average household age, number of household members, and outcome variables are shown in Table 4:

Table 4

	Count	Mean	Standard Deviation	Min	Max
Has electricity	43493	.0555032	.2289624	0	1
Educational attainment average score	43496	1.209865	1.875127	0	19
Fertility outcome	34741	1.019754	.9118504	0	6
Average household age	43505	24.17821	12.53861	6.333333	95
Number of household members	43506	5.020388	2.376275	1	22
Wealth Index	33063	2.407827	1.3089971	1	5
N	43506				

The variation in the number of observations for the variables in Table 4 can be attributed to missing data points, as some variables were not reported by the research subjects. This can result in incomplete information, leading to a lower number of observations for certain variables.

The mean value of the 'Has electricity' variable aligns with expectations, considering that a significant portion of the rural population in Ethiopia lacks access to electricity. As a result, the mean value of this variable is close to zero, indicating a limited presence of electrical access in the sample.

The concentration of the mean for the educational attainment average score close to 1 is not surprising, given that Ethiopia is a developing nation with limited availability of schooling in rural areas. Despite the maximum value of the educational attainment average score variable being

19, the mean value is low, indicating that educational attainment levels are generally low in the sample.

Additionally, the mean value of the fertility outcome variable being close to 1 suggests that, on average, there is approximately one child under 5 years of age per female in a household. This finding indicates that the number of children under 5 per household is roughly equivalent to the number of women in the household.

Figures 3 and 4 show distributions for educational attainment and fertility variables:

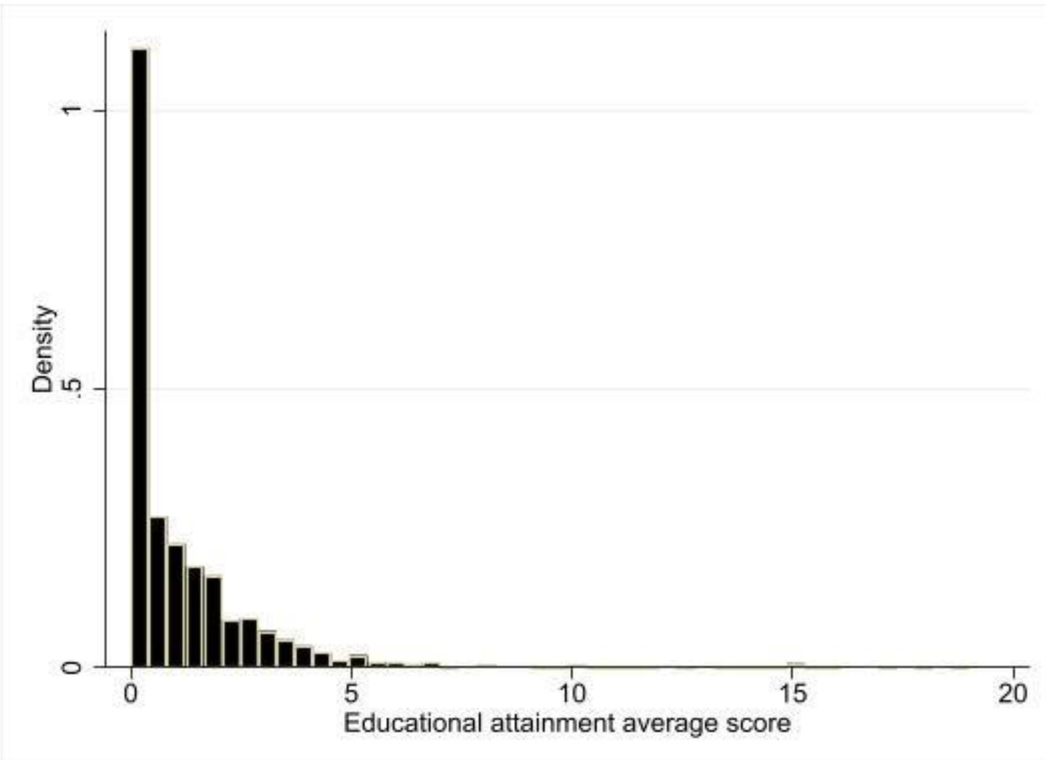

Figure 3

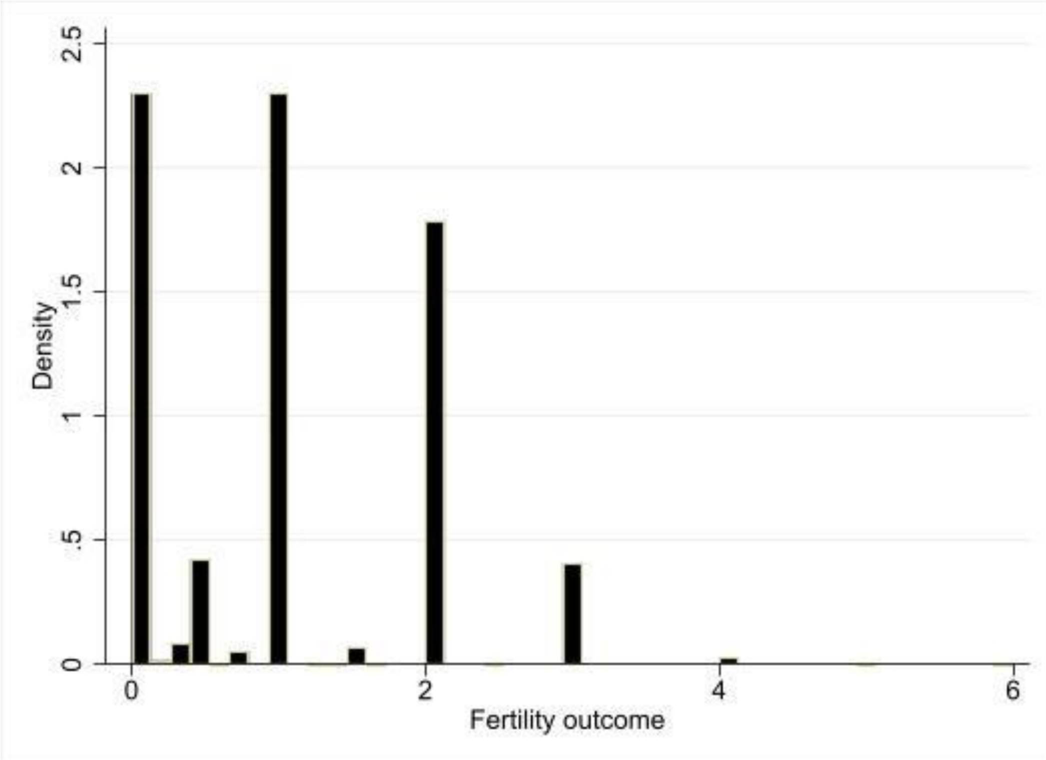

Figure 4

The data analysis in Table 4 reveals that the average household age in the observed households is approximately 24, indicating that the household members tend to be relatively young. However, it is crucial to acknowledge that the maximum score for the average educational attainment variable is 19. Consequently, houses with older members and a higher average household age might exhibit a greater average educational attainment estimate. This potential bias could have an upward influence on the results, underscoring the importance of controlling for average household age to mitigate such effects.

Another potential source of endogeneity lies in the number of household members. More household members can potentially contribute to a higher educational attainment score. Consequently, it is imperative to control for the number of household members to minimize biases in the results. The analysis reveals that the average number of household members is approximately 5, with a maximum value of 22.

Additionally, the mean value for the wealth index is approximately 2.4, suggesting that a majority of the households in the sample fall within the middle-income level. This finding provides valuable insight into the economic distribution within the sample population.

By taking into account these control variables and analyzing their impact on the results, we can enhance the accuracy and validity of the findings. Controlling for average household age and the number of household members allows us to disentangle the influence of these factors from the relationship being examined, thereby producing more robust and reliable results.

4. Results

4.1 Stage 1

Table 5: Treatment Effect of Rural Electrification Scheme on electricity outcome

	(1) Has electricity
treated (if in the treated zone)	-0.0252***
	(0.00282)
Post	0.0504***
	(0.00288)
Post treated	0.0564***
	(0.00560)
Constant	0.0453***
	(0.00149)
Observations	43493
Adjusted R^2	0.019

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

In Table 5, the coefficient associated with the treated dummy variable exhibits a negative value, indicating that during the preprogram period, the treated group had lower levels of electrical access compared to the control group. Specifically, a one unit increase in the treated variable corresponds to a negative 0.0252 decrease in electrical access. This suggests that the intervention group had less access to electricity before the program implementation.

However, the coefficient associated with the interaction term between the post period and the treated variable indicates a positive impact of the intervention on the treated group during the post-program period. This implies that there is a significant and positive increase in electrical access within the treatment group when compared to what would have been the case if the intervention had not taken place.

To visually represent this, Figure 5 displays a bar chart illustrating the means of the 'has electricity' variable for both the control and treatment groups during the pre and post periods. This visualization provides a clear comparison of electrical access between the groups and demonstrates the positive changes resulting from the intervention.

Overall, these findings suggest that the intervention had a beneficial effect on increasing electrical access within the treated group during the post-program period, highlighting the effectiveness of the program in improving access to electricity for the target population.

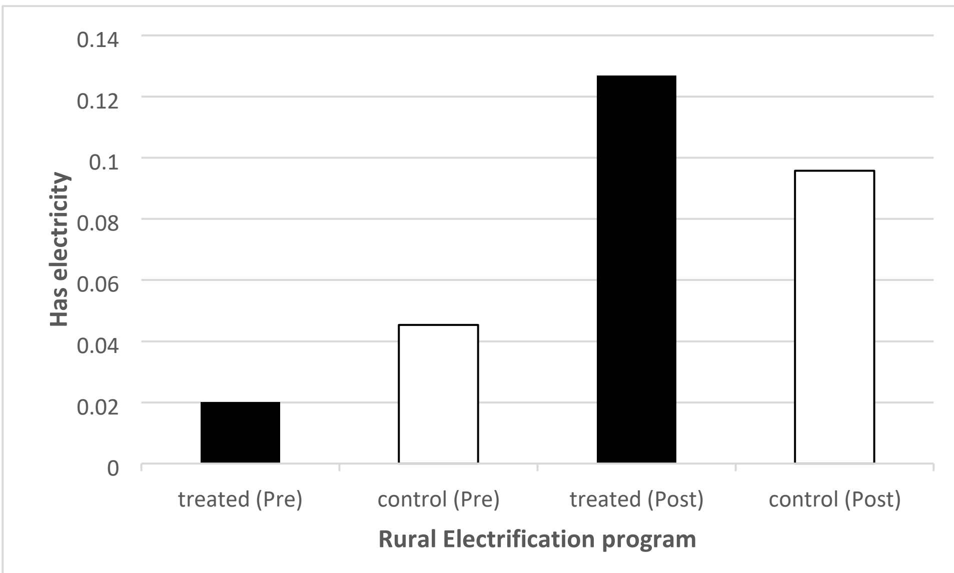

Figure 5

Examining Figure 5, it is evident that during the post-treatment period, the treated group experiences a notable positive increase in electrical access. Moreover, the average level of electrical access in the treated group surpasses that of the control group in the post-period. This visual representation provides strong evidence of the program's effectiveness in augmenting electrical access for the treated group.

Having established the positive impact of the program in terms of increased electrical access for the treated group, I now proceed to present the second stage results of my assessment.

4.2 Stage 2

4.2.1 Educational attainment outcome results

Table 6 shows results for four different regressions, one demonstrated in each column with columns numbered 1-4, with educational attainment average score as the outcome of interest.

Table 6: Treatment effect of Rural Electrification scheme on Educational attainment outcome

	(1) Educational attainment average score	(2) Educational attainment average score	(3) Educational attainment average score	(4) Educational attainment average score
treated (if in the treated zone)	0.0177	0.0408	0.0553	0.156***
	(0.0196)	(0.0250)	(0.0308)	(0.0353)
Post	0.796***	0.793***	0.735***	0.675***
	(0.0268)	(0.0261)	(0.0248)	(0.0356)
Post treated	0.306***	-0.0507	0.0169	-0.159*
	(0.0526)	(0.0510)	(0.0476)	(0.0624)
average household age		-0.0133***	-0.0134***	-0.00896***
		(0.000799)	(0.000874)	(0.00109)
number of household members		-0.117***	-0.112***	-0.0959***
		(0.00642)	(0.00474)	(0.00889)
Wealth index		0.536***	0.563***	0.538***
		(0.0101)	(0.00834)	(0.0141)
Constant	0.975***	0.740***	0.676***	0.421***
	(0.0112)	(0.0466)	(0.0461)	(0.0585)
Observations	43496	33055	32223	32223
Adjusted R^2	0.044	0.167	0.195	0.170

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The regression analysis involves several iterations to examine the treatment effect and the role of different variables. The first regression is a basic unweighted regression without any controls. The second regression introduces three controls: average household age, number of household members, and wealth index. The third regression incorporates zonal fixed effects along with the same three controls from the previous regression. Lastly, the fourth regression is a weighted regression that incorporates zonal fixed effects, the three controls (average household age, number of members, and wealth index), and household weights derived from the household weight variable. Household weight for a unique household in the data is the inverse of its selection probability multiplied by the inverse of the household response rate in the division.[7]

Across all regression columns, the coefficient on the treated variable is consistently positive and statistically significant, with a p-value of less than 0.001 in the fourth column. This suggests that there is a significant positive association between the treatment (electrification program) and the educational attainment average score. Specifically, a 1 unit increase in the treated coefficient corresponds to a 0.156 unit increase in the educational attainment average score during the pre-program period. This indicates that, on average, the treated group had a 0.156 unit higher educational attainment average score compared to the control group before the program implementation.

However, when examining the post-treated variable (representing the treatment effect), more intriguing results emerge. In the unweighted first regression, where no controls or fixed effects are included, the observed treatment effect is 0.306 units and statistically significant. However, when introducing controls into the model, the treatment effect becomes negative. In the fourth weighted regression with zonal fixed effects and all controls, the treatment effect is observed to be -0.159. This implies that, in the post-treatment period, the educational attainment average score decreased by 0.159 units for the group that received the electrification program.

Further analysis reveals that the introduction of the Wealth index variable plays a crucial role in driving the treatment effect towards the negative side.[8] The coefficients for the Wealth index variable in the table are statistically significant, and in the fourth column, it is observed that a 1 unit increase in the wealth index leads to a substantial 0.538-unit increase in the educational attainment average score. This finding aligns with theoretical expectations, as wealthier households would typically exhibit higher educational attainment outcomes.

[7] Description of household weight variable obtained from DHS program website. Accessed here: https://dhsprogram.com/data/Guide-to-DHS-Statistics/Analyzing_DHS_Data.htm
[8] Please refer to post treated variable coefficient in column 4 of table 9 in Appendix

Overall, these results indicate that the electrification program had a positive impact on the educational attainment average score during the preprogram period. However, after accounting for controls and introducing the Wealth index variable, the treatment effect becomes negative in the post-treatment period. The substantial increase in educational attainment associated with the Wealth index variable highlights the importance of household wealth in driving educational outcomes.

Educational attainment outcomes for poorest, middle income, and richest households

Table 7: Treatment effect of Rural Electrification scheme on Educational Attainment Outcome

	(1) Educational attainment average score (poorest)	(2) Educational attainment average score (Middle income)	(3) Educational attainment average score (richest)
treated (if in the treated zone)	0.192***	0.203***	0.0990
	(0.0432)	(0.0526)	(0.291)
Post	0.370***	0.839***	0.462*
	(0.0417)	(0.0744)	(0.218)
Post treated	-0.00444	-0.295**	-0.580
	(0.0846)	(0.110)	(0.377)
average household age	-0.00433***	-0.00124	-0.0212*
	(0.00123)	(0.00210)	(0.0101)
number of household members	-0.00134	-0.00565	-0.582***
	(0.00936)	(0.0143)	(0.0482)
Constant	0.677***	1.051***	6.933***
	(0.0768)	(0.115)	(0.447)
Observations	11201	6180	2237
Adjusted R^2	0.050	0.068	0.188

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7 presents three separate regressions to analyze the treatment effect of electrification on educational outcomes among different household categories. Column 1 focuses on the poorest households, column 2 examines middle-income households, and column 3 presents results for the richest households. The inclusion of the Wealth Index control enables a focused analysis within these specific household categories.

Across all three columns, the treated coefficient is consistently positive, suggesting that, on average, the treated groups within each household category exhibited better educational attainment outcomes compared to their respective control groups during the pre-treatment period. This finding underscores the positive impact of the electrification program on educational outcomes for households in all three income groups.

However, the coefficients on the post-treated interaction variables for all three groups are negative. This indicates a decrease in educational attainment outcomes in the post-period, after the treatment was implemented. Notably, in the middle-income group, a statistically significant negative treatment effect of -0.295 is observed. This implies that, for the treatment group in the middle-income category, the educational attainment outcome decreased by 0.295 during the post-treatment period.

4.2.2 Fertility outcome results

Table 8 shows results for regressions with fertility outcome as the dependent variable.

	(1) Fertility outcome	(2) Fertility outcome	(3) Fertility outcome	(4) Fertility outcome
treated (if in the treated zone)	-0.0429***	-0.00142	-0.00597	-0.0352*
	(0.0124)	(0.0122)	(0.0138)	(0.0169)
Post	0.0137	-0.0107	-0.00565	-0.00792
	(0.0135)	(0.0112)	(0.0112)	(0.0157)
Post treated	-0.0737**	0.0108	0.00315	0.0226
	(0.0250)	(0.0210)	(0.0216)	(0.0281)
average household age		-0.0672***	-0.0663***	-0.0662***
		(0.000691)	(0.000609)	(0.000917)
number of household members		0.0691***	0.0669***	0.0614***
		(0.00225)	(0.00208)	(0.00316)
Wealth index		-0.0451***	-0.0522***	-0.0509***
		(0.00342)	(0.00374)	(0.00486)
Constant	1.033***	2.138***	2.146***	2.192***
	(0.00666)	(0.0217)	(0.0215)	(0.0296)
Observations	34741	26205	25540	25540
Adjusted R^2	0.001	0.385	0.388	0.395

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table provides four sets of results, following a similar pattern to those in Table 6. The first column represents regression results without any controls, while the second column includes average household age, number of household members, and wealth index as controls. The third column incorporates zonal fixed effects in addition to the controls, and the fourth column displays results from a weighted regression using household weight, zonal fixed effects, and controls.

In the first column, the treated coefficient is statistically significant and negative. Similarly, in the fourth column where the results are shown for weighted regression with fixed effects and controls, the treated coefficient remains negative. This negative coefficient suggests that, during the pre-treatment period, the treated group had a lower fertility outcome compared to the control group. In other words, women in the treated group had fewer children under the age of five, on average, compared to women in the control group before the implementation of the treatment. The results in the first, second, and third columns may be more biased since the observations in the data were not properly weighted, potentially leading to an inaccurate representation of the true population.

The coefficient with post treated variable is significant and negative in the first column where there are no controls. In the second column, when controls are added, I observe that the coefficient becomes positive indicating a positive treatment effect. In the fourth weighted regression with controls and zonal fixed effects, a treatment effect of 0.0226 units is seen indicating that in the post period, there was an increase of 0.0226 children under the age of five per woman in each treated household observed. Interestingly, I again learn that this positive effect is highly influenced by income (indicated by the wealth index) as the treatment effect is negative without the introduction of the wealth index.[9]

4.3 Discussion

My findings regarding educational attainment outcomes suggest that the rural electrification program in Ethiopia had a significant negative impact during the post-intervention period on educational attainment in the zones that received electricity through the program. However, it is important to note that the significance diminishes when controls, fixed effects, and weighting are introduced, making the results more representative of the population.

Interestingly, the negative treatment effect seems to be driven by the wealth index, as there is a positive treatment effect when not controlling for wealth. This aligns with the theory that increased electrification can have negative effects on educational attainment through various indirect and direct channels, such as increased distractions resulting from electrical access.[10]

Furthermore, when examining educational outcomes within different sub-groups based on wealth index (poorest, middle income, and richest), I again observe a negative treatment effect. The impact is most pronounced in the richest population and least pronounced in the poorest population. This can be explained by the fact that the wealthiest population may be more

[9] Please refer to post treated variable coefficient in column 4 of table 10 in appendix

[10] More theoretical discussion on how electrification can impact education outcomes can be found in conceptual framework section of the paper.

susceptible to the negative impacts of electrification due to their affordability of distractions like televisions. On the other hand, the impact on the poorest households is relatively negligible, as they may not have access to electrical equipment that could enhance productivity or benefit from electricity in general.

In terms of fertility, the results indicate a positive impact of electrification. This implies that women in households within the treatment group, after the electrification program, had a higher average number of children under the age of five compared to the control group. This result contradicts the prevailing theory suggesting that electrification is more likely to decrease fertility. However, it can be theoretically justified by arguing that electrification may provide improved access to information that positively affects health outcomes and ultimately leads to increased fertility.

Overall, these findings shed light on the complex and nuanced effects of rural electrification programs on educational attainment and fertility outcomes. They highlight the importance of considering factors such as wealth index and sub-group analysis to better understand the differential impacts across populations.

5. Robustness

To enhance the robustness of my model and ensure the quality of my results, I employ several strategies. Firstly, I incorporate zonal fixed effects into my analysis. This approach allows me to address and minimize biases arising from variations between different zones. These variations encompass a range of observable and unobservable characteristics such as socio-economic stability, weather conditions, and other relevant factors. By accounting for these differences, I can enhance the reliability of my findings.

Additionally, I include suitable control variables to capture important factors that may influence the outcomes under investigation. These controls encompass variables such as the number of household members, wealth index, and average household age. By including these variables in my model, I can account for their potential effects and isolate the specific impact of the treatment or intervention being studied.

To address potential biases introduced by variations in the probability of selecting households within my dataset, I utilize sampling weights. These weights, represented by the household weight variable, allow me to adjust for these differences in the probability of selection.

By incorporating sampling weights, I can improve the robustness of my model and reduce heteroscedasticity, a source of potential bias in the analysis.

Upon applying the weighting technique, I observe changes in the treatment effects for my educational attainment and fertility outcomes. In Table 6, the treatment effect for educational attainment becomes negative and increases in magnitude, suggesting that the previously observed treatment effect was likely underestimated. Similarly, the positive treatment effect observed for fertility outcomes also increases in the weighted model, indicating that the original results may have underestimated the true treatment effect. These adjustments further enhance the validity and accuracy of my findings.

In addition to the aforementioned methods, I assess the parallel trends assumption to strengthen the validity of my model.[11] The difference-in-differences approach relies on the assumption that, in the pre-treatment period, both the treatment and control groups exhibit parallel trends in the outcome variables. This assumption implies that, in the absence of the treatment, the groups would have followed similar trajectories. By examining and confirming this parallel trends assumption, I can bolster the credibility of my model and the causal interpretation of the treatment effect.

By employing these strategies—incorporating zonal fixed effects, utilizing suitable controls, applying sampling weights, considering the impact of weighting on treatment effects, and checking for parallel trends—I improve the robustness, validity, and reliability of my model and the resulting findings.

Parallel Trends

To assess the parallel trends assumption, I examine the data from 2000 and 2005 and evaluate the trends in relation to the treatment and control groups. The analysis focuses on three outcome variables: 'has electricity,' 'educational attainment average score,' and 'fertility.'

Regarding the 'has electricity' outcome from the first stage and the 'educational attainment average score' outcome from the second stage, I find that the trends between the treatment and control groups in the pre-treatment period exhibit a fair level of parallelism. This indicates that both groups were following similar trajectories before the electrification intervention took place.

However, when examining the fertility outcome, the results are somewhat inconclusive and

[11] On next page

require further scrutiny. The trends observed for this outcome are borderline, indicating that the parallel trends assumption may not hold as strongly as desired. Nonetheless, it is noteworthy that both the treatment and control groups demonstrate positive trends in fertility over time. This suggests that the fertility outcome increased for both groups, irrespective of the treatment.

To provide a visual representation of these trends, I have included figures below that illustrate the observed patterns for the three outcome variables of interest. These figures serve as a useful visual aid in understanding the trajectories and trends exhibited by the treatment and control groups.

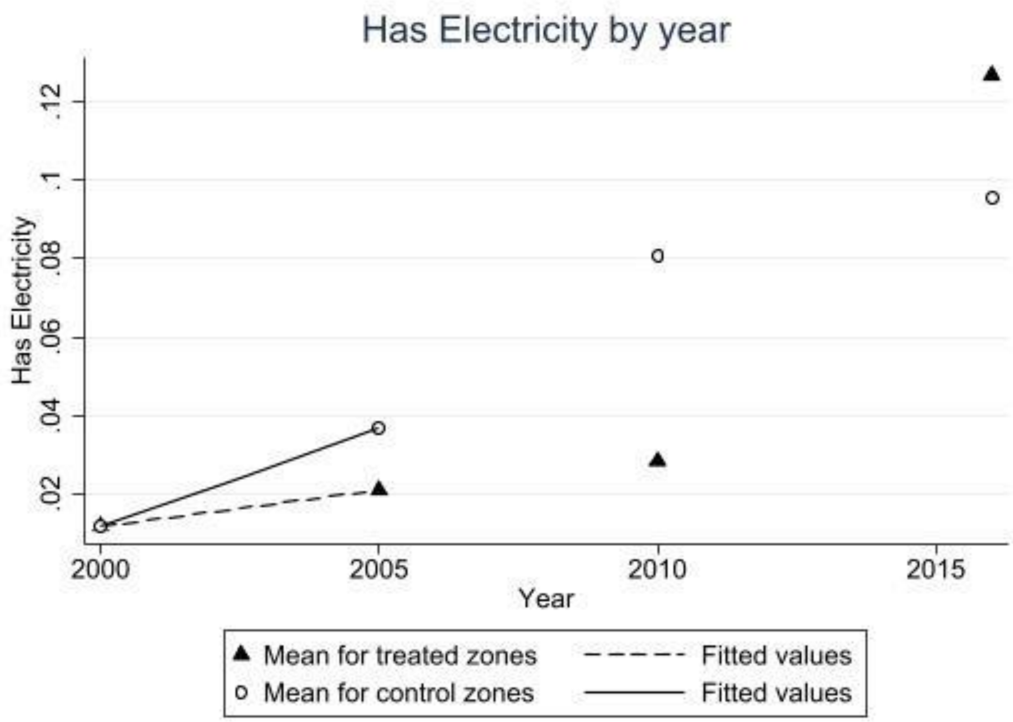

Figure 6

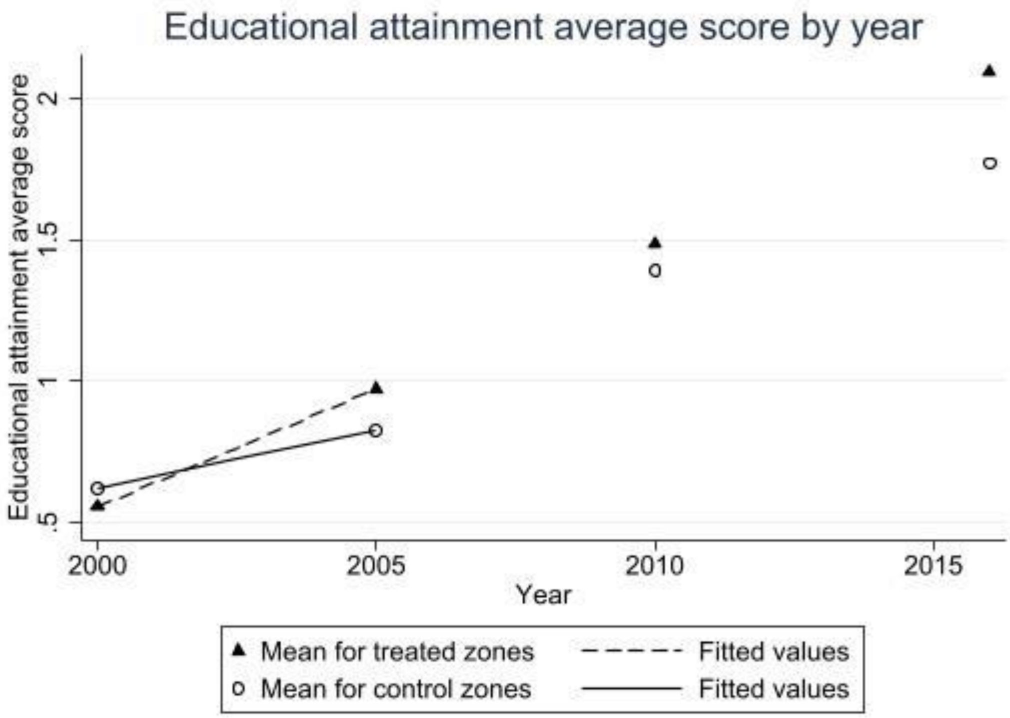

Figure 7

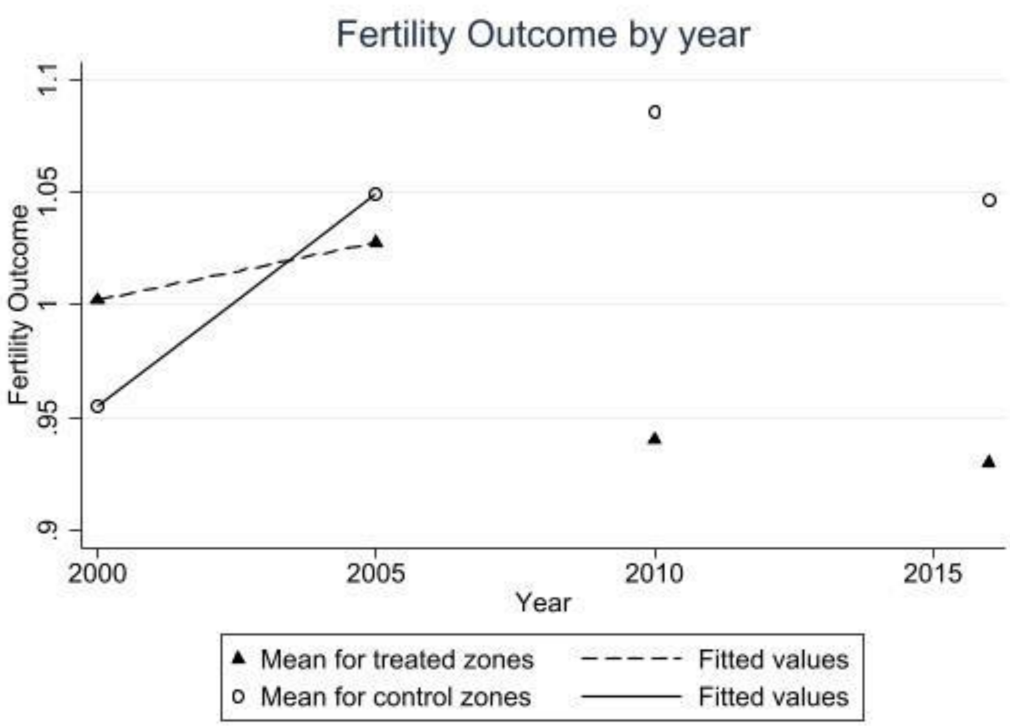

Figure 8

Concerns

I have identified several concerns regarding my results that warrant attention. Firstly, the lack of sufficient controls raises the possibility that other variables, such as parental education level or the total number of children, might be influencing the outcomes observed in my study. This potential issue could introduce omitted variable bias, which can compromise the validity of my findings. To address this, it would be beneficial to include parental education as a control since previous research has demonstrated its positive association with educational attainment.

Secondly, I have not thoroughly examined the heterogeneity within my sample. To gain a deeper understanding of the treatment effects of electrification, it would be valuable to perform triple differences analysis by creating interactions and studying the impact in different subsamples of the data. This approach will provide insights into how the treatment varies across different groups, enhancing the robustness of the conclusions drawn.

Furthermore, my use of cross-sectional datasets spanning four years limits the statistical power of my results. A more powerful alternative would be to employ a panel dataset, which would allow for a more comprehensive analysis. By having data for all years in the pretreatment period, I would be better equipped to test for parallel trends and establish stronger causal inferences.

Another concern lies in the relatively high standard error associated with the fertility outcome coefficients. Increasing the sample size would help improve the statistical significance of the fertility outcome and provide a more accurate coefficient estimate.

6. Conclusion

This paper examines the effects of the Global Partnership on the Output Based Aid component of the Ethiopia Electricity Access Rural Expansion Project Phase 2, as approved by the World Bank, on educational attainment and fertility outcomes in households located in treated zones in Ethiopia. Surprisingly, the findings reveal a negative influence of the electrification program on educational attainment outcomes in treated zones, while indicating a positive impact on fertility outcomes.

These results deviate from the majority of existing literature, which generally demonstrates a positive relationship between electrification and educational attainment, as well as a negative relationship between electrification and fertility. This inconsistency contributes to a deeper understanding of these specific research subareas and prompts further reflection on how electrification can affect these two outcomes. Moreover, it urges a broader examination of the overall developmental impacts of electrification and whether it successfully achieves its intended objectives.

Notably, the analysis indicates negligible impact on educational attainment outcomes for the poorest households in the program. This raises the crucial question of addressing other factors prior to electricity provision to ensure its effective utilization. Consequently, it is essential to provide opportunities for enhancing the welfare and income levels of the poorest households, as they may lack the financial resources necessary to fully benefit from electrification.

The observed negative relationship between fertility and electrification is also intriguing, as the existing research predominantly indicates a decrease in fertility with electrification. Given the limited amount of research available on this topic, this study is expected to stimulate further investigation into this area. However, it is important to acknowledge the limitations of this paper. The parallel trends observed for fertility outcomes are on the borderline, emphasizing the need for caution in interpreting these findings. Additionally, the reliance on cross-sectional data in this study suggests the potential for future research to employ panel data, enabling the observation of trends over time. Considering that electrification is believed to exhibit its most significant impact in the long term, further investigation involving multiple years after 2016 is warranted.

In terms of generalizability, while it may be possible to extend these results to other regions within Ethiopia due to shared conditions, caution must be exercised when generalizing to other countries due to cultural and socio-economic differences. Future research focusing on the impacts of electrification on educational attainment and fertility in different countries, utilizing larger panel datasets, would be valuable.

References

Aguirre, Julio., "The Impact of Rural Electrification on Education: A Case Study from Peru," The Lahore Journal of Economics, 2017, Vol. 22(1)

"Barnes, Douglas F.; Golumbeanu, Raluca; Diaw, Issa. 2016. Beyond Electricity Access: Output Based Aid and Rural Electrification in Ethiopia. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/26317 License: CC BY 3.0 IGO.

Burlig, Fiona, and Louis Preonas."Out of the Darkness and Into the Light? Development Effects of Rural Electrification,"2016. Working paper. Accessed here: https://epicstaging.uchicago.edu/wp-content/uploads/2019/08/WP268.pdf

Chakravorty, Ujjayant, Kyle Emerick, and Majah-Leah Ravago. 2016. "Lighting up the last mile: The benefits and costs of extending electricity to the rural poor." https://works. bepress.com/chakravorty/45/download/.

Jimenez, Raul, "Development Effects of Rural Electrification," 2017. *Inter-American Development Bank Infrastructure and Policy Brief.* Accessed here: https://publications.iadb.org/publications/english/document/Development-Effects-ofRural-Electrification.pdf

Lee, Kenneth, et al., "Experimental Evidence on the Economics of Rural Electrification," *Journal of Political Economy*, 2020, Vol. 128 (4).

Lee, Kenneth, Edward Miguel, and Catherine Wolfram. 2016. "Experimental Evidence on the Demands for and Costs of Rural Electrification." *NBER Working Paper 22292*.

Lewis, Joshua, and Severnini, Edson, "Short- and Long-Run Impacts of Rural Electrification: Evidence from the Historical Rollout of the U.S. Power Grid," 2017. *IZA institute of labor economics*. Accessed here: https://www.iza.org/publications/dp/11243/short-and-longrunimpacts-of-rural-electrification-evidence-from-the-historical-rollout-of-the-us-powergrid

Lipscomb, Molly, A. Mushfiq Mobarak, and Tania Barham. 2013. "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil." *American Economic Journal: Applied Economics* 5 (2): 200–231.

Richie, Hannah, and Roser, Max, "Energy," 2020.Our World in Data. Accessed here: https://ourworldindata.org/energy

Tomoki, Fujiji and Abu, Shonchoy, "Fertility and rural electrification in Bangladesh," Journal of Development Economics, 2020, Vol.143(C).

Appendix

Table 9: Treatment effect of Rural electrification scheme on Educational attainment outcome

	(1) Educational attainment average score	(2) Educational attainment average score	(3) Educational attainment average score	(4) Educational attainment average score
treated (if in the treated zone)	0.0177	0.0300	-0.0242	0.108***
	(0.0196)	(0.0196)	(0.0253)	(0.0282)
Post	0.796***	0.792***	0.768***	0.789***
	(0.0268)	(0.0265)	(0.0234)	(0.0365)
Post treated	0.306***	0.302***	0.324***	0.135*
	(0.0526)	(0.0518)	(0.0447)	(0.0624)
average household age		-0.0137***	-0.0138***	-0.00869***
		(0.000703)	(0.000778)	(0.000939)
number of household members		-0.0981***	-0.0943***	-0.0669***
		(0.00556)	(0.00412)	(0.00742)
Constant	0.975***	1.797***	1.813***	1.502***
	(0.0112)	(0.0462)	(0.0356)	(0.0610)
Observations	43496	43495	42350	42350
Adjusted R^2	0.044	0.057	0.091	0.065

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

Table 10: Treatment effect of Rural electrification scheme on Fertility outcome

	(1) Fertility outcome	(2) Fertility outcome	(3) Fertility outcome	(4) Fertility outcome
treated (if in the treated zone)	-0.0429***	0.00819	0.00591	-0.0195
	(0.0124)	(0.00981)	(0.0112)	(0.0137)
Post	0.0137	0.00658	0.00991	-0.00483
	(0.0135)	(0.0106)	(0.0105)	(0.0150)
Post treated	-0.0737**	-0.0279	-0.0328	-0.0116
	(0.0250)	(0.0198)	(0.0201)	(0.0266)
average household age		-0.0671***	-0.0665***	-0.0669***
		(0.000592)	(0.000525)	(0.000788)
number of household members		0.0654***	0.0635***	0.0588***
		(0.00193)	(0.00179)	(0.00267)
Constant	1.033***	2.042***	2.037***	2.086***
	(0.00666)	(0.0175)	(0.0170)	(0.0234)
Observations	34741	34741	33811	33811
Adjusted R^2	0.001	0.380	0.381	0.387

Standard errors in parentheses

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