THE IMPACT OF ELECTRICITY ACCESS ON EDUCATION IN KENYA

BY

GUYU YE

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Master's Committee:

Associate Professor Kathy Baylis, co-Chair
Assistant Professor Erica Myers, co-Chair
Associate Professor Brian Deal

ABSTRACT

As an integral component of modern society, access to electricity plays a pivotal role in economic prosperity. Its influence reaches to the foundations of industrial development, with the potential to enrich everything from healthcare, to transportation, to utilities and manufacturing. In this thesis, I explore the relationship between electricity and a fundamental contributor to economic growth: education. More specifically, I ask whether and how access to electricity improves educational outcomes in Kenya, one of the fastest growing countries in Africa where electricity demand has substantially outpaced its supply. A handful of prior studies have found mixed results in terms of the impact of electricity access on educational attainment in developing countries. Recent literature has recognized the endogeneity issue of electricity access, but it does not explore the potential mechanisms of the effect of electricity on educational outcomes. To address both issues, I use proximity to pre-existing transmission lines as an instrument for electricity access to estimate the effect on the average years of education for household clusters. On average, the number of years of schooling for households with electricity access is 4.1 higher than the ones without. The results under a traditional 2SLS framework show that electricity access plays an important role in improved educational attainment. In particular, when controlling for county and year fixed effects, gaining electricity access increases cluster-level average years of schooling by about 3.3 to 7.8 years. Moreover, nightlight brightness does not contribute additionally to educational attainment on top of the overall electricity access. This result is consistent with the one under a spatially lagged 2SLS framework, suggesting that electricity access contributes to improved educational attainment through mechanisms other than simply illumination.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Professor Kathy Baylis for her vision, support and dedication. Professor Baylis has been an incredible mentor to me since the beginning of my graduate program, and always provides me with opportunities to develop my research and technical skills. I also want to thank Professor Baylis for always being approachable and for allowing me to explore. Many thanks to Professor Myers and Professor Deal, as well as my Austin Energy family for preparing me well for a career in energy and sustainability.

I also want to thank my peers in ACE and UP, for all the good and bad times together, and the thoughtful discussions about research and life. Without each one of you, my graduate experience would have been totally different.

Thanks to the smart and wonderful friends in my life that have helped me get through the three years of graduate school. Particularly, thanks to Bobby Wade for proofreading many of my papers, Matt Meyer for technical and emotional support, and Anna Fairbairn for being an amazing roommate, study buddy and constant inspiration.

Thanks to my Hanover families for always believing in me, encouraging me to achieve higher goals, and having my back. I would not have been able to achieve half as much as I have now without my Hanover experience. Go Panthers!

Last but not the least, I want to thank my family for their love and support throughout my life. Thanks Mom and Dad for being the best and most selfless parents in the world.

TABLE OF CONTENTS

LIST OF FIGURES……………………………………………………………………………v

LIST OF TABLES…………………………………………………………………………..vi

CHAPTER 1: INTRODUCTION ……………………………………………………………...1

CHAPTER 2: BACKGROUND……………………………………………………………..5
2.1 Introduction/ Brief History and Urban Development ……………………………...5
2.2 Kenya's Electricity Sector ………………………………………………………7
2.3 Recent Development …………………………………………………………...12

CHAPTER 3: A REVIEW OF LITERATURE ……………………………………………16
3.1 Introduction ………………………………………………………………………16
3.2 Impact of Electricity Access ……………………………………………………...16
3.3 Impact of Electricity Access on Education………………………………………..17

CHAPTER 4: CONCEPTUAL AND ANALYTICAL FRAMEWORK ……………………….20
4.1 Introduction ………………………………………………………………………20
4.2 Basic Model Specification………………………………………………………...20
4.3 Spatial Lag 2SLS Model………………………………………………………….22

CHAPTER 5: DATA SOURCES AND SUMMARIES ……………………………………...25
5.1 Introduction………………………………………………………………………..25
5.2 USAID DHS Household-Level Survey Data……………………...………………25
5.3 Transmission Lines………………………………………………………………..31
5.4 NOAA Nighttime Satellite Imagery………………………………………………34

CHAPTER 6: RESULTS………………………………………………………………...39
6.1 Introduction ………………………………………………………………………39
6.2 Traditional 2SLS Models…………………………………………………………39
6.3 Spatial 2SLS Models..…………………………………………………………….45

CHAPTER 7: CONCLUSION AND DISCUSSION..……………………………………..51
7.1 Introduction ………………………………………………………………………51
7.2 Important Results and Implications ………………………………………………51
7.3 Limitations of the Analysis and Future Work ………………………...…………..53

BIBLIOGRAPHY………………………………………………………………………...55

APPENDIX: ADDITIONAL FIGURES...…………………………………………………61

LIST OF FIGURES

Figure 1: Provincial Map of Kenya…………………………………………………………………5

Figure 2: Structure of Electricity Sector in Kenya………………………………………………..9

Figure 3: Spatial Distribution of Electricity Access by County (2008/09 and 2014)….….……..29

Figure 4: Average Years of Schooling by County (2008/09 and 2014)…………………………31

Figure 5: Scatterplot of Distance and Electricity Access……………………………………...32

Figure 6: Scatterplot of Distance to Grid and Educational Attainment …………………………34

Figure 7: Comparison of Nighttime Satellite System Values……………………………………36

Figure 8: Scatterplot of DMSP Values and Projected VIIR Values….......................................36

Figure 9: Nighttime Brightness by County (2008/09 and 2014)………………………………...37

Figure 10: Scatterplot of Electricity Access and Nighttime Brightness……………………..…..38

Figure 11: Simulated Results of Increase in Educational Attainment…………………………...52

Figure 12: Electricity Generation, Transmission and Distribution Illustrated…………………...61

Figure 13: Nighttime Satellite Imagery in Kenya in 2009……………………………………….61

Figure 14: Kenya’s Transmission Network in 2012……………………………………………..62

Figure 15: Kenya’s Transmission Network in 2006……………………………………………..63

Figure 16: Simulations of Education Outcomes with Non-Spatial Model……………...…….64

Figure 17: Simulations with Education Outcomes with Spatial Model…………………………65

LIST OF TABLES

Table 1: Electric Power Generation Sources and Energy Generated in Kenya…………………10

Table 2: Roadmap of Kenya 5,000+ MW………………………………………………………..13

Table 3: Summary Statistics of Household-Level Characteristics ……………………………...28

Table 4: Household Characteristics with vs. without Electricity Access………………………..30

Table 5: Cluster-Level Characteristics by Distance ………………..…………………………...33

Table 6: Cluster-Level Characteristics by Type of Settlements …..……………………………33

Table 7: First Stage Results under Traditional 2SLS Framework………………………………40

Table 8: Second Stage Results under Traditional 2SLS Framework……………………………41

Table 9: First Stage Results under Traditional 2SLS Framework

with Nighttime Light………..…………………………………………………………...43

Table 10: Second Stage Results under Traditional 2SLS Framework

with Nighttime Light………………………..44

Table 11: Tests for IV under Traditional 2SLS Framework………..……………………...........45

Table 12: Moran's I Coefficients ……………………………………………………………….46

Table 13: Second First Stage Results under Spatial 2SLS Framework ………………………...47

Table 14: Second Stage Results under Spatial 2SLS Framework………………………………48

Table 15: Second Stage Results under Spatial 2SLS Framework

with Nighttime Light ……………………………………………………………………49

Table 16: Tests for IV under Spatial 2SLS Framework ……………………………………...50

CHAPTER 1. INTRODUCTION

Access to electricity is particularly crucial to human development, as certain basic activities—such as lighting, refrigeration, running household appliances, and operating equipment—cannot be easily carried out using other forms of energy (Crousillat, Hamilton, & Antmann, 2010). However, energy poverty, defined as the lack of access to modern energy sources, such as electricity and clean cooking facilities, still dominates much of the developing world. The *World Energy Outlook 2015* states that 1.2 billion people are still without access to electricity and more than 2.7 billion people rely on traditional use of biomass for cooking, which can cause harmful indoor air pollution (International Energy Agency [IEA], 2015). To address this issue, the former United Nations Secretary-General Ban Ki-moon pointed out that "affordable and reliable modern energy services are essential for alleviating poverty, improving health and raising living standards" (United Nations [UN], 2014).

One of the argued benefits of electricity access is that it can increase education, which is an important contributor to productivity growth and improved household welfare (Decker, Rice, Moore, & Rollefston, 1997). Former South African President Nelson Mandela once said that "education is the most powerful weapon which you can use to change the world" (Strauss, 2013). A skilled workforce sustains a country's long-term economic growth through cycles, because education not only expands a worker's capacity to perform tasks and use modern technologies, but also makes them more adaptable to change and better at communications (Berge & Fisher, 2013; Shankar, 2013; Decker et al., 1997). Extended lighted hours, powered by electricity, allow school-age children to read at night and potentially lead to better educational outcomes. However, whether electricity access in fact contributes to a higher educational attainment depends on a number of other direct and indirect factors, such as the number of hours that students would study with illumination, how the family views the importance of education, and whether the adoption of electric appliances frees up the time for more learning.

A small number of papers have studied the impact of electricity access on educational outcomes in developing countries, with mixed findings. While some papers find a positive effect of electricity on education, many find no effect. A study by Khandker, Samad and Barnes (2012) finds that electricity access generates an increase in hours spent studying; in particular, school-age boys and girls in rural India with electricity access study about 2.3 hours more at home compared to the ones without. However, Bensch, Kluve and Peters (2011) find no significant difference in kid's study hours at home between villages with and without electricity access in a case study in Rwanda. As for the impact on enrollment, Barron and Torero (2014) find no effect while Khandker et al. (2012) find increases in both enrollment and attainment. In terms of school attendance, Squires (2015) finds that access to electricity reduces school attendance and the completed years of schooling; however, it also decreases the hazard of dropping out in the last few years of school.

These studies have a number of limitations. Firstly, having electricity access does not equate to having electricity at night. Households that obtain electricity access through off-grid renewable equipment tend to still experience darkness at night. Only examining the impact of electricity access may fail to capture what it means to education for having electricity at night, which is an important question to ask when policymakers determine the best ways to increase educational attainment. Secondly, household decisions, such as investments on electricity and education tend to have a locational component. One household's decision may impact those of its neighbors, and the strength of this relationship may vary from place to place. Therefore, failure to recognize the spatial dependence between observations and the spatial heterogeneity in relationships could ignore the spillover effects from neighboring households and thus produce biased estimates of the impact of electricity access.

This thesis explores the impact of electricity access on education outcomes in Kenya using a two-stage least squared (2SLS) framework in a fixed effects model to take into account the potential endogeneity of electricity access. In addition, I examine the mechanisms of the effect of electricity through studying the added value of nighttime brightness. To address the potential bias from spatial autocorrelation among observations, I furthermore construct a spatially lagged 2SLS model with spatially lagged terms of selected variables. Datasets used in this analysis include household-level survey results from United States Agency for International Development (USAID)'s Demographic and Health Surveys (DHS) Program and nighttime satellite imagery from the National Oceanic and Atmospheric Administration (NOAA). Shapefiles of transmission lines digitized from power supply maps are also used to calculate the instrumental variable, which is the distance to pre-existing transmission line. What makes Kenya of particular interest is that although it faces similar development issues to other lower middle income countries, it has a comparatively democratic government that has put reducing energy poverty as a national priority. Its energy sector is also attracting investment from all over the world to reach the goal of universal electricity access (Sustainable Energy for All [SE4All], 2016).

Results show that electricity access does play a very significant and important role in improved educational attainment. In particular, when county and year fixed effects are controlled for, obtaining electricity access increases in the cluster-level average years of schooling by about 3.3-7.8 years. In terms of mechanism, nighttime brightness does not additionally increase the average years of schooling on top of electricity access. This result suggest that the effect of electrification is more likely to be carried out throughout the day, instead of just at night.

This thesis continues as follows. Chapter 2 provides essential background information on Kenya and Kenya's electricity sector. Chapter 3 reviews previous academic literature on the impact of electricity access in developing countries. Chapter 4 and 5 outline conceptual framework and data description for this analysis. Chapter 6 highlights and interprets some of the regression results. Chapter 7 concludes the analysis by summarizing important findings, discussing policy implications and addressing limitations of this study.

CHAPTER 2. BACKGROUND

2.1 Introduction/ Brief History and Urban Development

Kenya is situated in the eastern part of the African continent. The country lies between 5 degrees north and 5 degrees south latitude and between 24 and 31 degrees east longitude (see Figure 1). The equator passes through the middle of the country, separating the upper and lower parts almost equally. Kenya borders Ethiopia, Somalia, Tanzania, Uganda, and South Sudan. The coastline houses the port of Mombasa, which enables Kenya and several other countries, including Uganda, Rwanda, and South Sudan, to engage in global trade.

Urbanization in Kenya started following its colonization. Kenya's pivotal geographical location in the continent of Africa and fertile highlands have attracted European settlers since the 1800s. Its colonial history officially started from the Berlin Conference in 1885, when East Africa was first divided into territories of influence by the European powers (Coleman, 2013). During this time, cities emerged in response to an increased demand for transportation and settlement. Nairobi, for example, was built as a British railroad camp and

Figure 1. Provincial Map of Kenya

Source: Geology.com, 2017

supply depot for the Uganda Railway in the Maasai area due to its central position, proximity to a network of rivers and its high elevation that made it cool enough for comfortable residential living and to limit mosquitoes and mosquito-borne illnesses (Coleman, 2013). Even before it was officially declared a British colony in 1920, these settlers were allowed a voice in government, while the Africans and the Asians were banned from direct political participation until 1944.

Although economic development post-independence is mixed, Kenya has witnessed tremendous economic growth with the economic and structural reforms known as the Economic Recovery Strategy (ERS), started in 2005. This 5-year plan was launched by the Ministry of Planning and National Development to lift Kenya out of one of the country's longest recessions (Government of Kenya [GoK], 2003). With a central goal of promoting economic growth and increasing employment, the Strategy identified key policy actions, including providing compulsory and free primary education and initiating a National Social Health Insurance Scheme (GoK, 2003). This blueprint led Kenya to make significant progress in reversing its poor economic performance of the preceding decade, as well as increasing social infrastructure, such as education and health facilities (Kenya Vision 2030, 2007). The Strategy finished up with a 114% increase in GDP per capita (after adjusting for purchasing power parity [PPP]) from 2003 to 2008, and has laid solid educational and political foundations for Kenya's future economic growth (World Bank [WB], 2017).

During the recent years, Kenya has been taking steps toward democracy. Following the March 2013 general elections, the Kenyan national government devolved substantial powers to the districts (WB, 2016). A significant portion of public finances and responsibility for service delivery in health, agriculture, urban services and local infrastructure was handed over to 47 new county governments in less than a year. This ambitious devolution shifts some key decision-making from central to county governments, creating a window of opportunity for more "bottom-up" engagement, backed by a Constitution and legal framework that include provision for government to share information, consult the public and regularly gather citizen feedback.

Like many other lower-middle countries in sub-Saharan Africa, Kenya still faces a number of development challenges, such as high energy costs, a large reliance on the agricultural sector, high levels of corruption, and high rates of inflation (Institute of Economic Affairs [IEA], 2015). High costs of energy, in particular, limit Kenya's economic activity (Kenya Institute for Public Policy Research and Analysis [KIPPRA], 2010). After adjusting for PPP, which has a conversion factor of 43.76 Kenyan Shillings to one US Dollar in 2015, the cost of electricity in Kenya is four times that of the United States (WB, 2017; Regulus, 2017; U.S. Energy Information Administration [USEIA], 2017). Compounding the high cost is the inconsistency of power; constant electricity outages are still a common phenomenon experienced by Kenyan residents due to an aging energy network and insufficient generation capacity (Biryabarema, 2016). The lack of access to affordable and reliable electricity costs Kenya foreign direct investment, with considerable penalties on socio-economic development. Therefore, to be prepared for future growth and sustainable development, Kenya needs to prioritize increasing the affordability, accessibility and reliability of electricity; strategies on how to achieve this goal will set an example for other countries that are facing similar challenges to follow.

2.2 Kenya's Electricity Sector

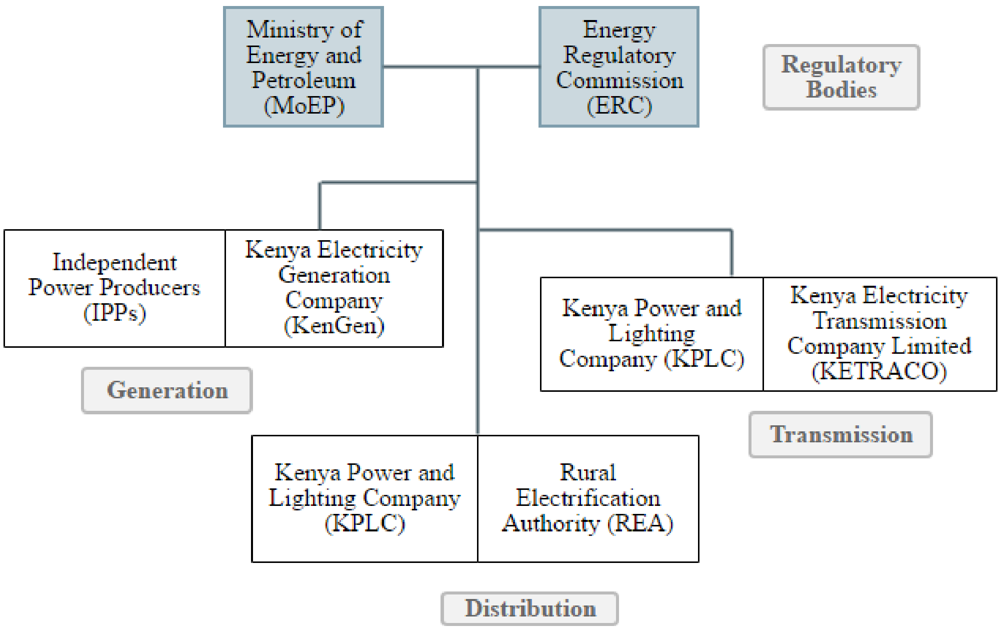
2.2.1 Structure and Main Players

Historically, the electricity sector in Kenya had been vertically integrated under the Kenya Power and Lighting Company (KPLC) (KPLC, 2017). In 1997, the Government of Kenya embarked on electricity sector reforms to increase competition and "ensure affordable, sustainable and reliable supply to meet national and country development needs" (Ministry of Energy and Petroleum [MoEP], 2012). The first step was to unbundle Kenya's generation from transmission and distribution in 1998 with the creation of Kenya Electricity Generation Company (KenGen) (generation) and KPLC (transmission and distribution) (Chimbaka, 2015; Power Africa, 2015). In 2008, the establishment of the Kenya Electricity Transmission Company Limited (KETRACO) further separated transmission from distribution and marked the completion of unbundling. Even though KPLC still operates the transmission lines constructed prior to the 2008 separation, the electricity sector in Kenya is currently largely unbundled (Power Africa, 2015).

Kenya's electricity sector is currently governed by two regulators (see Figure 2). The Ministry of Energy and Petroleum (MoEP) is responsible for overall policy coordination and development in the energy sector. It sets the strategic direction for the growth of the sector and provides long-term visions for all players in the sector (MoEP, 2016). The Electricity Regulatory Commission (ERC), established during the reform under the Electric Power Act in 2006, is an autonomous, independent sub-sector regulator. It sets, reviews and adjusts consumer tariffs, approves power purchase agreements, promotes competition in the sub-sector where feasible, resolves consumer complaints and enforces environmental, health, and safety regulations (Electricity Regulatory Commission [ERC], 2016).

Electricity is generated by KenGen and an increasing number of Independent Power Producers (IPPs). KPLC, the wholesale buyer and sole distributor of electricity in Kenya, purchases electricity from all generators through negotiated power purchase agreements (PPAs) approved by the ERC. Electricity is then conveyed over transmission systems owned by KPLC or KETRACO. Finally, KPLC carries out distribution and retail supply of the electrical energy to customers in accordance to licenses and permits issued by the Commission.

Figure 2. Structure of Electricity Sector in Kenya



Source: Ministry of Energy and Petroleum, 2017

2.2.2 Current Generation Capacity

Electricity demand is expected to increase dramatically in Kenya in the next few decades. Energy-intensive activities such as mining and manufacturing are anticipated to grow rapidly as the county governments take shape post-devolution, and residential demand for electricity is likely to go through the first transition of the S-shaped curve as the first-time ownership of appliances starts to increase (MoEP, 2015; Wolfram et. al., 2012). To meet the industrial and residential demand for affordable and reliable electricity, the Ministry of Energy and Petroleum has proposed and is implementing a roadmap to increase the installed generation capacity from 1,664 MW as of October 2013 by at least 5,000 MW by 2017 (MoEP, 2015).

With one of the greenest energy mixes in the world, Kenya generates over 80% of its electricity from renewable energy sources, such as hydro, thermal, geothermal, and wind (see Table 1). Oil-fired medium speed diesel (MSD) power generators contribute to only about 20% of electricity generated in 2014/15. Due to the unreliability of hydro, government is planning on adding more generation capacity from wind, thermal and geothermal. In particular, the current 2017 roadmap aims at shifting Kenya's current generation mix to 55% thermal, 32% of geothermal and close to 90% of installed renewable energy capacity (Power Africa, 2015). This diversified portfolio will likely result in reduced cost, and increase stability of electricity supply, enabling more households to be connected to the grid.

Table 1. Electric Power Generation Sources and Energy Generated in Kenya

Source of Electric Power Generation	**Installed Capacity (June 2014)**		**Annual Generation (FY 2014/15)**	
	MW	%	GWHrs	%
Hydro	821	37.8	3,466	36.8
Geothermal	593.5	27.3	4,060	43.1
Wind	25	1.2	37	0.4
Cogeneration	38	1.7	14	0.2
Imports	-	-	79	0.8
Total Renewable	1477.5	68	7,657	81.3
MSD	579.5	26.7	1,643	17.4
Gas Turbines	60	2.8	4	0
HSD (Isolated Stations)	25.8	1.2	36	0.4
Emergency Power Plant	30	1.4	84	0.9
Total Fossil Fuels	695.3	32	1,776	18.7

Source: KPLC and ERC, 2015

2.2.3 Transmission and Distribution

Electricity transmission is the process by which large amounts of electricity produced at power plants is transported over long distances for eventual use by consumers. Electricity is usually transmitted from the generator to a substation near a populated area. At this substation, the high-voltage electricity is converted to lower voltages suitable for consumer use, and then transmitted to end users through relatively low-voltage electricity distribution lines that are owned and operated by the national electricity utility. Therefore, households physically close to a transmission line are likely to have a lower cost of being connected to the grid, and a higher probability of future electricity access.

As the owner and operator of Kenya's new transmission lines since the final unbundling in 2008, KETRACO plans to construct over 4,000 kilometers of high-voltage transmission infrastructure in the next 3 to 4 years to open up geographical areas without access to the national grid, enhance the stability of power supply and build inter-connectors to facilitate regional power trade (KETRACO, 2017). To achieve these goals, three types of projects are underway, including Electricity Scale-Up Projects, such as the Kilimambogo – Thika – Githambo Line in the Central region and Mumias – Rangala Line in the Western region, Major Projects, such as the Mombasa – Nairobi Line and the Radai – Malindi – Garsen – Lamu Line, and the Regional Power Interconnections Projects, such as the Zambia – Tanzania – Kenya (ZTK) Line and the Lessos – Tororo Line between Kenya and Uganda. Transmission expansion projects connect more regions to the national grid, and thereby are a necessary condition for increasing electricity access.

New transmission routes are identified using a consultative approach. KETRACO is required to first look at major corridors, such as existing utility lines, roads and railroads, and then work with affected communities to balance the consideration of several factors including engineering, environment, current and future land use, and electrical needs from the community. Routes that are under consideration are shared with the public for comments and input to most benefit the community (KETRACO, 2017).

Distribution and retail sales of electricity in Kenya is solely handled by KPLC. To extend the low-voltage network to reach households, especially in rural areas, KPLC has been implementing the first phase of the Last Mile Connectivity initiative since September 2015. All households located within 600 meters of identified transformers across all the 47 counties are being connected, and the full implementation will facilitate Government objective of connecting 70% of Kenya households by 2017 and achieving universal access by 2020 (KPLC, 2016). During this past year, Kenya was able to add 1.3 million households to its electricity grid, raising the percentage of connected Kenyans to 55% from just 27% in 2013 (Kuo, 2017). However, this initiative has been criticized because customers in the least connected areas tend not to be able to afford the connectivity fee of Sh15,000 (approximately $140) in full, and they are thus required to pay monthly instalments until the connectivity fee is paid off, which is not what most low-income residents are willing to do (Okoth, 2016).

2.3 Recent Development

2.3.1 Kenyan Government

Government support plays an integral role in increasing electrification rate throughout the country. To build on the momentum from the ERS, in 2008, H.E. President Mwai Kibaki launched the Kenya Vision 2030, which aims at doubling Kenya's rate of growth and making it a middle income country by the year of 2030 (Kenya Vision 2030, 2007). The consultative approach adopted in developing this blueprint enabled the government to identify priorities, such as increasing energy generation and expanding transmission and distribution networks. In particular, the government has proposed specific short-term goals of new addition in capacity for every six months, which will likely increase the accessibility and affordability of electricity (see Table 2)

Table 2. Roadmap of Kenya 5,000+ MW

Time in Months		**12**	**18**	**24**	**30**	**36**	**40**	**Total**
Hydro	24	-	-	-	-	-	-	24
Thermal	87	163	-	-	-	-	-	250
Geothermal	90	176	190	50	205	150	785	1,646
Wind	-	-	20	60	300	250	-	630
Coal	-	-	-	-	960	-	960	1,920
LNG	-	-	-	700	350	-	-	1,050
Cogeneration	-	-	18	-	-	-	-	18
Total	201	339	228	810	1,815	400	1,745	
Cumulative Additions		201	540	768	1,578	3,393	3,793	5,638

Source: USAID (2016)

The Rural Electrification Authority (REA) was established in 2006 by the Government of Kenya to expand electricity access to rural residents. Many rural electricity expansion projects have been carried out by REA to connect schools, public facilities, and off-grid areas (such as islands in Lamu County). Remarkably, the REA announced that over 96 percent of primary schools were connected to electricity by the end of May in 2016, and it is starting the country's first mini-hybrid solar power station in south Wajir (Songok, 2016). Currently, the Authority is developing framework to promote photovoltaic solar panels among residents and businesses in the rural area to take advantage of the identified renewable energy potential in the country (Rural Electrification Authority [REA], 2016).

Overall, the Government of Kenya has put in place an ambitious plan to increase energy supply through heavy infrastructure investments and exploration of natural resources for power generation. If these plans are implemented, they should result in sufficient growth in energy consumption as well as enhance economic growth across all economic sectors.

2.3.2 Other Countries and Organizations

Outside sources and organizations have been helping tackle Kenya's energy poverty issue as well, including the World Bank, Africa Development Bank, and the Government of China. One of the most prominent projects is the Power Africa Initiative launched by President Obama in 2013, which brings together technical and legal experts, the private sector, and governments from around the world to work in partnership to increase the number of people with access to power in sub-Saharan Africa (Power Africa, 2016a; Power Africa, 2016b). In Kenya, Power Africa is supporting the development of the energy sector through financing, grants, technical assistance, and investment promotion. It is focused on using innovative solutions to connect rural Kenyans to the electricity grid, including supporting small on-grid power generation projects and off-grid and mini-grid solutions for small communities (Power Africa, 2016b). Kenya is also one of the three countries that is being assisted by the US National Association of Regulatory Utility Commissions (NARUC) to develop an electricity tariff scheme for cross-border electricity transactions (Power Africa, 2016b). By September 2016, Power Africa has funded projects that are expected to provide over 1.2 million new off-grid connections for homes and businesses, and helped facilitate the financial close of private sector power transactions that are expected to generate 8.4MW biogas in Cummins Bariago and 370MW wind power in Kinangop and Lake Turkana (Power Africa, 2016b).

The World Bank, which is also one of the main participants in the Power Africa initiative, is a major investor in the African energy sector. Notably, in FY 2016, it had a portfolio of 46 active projects across the continent totaling $9.7 billion (Power Africa, 2016b). Aside from capacity building and technical assistance, the World Bank Group's approach consists primarily of financial support through investment lending, loans and guarantees from the Multilateral Investment Guarantee Agency (MIGA) and the International Bank for Reconstruction and

Development (IBRD), and investments from the International Finance Corporation (IFC). In 2016, World Bank President Jim Kim co-signed the Power Africa Roadmap, along with U.S. President Obama and AfDB President Adesina, reaffirming strong support for Power Africa's path to reaching 30,000 MW and 60 million connections.

Governments of other more developed economics, including China, are also funding energy projects in Kenya (KETRACO, 2017). For example, the Olkaria – Lessos – Kisumu Line, a major project by KETRACO to distribute supply from Olkaria Geothermal Power Plant and to strengthen the link between the Eastern and Western Grid, was financed by the Japan International Cooperation Agency (JICA), and several electricity scale-up projects aimed at to expanding electricity network were financed by the KBC Bank in Belgium. This financial support and technological diffusion from other governments and organizations will help Kenya to achieve the goal of increasing installed generation capacity by 5,000+ MW by 2017 and universal electrification by 2020, as outlined in Kenya Vision 2030.

CHAPTER 3: A REVIEW OF LITERATURE

3.1 Introduction

Electricity alone may not be able to create all the conditions for economic growth, but it is essential for many basic human needs and modern economic activity (IEA, 2015). In theory, access to electricity can improve socio-economic conditions in developing countries through its influence on key components of poverty, namely health, education, income and the environment; however, whether this assumption holds needs further investigation (Kanagawa & Nakata, 2008; Dinkelman, 2010). This section reviews previous literature on the impact of electricity access in developing countries, and provides a summary of their methodologies and findings.

3.2 Impact of Electricity Access

Electricity is a necessity for modern way of life; it drives economic growth in many ways, such as productivity, employment, and market competition. A study by Dinkelman (2010) analyzes South Africa’s mass roll-out of electricity to rural households. She uses a 2SLS model with land gradient as the instrumental variable for electricity access, and furthermore complements this specification with a district-level trend analysis, where she pools information from four cross-sections of South Africa household survey data to estimate the impact of electrification on male and female employment, hours of work, wages and earnings. After regressing each of the labor market outcomes on age, age-squared and years of education, she finds large increases in the use of electric lighting and cooking, and reductions in wood-fired cooking over a five-year period, as well as a 9.5 percentage point increase in female employment.

Burlig and Preonas (2016) examine the medium-run effects of India's national rural electrification program, Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY), using a regression discontinuity design with high-resolution geospatial data. In particular, they use the population of 300 as the cutoff, because villages that contained habitations with populations of 300 or more are eligible for treatment under the program rules. Results suggest that the effects of rural electrification may be less beneficial than previously thought.

Several studies also suggest that electricity access has a positive impact on productivity and thus efficiency (Kirubi, Jacobson, Kammen, & Mills, 2009; Blalock & Veloso, 2007; Barnes & Binswanger, 1986). In particular, through a case study analysis of a community-based electric micro-grid in rural Kenya, Kirubi et al, (2009) finds that the use of electricity can increase productivity per worker by approximately 100-200% for carpenters and by 50-170% for tailors, depending on the item being produced.

In terms of environmental and health outcomes, electricity leads to improvements in indoor air pollution, which reduces the incident of acute respiratory infections among children, and lowers the exposure to pollutants among adult household members (Barron & Torero, 2017).

3.3 Impact of Electricity Access on Education

One of the most important outcomes that can be affected by electrification is education, which is positively correlated with economic growth (Hanushek & Woessmann, 2012; Berger & Fisher, 2013). Illumination enables students to study at more flexible hours, and potentially longer. Estimates show that sub-Saharan Africa has the lowest rate of primary school electrification, with just 35% of schools having access to electricity, which makes it difficult to promote education (Provost, 2013). SolarAid, an international charity that combats poverty and climate change by providing access to solar lights to some of the most remote regions in Africa, conducted focus group discussions with school children in Kenya, Malawi, Tanzania and Zambia. They discover that students rated limited lighting as their biggest obstacle to learn and do homework (Harrison, Scott, & Hogarth, 2016). However, there have been mixed results in previous literature on whether electricity access contributes to better educational attainment at all.

Khandker et al. (2012) use a two-stage least squares framework to evaluate the socio-economic outcomes of electricity access in rural India (Khandker et. al., 2012). To account for the endogeneity of electricity access, they use a vector of instruments that include a variable indicating the proportion of households in a community with electricity and its interaction with such household-observed characteristics as household head's age, gender, and education. Using the 2005 India Human Development Survey (IHDS), a nationally representative household-level dataset, and a region fixed effect model, they find that in households with electricity, both boys and girls spend more time studying than households without electricity, suggesting a better educational outcome in the future.

A recent job market paper by Tim Squires (2015) examines school attendance, total years of schooling and the dropout rates during the electrical expansion in Honduras from 1992 to 2005 (Squires, 2015). He uses individual-level data collected from Encuesta Permanente de Hogares de Propósito Múltiples (EPHPM) with a model with municipality and year fixed effects. Overall, he finds that electricity access decreases educational attainment. In particular, access to electricity reduces school attendance by 4 percentage points, and the number of years exposed to electricity reduced completed years of schooling by approximately 0.1 years. It also increases the hazard of dropping out in the first few years of school while decreases the hazard in the last few years of schools. One of his explanations is that access to electricity is associated with an increase in employment for children, which translates into higher dropout rates and lower educational attainment.

Barron and Torero (2014) designed a randomized control trial to study the socio-economic impact of electrification on rural households during a recent grid expansion program in Northern El Salvador. In this program, the government covered all the installation costs for all 500 eligible households, while the households were responsible for their own internal wiring and a connection fee of about $100. Barron and Torero randomly allocated 200 low-discount (20% discount) vouchers and 200 high-discount (50% discount) vouchers among the households, and kept the remaining 100 as the control group. They construct a 2SLS model with the level of voucher as the instrument for connection status, and find that electrification leads to a 54 percentage point increase in the probability among school-age children to study at home, and a 84 percentage point increase in performing other school-related activities.

While supporting better access to lighting for children at home appears to increase study hours, there are other ways to enhance educational opportunities relating to energy-use. A study by Cabraal, Barnes and Agarwal (2005) suggests that the presence of household electricity in rural areas can attract more effective teachers, and that improved education outcomes can also be obtained indirectly from electricity through this channel (Cabraal, Barnes, & Agarwal, 2005).

CHAPTER 4: CONCEPTUAL AND ANALYTICAL FRAMEWORK

4.1 Introduction

To analyze the effect of electricity access on educational attainment, I combine a 2SLS fixed effect model with spatial econometric techniques. In particular, I ask how electricity access affects average years of schooling for a Kenyan household and whether illumination measured by nighttime light plays an additional role in improved educational attainment. This section identifies these variables of interest, and lays out the empirical strategy for my analysis.

4.2 Basic Model Specification

Electricity use is influenced by the availability and price of electricity, household characteristics and community characteristics (Barnes et al., 2010). To estimate the causal effect of electricity on selected socio-economic indicators controlling for cluster-specific and region-specific characteristics as well as year fixed effects, I consider the outcomes conditional on electricity access expressed as follows:

$$Y_{it} = \alpha_1 + \beta_1 X_{it} + \gamma_1 V_{it} + \mu_r + \eta_t + \varepsilon_{rt}$$

Where Y_{it} denotes the education outcome (average years of schooling in particular) in household cluster i in year t, X_{it} is the dummy variable indicating the electrification rate for a household cluster i in year t, V_i is a vector of observable cluster-level characteristics, such as type of settlement, gender and average age of household head, μ_r represents region fixed effects, η_t represents unobserved characteristics that are specific to a particular year, and ε_{rt} is the randomly distributed error term.

Standard linear regression models assume that errors in the dependent variable are uncorrelated with the independent variable(s). When this is not the case, OLS no longer provides unbiased model estimates. As noted in several papers (Khandker et al., 2012; Barron & Torero, 2014; Dinkelman, 2010), households are not randomly connected to electricity and villages are not randomly selected for electrification. Instead, the decision to electrify is often based on both observed and unobserved characteristics, such as an area's geographic features and a household's ability to perceive returns to investment. If households who are more willing to invest in their future both see the returns to investment in electricity, and are thus more likely to be electrified, but also are willing to invest in education, resulting in higher educational attainment, then the failure to control for this correlation will yield an upwardly-biased estimated effect of electrification on education. To address this issue, I introduce an instrumental variable (IV) that directly influences the probability of electricity access under a 2SLS framework. 2SLS uses instrumental variables that are uncorrelated with the error terms to compute estimated values of the problematic predictor(s) (the first stage), and then uses those computed values to estimate a linear regression model of the dependent variable (the second stage). Since the imputed values of the endogenous variable are based on variables that are uncorrelated with errors in the second stage, the results of the two-stage model are unbiased. Adjusted theoretical model reads as follows:

$$\text{First stage:} \quad X_{it} = \alpha_0 + \alpha_1 Z_{it} + \alpha_2 V_{it} + \mu_r + \eta_t + \varepsilon_{rt}$$

$$\text{Second stage:} \quad Y_{it} = \beta_0 + \beta_1 \hat{X}_{it} + \beta_2 V_{it} + \mu_r + \eta_t + u_{rt}$$

Where Z_{it} represents the instrument for each household cluster in a particular year, and $\hat{X}_{it}$ is the predicted values of X_{it} by using the instrument selected.

To further explore the mechanism of the impact and measure whether illumination in particular plays additional role in improved educational attainment on top of overall electricity access, I include the values of nighttime brightness as well as an interaction term of electricity access and brightness in the second stage. In this case, the model specification reads as follows:

First stage: $X_{it} = \alpha_0 + \alpha_1 Z_{it} + \alpha_2 V_{it} + \alpha_3 N_{it} + \mu_r + \eta_t + \varepsilon_{rt}$

Second stage: $Y_{it} = \beta_0 + \beta_1 \hat{X}_{it} + \beta_2 V_{it} + \beta_3 N_{it} + \beta_4 \hat{X}_{it} N_{it} + \mu_r + \eta_t + u_{rt}$

Where N_{it} represents the level of nighttime brightness for each household cluster in a particular year.

For the instrument to be valid, the IV method required two conditions: the instrument should be as good as randomly assigned, and it should only affect *Y* through *X* instead of directly correlating with *Y*. In other words, $cov(Z, X) \neq 0$ and $cov(Z, u) = 0$. Some of the instruments used for electrification by previous literature include land gradient (Dinkelman, 2011), distance to the electricity line (Barnes et al., 2010; Squires, 2015), and distance to power generating plants (van de Walle et al., 2013). In this thesis, I use the distance to pre-existing transmission lines as an instrument for this analysis, which I believe satisfies both requirements. The distance to previous powerlines does not directly impact education outcomes. However, because grid expansion tends to happen around existing grids, the probability of having electricity access in year *t* is affected by distance to grid in year *t-1*. In other words, the outcome variable of interest, educational attainment, is only affected by distance to previous grid through the probability of electricity access.

4.3 Spatial Lag 2SLS Model

Another issue to consider is the potential presence of spatial autocorrelation. While the traditional OLS framework assumes that observations are independent from one another, data collected for regional science tend to have a very strong spatial component, where observations are spatially dependent and display spatial heterogeneity in the relationships we are modeling (LeSage, 1999). In the context of this study, a household's decision to invest in either education or electricity access may likely be directly affected by or correlated with that of a neighboring household's decision. Further, cluster-specific characteristics, such as the ownership of agricultural land might also be spatially correlated. An OLS specification without accounting for the spatial nature of the data would ignore the spillover effects from neighboring household clusters and thus generate biased estimates of the causal relationship. After running proposed models under a traditional 2SLS framework, I calculate the Moran's I coefficients to measure the spatial autocorrelation of variables of interest. If I observe evidence of spatial dependence, I will utilize a spatial lag model to estimate the magnitude of this dependence. One concern with using a spatially-lagged dependent variable is that this variable itself, spatially weighted education outcome, is also likely endogeneous. If one household cluster's educational attainment is affected by the educational attainment of their neighbor, then the reverse is also likely true. To cope with this issue, I following Kapoor, Keleijian and Prucha (2007) use a spatial 2SLS approach and instrument spatially weighted educational outcome with a second order spatial lag of this variable, essentially capturing the education of the neighbors of neighbors. This model is specified as follows:

First stage 1: $X_{it} = \alpha_0 + \alpha_1 Z_{ij} + \alpha_2 V_{it} + \mu_r + \eta_t + \varepsilon_{rj}$

First stage 2: $WY_{it} = \alpha_3 + \alpha_4 WWY_{it} + \alpha_5 WV_{it} + \mu_r + \eta_t + \sigma_{rj}$

Second stage: $Y_{ij} = \beta_0 + \beta_1 \hat{X}_{it} + \beta_2 W\hat{Y}_{it} + \beta_3 V_{it} + \mu_r + \eta_t + u_{rt}$

Where WY_{it} is a spatially weighted term of the outcome variable of interest, WV_{it} is a vector spatially weighted cluster characteristics V_{it}, and WYY_{it} is the spatially weighted term for WY_{it}. By isolating the spillover effects from spatially correlated variables, this specification should yield useful estimates. Similarly, to explore the mechanism of the effect, I add nighttime rightness in the model:

$$\text{First stage 1:}\ \ X_{it} = \alpha_0 + \alpha_1 Z_{ij} + \alpha_2 V_{it} + \alpha_3 N_{it} + \mu_r + \eta_t + \varepsilon_{rj}$$

$$\text{First stage 2:}\ \ WY_{it} = \alpha_3 + \alpha_4 WWY_{it} + \alpha_5 WV_{it} + \mu_r + \eta_t + \sigma_{rj}$$

$$\text{Second stage:}\ \ Y_{ij} = \beta_0 + \beta_1 \hat{X}_{it} + \beta_2 W\hat{Y}_{it} + \beta_3 V_{it} + \beta_4 N_{it} + \beta_5 \hat{X}_{it} N_{it} + \mu_r + \eta_t + u_{rt}$$

CHAPTER 5: DATA SOURCES AND SUMMARIES

5.1 Introduction

One of the largest obstacles for impact evaluation in development economics is data sources and data quality, as developing countries, especially the least developed countries, tend not to have strict protocols for data collection, maintenance and processing. In most cases, data do not exist. For this study, I use unique GIS techniques to help cope with some of the data issues. Household-level survey results from 2008/09 and 2014 with measures on electricity access, educational attainment as well as spatial coordinates of household cluster are obtained from USAID's Demographic and Health Program. I digitize power supply maps in 2006 and 2012 to calculate household clusters' distances to pre-existing transmission grids.

A highlight of this study is that it explores the mechanism of the impact of electricity access on education using nighttime satellite imagery as a proxy of illumination. One issue that I come across is that two satellite systems with different resolutions have been used to collect this data over the past few years. To normalize these values, I utilize the data from a crossover year, 2013, with machine learning techniques to find the optimal algorithm to relate the two datasets. Then, I use this model to project what 2014 data would have looked like under the 2008/09 system.

5.2 USAID DHS Household-Level Survey Data

5.2.1 Background and Survey Design

The Demographic and Health Surveys (DHS) program was established by the USAID in 1984 as a follow-up to the World Fertility Survey and the Contraceptive Prevalence Survey projects (Rutstein & Rojas, 2006). The basic approach of the DHS program is to collect data that are comparable across countries for a wide range of monitoring and impact evaluation indicators in areas such as population, marriage, health and education. To achieve this, standard model questionnaires have been developed. A country is typically asked to adopt the model questionnaire in its entirety but can add questions of particular interest or delete questions that are irrelevant.

There are two main types of DHS Surveys. Standard DHS surveys have large sample sizes and typically are conducted about once every 5 years, and Interim DHS Surveys focus on the collection of information on key performance monitoring indicators but may not include data for all impact evaluation measures. These nationally representative surveys are conducted between rounds of DHS surveys and have shorter questionnaires than the standard surveys. Since 1984, about 70 countries have completed over 130 nationally representative household-level surveys. Many countries have conducted multiple DHS surveys, both standard and interim, to establish trend data that enable them to gauge progress in their programs.

Kenya started participating in the DHS program in 1989. Since then, 11 DHS surveys have been completed. The Kenya National Bureau of Statistics (KNBS) have been serving as the implementation agency by providing guidance in the overall survey planning, survey tools development, personnel training, data collection, processing and analysis, and results dissemination. Several questionnaires have been used to collect data for specific purposes, including the Household Questionnaire, the Individual Questionnaire and the Women's Questionnaire. For this analysis, I focus on data from household questionnaires, because it not only includes characteristics of all the usual members and visitors in selected households, but more importantly, it contains information about the dwelling itself, such as electricity access, measures of electronic appliances, and wealth assessment.

The selection of samples for KDHS Household Questionnaire follows a two-stage sample design. First, clusters are randomly drawn from the most current master sampling frame, which is developed and maintained by the National Sample Survey and Evaluation Programme (NASSEP V), and are stratified proportional to enumeration areas covered in the most recent census. Households from the selected clusters then serve as the sampling frame for the second stage of selection in which 25 households were selected from each cluster. One feature about Kenya's DHS is that starting in 2008, sampled clusters are geo-referenced. This enables a more useful and relevant analysis that accounts for spatial autocorrelation that may occur due to the nature of the data. However, in order the maintain the confidentiality of households, USAID randomly displace the GPS coordinates for all surveys, such that urban clusters are deviated for 0-2 kilometers and rural clusters for up to 10 kilometers. The displacement is restricted so that the points stay within the DHS survey region, which is county for the Kenya series (Obudho, Muguti, Bore, & Kakinyi, 2015).

For this study, I used the KDHS in 2008/09 and in 2014, the only two standard DHS surveys up to date that are geo referenced. The number of households sampled has increased from 10,000 in 400 clusters in 2008 to about 40,000 households in 1,612 clusters in 2014. Moreover, the response rate has increased from 97.72% to 98.96%. These datasets were obtained from USAID DHS Program after submitting a research proposal.

Summary statistics show that most of the households surveyed own agricultural lands and livestock, have a male household head, and have no electricity access (see Table 3). On average, there are four household members in each household with a mean of about 5.5 years of education, and the age of household head is approximately 44.

Table 3. Summary Statistics of Household-Level Characteristics

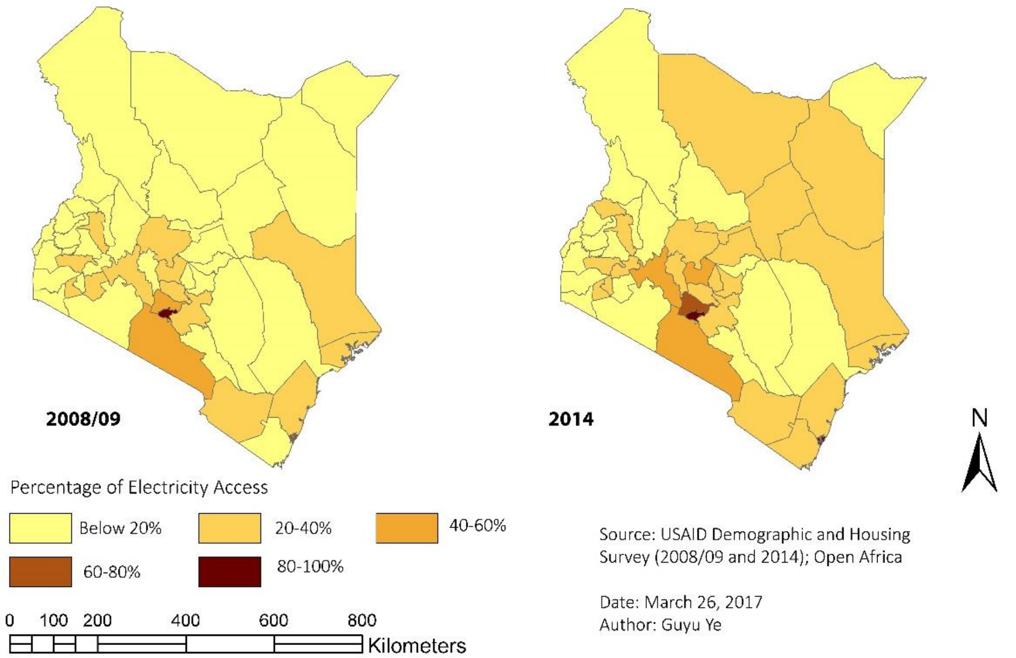
	Min	Max	Mean	Median
Number of Household Members	1.36	8.39	4.22	4.20
Number of Children Under 14	0.05	2.20	0.72	0.67
Electricity Access	0.00	1.00	0.27	0.08
Male Household Head	0.00	1.00	0.66	0.67
Age of Household Head	27.04	69.22	43.96	44.58
Own Ag Land	0.00	1.00	0.64	0.74
Own Livestock	0.00	1.00	0.68	0.76
Average Years of Schooling	0.00	18.92	5.53	5.37
Distance to Grid (m)	2.00	601,436.00	47,506.00	15,500.00
Nighttime Brightness (DMSP)	0.00	62.35	6.74	0.22

5.2.2 Access to Electricity

Survey results show that only 25% of households sampled have electricity access in 2008/09 and 27% in 2014. In addition, a large discrepancy still exists in terms of accessibility. About 56% urban households surveyed indicated that they have access to electricity, while only 9.4% rural households reporting electricity access.

For both years, households with electricity access tend to concentrate in Central, Nairobi, and Coast regions, while households in the northern part of Rift Valley and southern part of the Eastern region tend to have lower levels of electricity access (see Figure 3). From 2009 to 2014, a large increase in electricity access has been observed in the Central region as well as the northern parts of Eastern and Northeastern regions. This increased connectivity potentially resulted from recent projects, such as the Last Mile Connectivity initiative, to rapidly connect households to the grid.

Figure 3. Spatial Distribution of Electricity Access by County (2008/09 and 2014)



Households with electricity access are significantly different from those without in terms of household-specific characteristics (see Table 4). In particular, households with electricity access tend to have less household members, less children under 14, an urban location, and a younger male household head. Average distance to pre-existing transmission lines (calculation explained in Section 5.3) for connected households is close to half the size of those that are not connected, and the average number of years of schooling in households with electricity access is about 4.1 higher than those without.

Table 4. Household Characteristics with vs. without Electricity Access

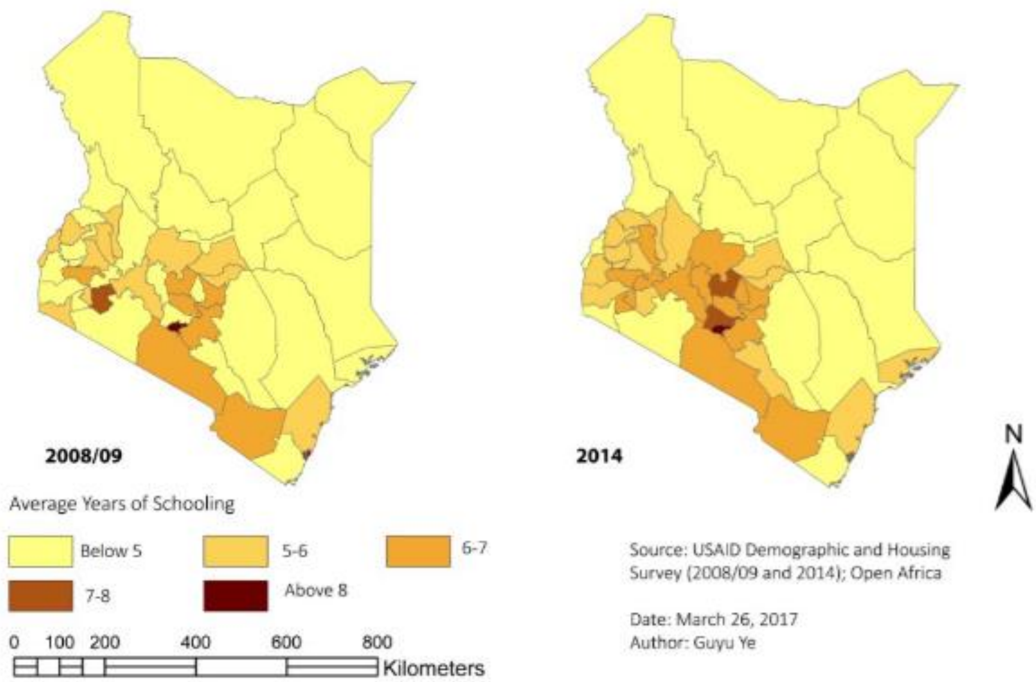
	Access	No Access	t stat	p-value
Number of Household Members	3.40	4.53	45.87	0.00
Number of Children Under 14	0.49	0.81	37.26	0.00
Urban	0.78	0.22	-125.19	0.00
Male Household Head	0.71	0.64	-15.79	0.00
Age of Household Head	39.48	45.71	40.11	0.00
Own Ag Land	0.50	0.70	39.54	0.00
Own Livestock	0.45	0.76	61.69	0.00
Average Years of Schooling	8.51	4.43	-88.37	0.00
Distance to Grid (m)	29,303.67	51,675.23	27.57	0.00

5.2.3 Education

The average number of school years completed is used as a measure of educational attainment in this analysis. The DHS collects data on years of schooling for up to 20 household members for each household. This variable ranges from 0 to 22 in 2008/09, and 0 to 16.2 in 2014. Average household-level years of schooling is 5.52, which is less than the number of years for Kenya's elementary schools.

In terms of spatial distribution, education also seems to display strong spatial autocorrelation (see Figure 4). Regions with the highest numbers of years completed mainly concentrate in Nairobi, Coast, and the central and southern parts of Rift Valley region. West region has observed some increase in average years of schooling from 2008/09 to 2014. The distribution of educational attainment, especially in 2008, is very close to that of electricity access.

Figure 4. Average Years of Schooling by County (2008/09 and 2014)



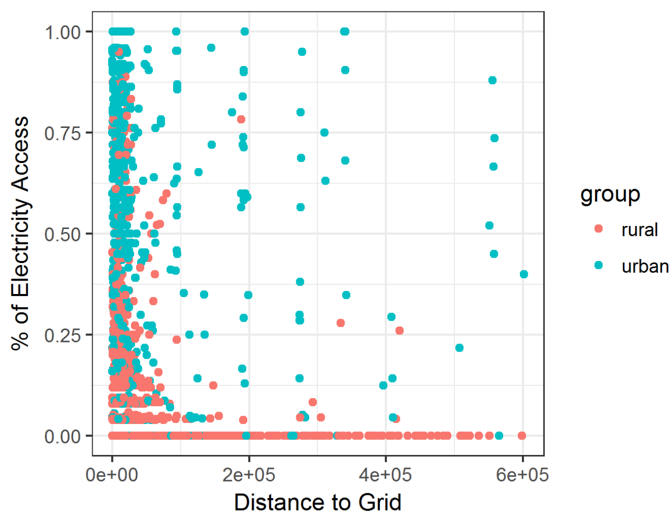
5.3 Transmission Lines

Since the household-level data for 2009 and 2014 are used for this analysis, I obtained the powerline data for 2006 and 2012 to calculate the distance to the grid of previous year (see Appendix). Due to the difficulty in obtaining shapefile or AutoCAD from KPLC on time, I digitized the powerlines from power supply maps of 2006 and 2012 using GIS tools.

To calculate the distance to grid, the first step is to convert the clusters and powerline shapefiles to projected system for accurate distance calculation. The projected coordinate system of Arc 1960 UTM Zone 37S is used for this analysis, as this zone covers the largest portion of Kenya compared to others (Esri, 2012). The IV for 2009 survey data is the distance of 2009 household clusters to the 2006 transmission line, and the IV for 2014 survey data is the distance of 2014 household clusters to the 2012 transmission line. Because geographic coordinates from household-level survey data are only available on the cluster level, I collapsed all the variables of interest by cluster and by year to allow for the heterogeneity of cluster-level characteristics and make all the observations geo-referenced.

Distance to previous grid ranges from 107 meters to 601,436 meters in 2009 and from 2 meters to 55,892 meters in 2014. Since this variable is not normally distributed and that the values are much larger compared to other variables of interest, I use the log transformation of distance for this analysis. A scatterplot of distance and percentage of cluster-level electricity access shows that these two variables are mostly negatively correlated (see Figure 5). In particular, clusters with high level of electricity access tend to be in urban areas, and clusters with low levels of electricity access tend to be in rural areas.

Figure 5. Scatterplot of Distance and Electricity Access



In addition, cluster-level characteristics are very different for clusters that are far away from pre-existing grid (defined in this thesis as more than 25 kilometers away) and those that are not. In particular, clusters that are farther away tend to have more household members, more children under 14, a smaller probability of being in an urban area and lower educational attainment (see Table 5).

Table 5. Cluster-Level Characteristics by Distance

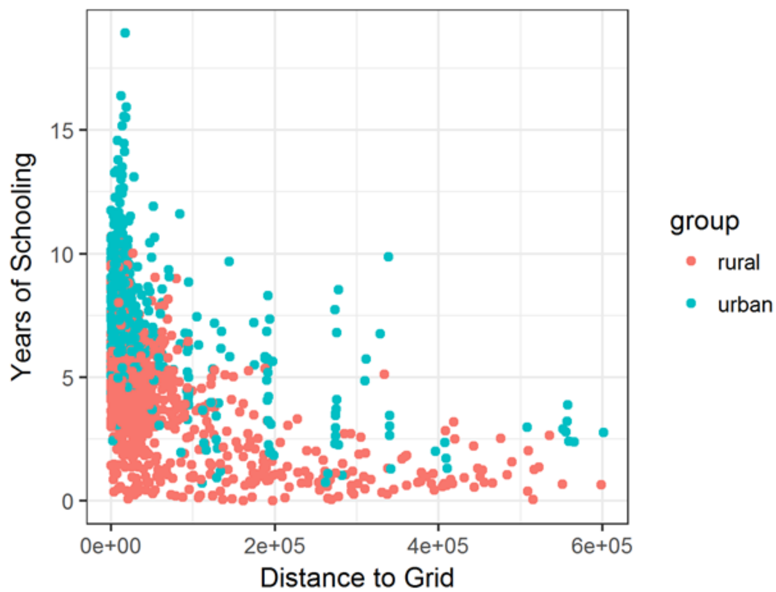
	Less than 25km	More than 25km	t stat	p-value
Number of Household Members	3.95	4.74	-15.21	0.00
Number of Children Under 14	0.64	0.88	-15.25	0.00
Urban	0.45	0.25	9.07	0.00
Male Household Head	0.68	0.62	7.86	0.00
Age of Household Head	43.60	44.64	-3.61	0.00
Own Ag Land	0.69	0.54	9.41	0.00
Own Livestock	0.65	0.73	-6.80	0.00
Average Years of Schooling	6.39	3.92	22.61	0.00
Distance to Grid (m)	9846.06	117857.45	-22.70	0.00

Table 6. Cluster-Level Characteristics by Type of Settlement

	Urban	Rural	t stat	p-value
Number of Household Members	3.59	4.61	20.37	0.00
Number of Children Under 14	0.57	0.82	17.35	0.00
Electricity Access	0.56	0.09	-34.27	0.00
Male Household Head	0.69	0.63	-9.65	0.00
Age of Household Head	39.83	46.46	24.64	0.00
Own Ag Land	0.48	0.73	19.74	0.00
Own Livestock	0.45	0.81	35.97	0.00
Average Years of Schooling	7.31	4.45	-26.72	0.00
Distance to Grid (m)	39,226.29	52,510.27	3.24	0.00

Cluster-level characteristics are also very different by the type of settlement (see Table 6). Urban clusters tend to have less household members, younger male household heads, less ownership of agricultural lands and livestock, higher level of education, and are closer to transmission lines (see Figure 6).

Figure 6. Scatterplot of Distance to Grid and Educational Attainment



5.4 NOAA Nightlight Satellite Imagery

The DMSP series under the Department of Defense (DoD) Program is created using satellites with 101-minute, sun-synchronous near-polar orbit at an altitude of 830km above the surface of the earth (National Oceanic and Atmospheric Administration [NOAA], 2017). The visible and infrared sensors from the Operational Linescan System (OLS) collect images across a 3000lm swath, providing global coverage twice a day. The data from the DMSP satellites are received and used at operational centers continuously. The data are sent to the National Geophysical Data Center's Solar Terrestrial Physics Division Earth Observation Group by the Air Force Weather Agency (AFWA) for creation of an archive (NOAA, 2016). This system has been used from 1992 to 2013, and there are cloud-free composites made using all the available archived DMSP-OLS smooth resolution data for calendar years for free download. The products are 30 arc second grids, spanning -180 to 180 degrees longitude and -65 to 75 degrees latitude.

Since 2012, the Earth Observations Group (EOG) at NOAA has produced a version 1 suite of average radiance composite images using nighttime data from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/ Night Band (DNB) (NOAA, 2017). The DNB data is first filtered to exclude data impacted by stray light, lightning, lunar illumination, and cloud-cover. Then, temporal averaging is done on a monthly and annual basis to produce monthly and annual values. The version 1 products span the global from 75N latitude to 65S. The products are produced in 15 arc-second geographic grids, which is more precise compared to the DMSP system, and are made available in geotiff format as a set of 6 tiles. The titles are cut at the equator and each span 120 degrees of latitude. Kenya spans across Tile 2 (75N/060W) and Tile 5 (00N/060W).

Because the nighttime images for the two years of interest are collected using different systems with different resolutions and frequency, I use machine learning techniques to normalize these values for this analysis. The basic idea is to train a model using some values for a crossover year, 2013 in this case, and then predict the DMSP values for 2014 using VIIRS values. This approach is theoretically very similar to statistical modeling but tend to be able to make more accurate predictions, because it does not have underlying structural assumptions and can be more versatile. To achieve the goal of normalizing these data, I extract the average nighttime values for a 2km-buffer of all household clusters using satellite images collected with both systems in 2013. The relationship between these two sets of data is not linear (see Figure 7), suggesting that data in 2014 should not be extrapolated using the ratio between the ranges of values for two years.

Figure 7. Comparison of Nighttime Satellite System Values

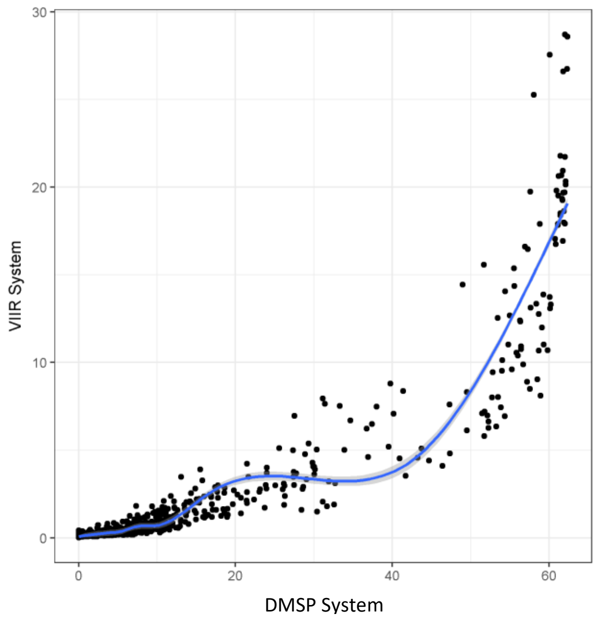
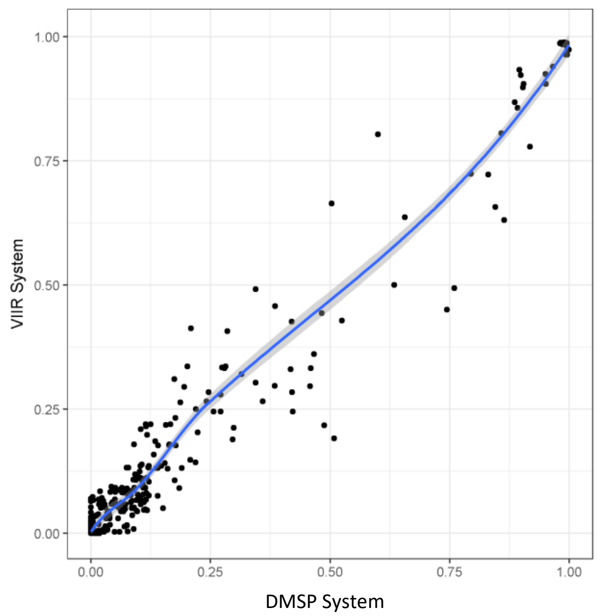


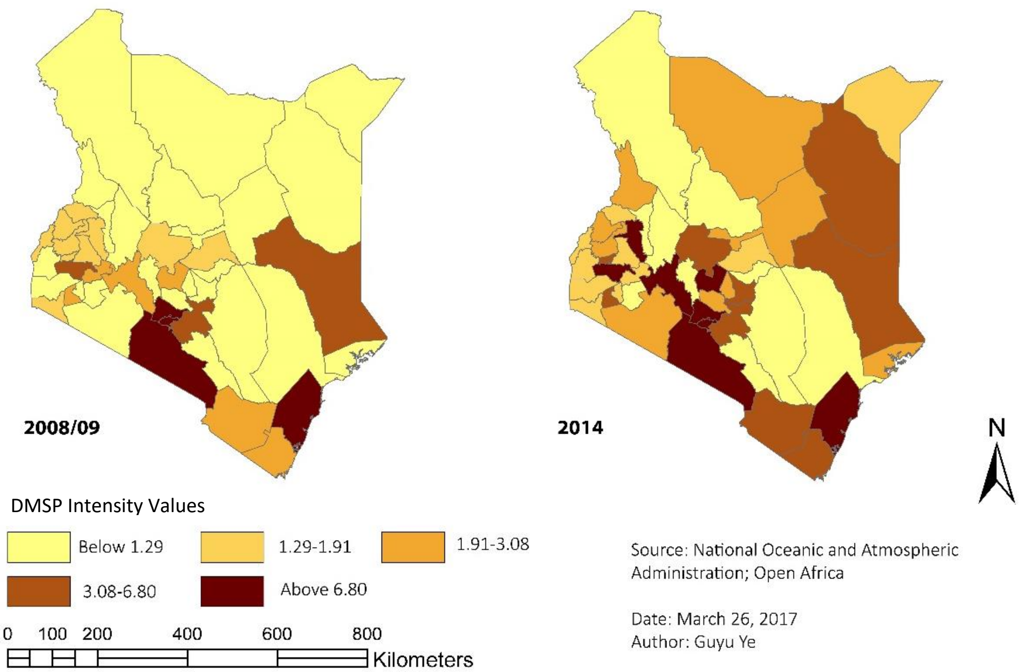
Figure 8. Scatterplot of DMSP Values and Projected VIIR Values



Because data collected from these two systems have different ranges, I first of all normalize the data so that they are bounded between 0 and 1. I shuffle the orders of observations, and divide the dataset into two parts, with 2/3 randomly selected as training data and the rest as test data. Because DMSP has a lower resolution compared to VIIRS, I use VIIRS as a predictor of DMSP. Then, I train the model using a random forest algorithm, which takes random samples of variables and makes predictions based on a decision tree. To evaluate the performance of the model, I test the algorithm on test data, and the predict values and true values are very similar (see Figure 8). An R-Squared of 0.948 furthermore shows that the model is a very good fit and that 94.8% of the variations is predicted using this model. Therefore, I am confident that normalized values through this approach are fairly accurate.

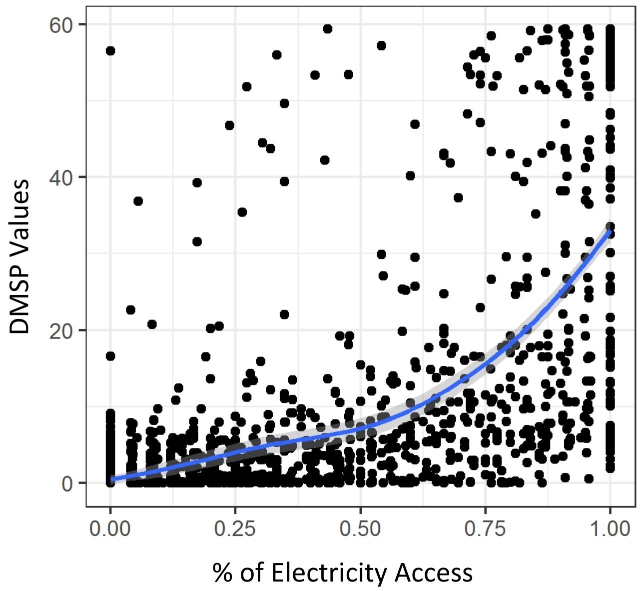
With this validated model, I predict the DMSP values for the nighttime light values extracted for 2014 clusters. Even though DMSP intensity ranges from 0 to 63, most of Kenya is really dark with an DMSP value of less than 6.8. Nighttime becomes brighter in 2014 compared to the previous survey year, especially in Eastern, Northeastern, Coast and Central regions, which is consistent with the spatial distribution of electricity access (see Figure 9).

Figure 9. Nighttime Brightness by County (2008/09 and 2014)



A scatterplot of the percentage of household cluster-level electricity access and nighttime light values show that in general, a higher percentage of electricity access results in higher level of brightness in the area (see Figure 10). Clusters with low electricity access but high level of nighttime brightness are mostly located in urban areas; they might be urban slums that are not connected to the grid but still get lights from surrounding area. For households in the lower right corner (where the level of electricity access is high while DMSP level is low), it could potentially be that they are using off-grid renewable energy equipment that do not generate power at night.

Figure 10. Scatterplot of Electricity Access and Nighttime Brightness



CHAPTER 6: RESULTS

6.1 Introduction

The goal of this thesis is to examine the impact of electricity access on educational attainment and explore the potential mechanism of such an impact. This chapter presents the results from regression models specified in Section 4, and discusses their implications.

6.2 Traditional 2SLS Models

I use a two-stage least squares framework to address the potential endogeneity of electricity access. The first stage predicts the percentage of households within a household cluster that have electricity access using the distance to transmission lines in the previous year as an instrument. Following Khandker et al (2012), I include characteristics that are specific to household clusters, such as the average age of household head and number of household members, in the regression model. Squires' paper (2015) suggests that increased electrification leads to reduced attendance, which is potentially a result of increased child labor. To control for this potential effect, I add the ownership of agricultural land and livestock into the regression. In a later specification, I include the log transformation of nighttime light index to explore the mechanism of the impact of electricity access. Four models are specified for the first stage: first, without any fixed effects, second, with year fixed effects only, third, with year fixed effects and region/ province fixed effects, and fourth, with year fixed effect and county fixed effects to capture unobserved regional characteristics at a finer level.

Results from the first stage suggest that the log transformation of distance to the transmission grid in the prior year is a strong instrument and is negatively correlated with household-cluster level electrification rate across all specifications (see Table 5). In particular, a one percent increase in distance to pre-existing transmission line leads to about 1-2 percentage points decrease in cluster-level electrification rate, holding everything else constant. All other independent variables perform strongly as well, and most of them have expected coefficient signs. Being in an urban cluster and having a male as household head both contribute to a higher cluster-level electrification rate. On the contrary, household clusters with older household heads, agricultural land and livestock, and more children under 14 are less likely to be electrified, holding everything else equal.

Table 7. First Stage Results under Traditional 2SLS Framework

	(1)	(2)	(3)	(4)
log(Dist)	-0.019***	-0.018***	-0.018***	-0.009**
	(0.003)	(0.003)	(0.004)	(0.004)
Children Under 14	-0.183***	-0.181***	-0.155***	-0.151***
	(0.024)	(0.024)	(0.024)	(0.024)
Male Household Head	0.187***	0.188***	0.150***	0.090***
	(0.034)	(0.034)	(0.033)	(0.035)
Age of Household Head	-0.006***	-0.006***	-0.006***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Own Ag Land	-0.076***	-0.075***	-0.042*	0.026
	(0.020)	(0.021)	(0.023)	(0.029)
Own Livestock	-0.440***	-0.443***	-0.431***	-0.445***
	(0.029)	(0.030)	(0.030)	(0.030)
Urban	0.172***	0.171***	0.160***	0.157***
	(0.013)	(0.013)	(0.013)	(0.013)
Number of Household Members	-0.017**	-0.017**	-0.012	-0.007
	(0.008)	(0.008)	(0.008)	(0.008)
Constant	1.071***			
	(0.060)			
Year FE	N	Y	Y	Y
Region FE	N	N	Y	N
County FE	N	N	N	Y
Observations	1959	1959	1959	1959
R2	0.657	0.789	0.806	0.825
Adjusted R2	0.655	0.788	0.804	0.82
Residual Std. Error	0.201	0.201	0.193	0.186
F Statistic	466.614***	730.523***	474.832***	159.835***

In the second stage, I regress the dependent variable on all exogenous regressors and the predicted value for independent variable. Standard errors obtained in the second stage are not valid because they do not take into account that the prediction of dependent variable is an estimate itself. After correcting the standard errors, the final output is as follows:

Table 8. Second Stage Results under Traditional 2SLS Framework

	(1)	(2)	(3)	(4)
Electricity	10.616***	12.154***	10.900***	7.835**
	(1.734)	(1.983)	(1.903)	(3.132)
Children Under 14	-0.664*	-0.604	-1.165***	-1.542***
	(0.396)	(0.444)	(0.387)	(0.512)
Male Household Head	0.164	-0.245	0.301	0.712*
	(0.472)	(0.537)	(0.444)	(0.384)
Age of Household Head	0.005	0.008	0.009	-0.001
	(0.014)	(0.016)	(0.016)	(0.023)
Own Ag Land	3.027***	3.059***	2.367***	0.858***
	(0.180)	(0.201)	(0.215)	(0.236)
Own Livestock	1.920**	2.914***	2.268**	1.156
	(0.837)	(0.967)	(0.896)	(1.425)
Urban	-0.905***	-1.089***	-0.882***	-0.464
	(0.331)	(0.375)	(0.337)	(0.507)
Number of Household Members	-0.082	-0.003	0.007	-0.044
	(0.080)	(0.091)	(0.079)	(0.063)
Constant	0.277			
	(1.631)			
Year FE	N	Y	Y	Y
Region FE	N	N	Y	N
County FE	N	N	N	Y
Observations	1959	1959	1959	1959
R2	0.471	0.881	0.907	0.95
Adjusted R2	0.469	0.881	0.906	0.949
Residual Std. Error	1.882	2.107	1.871	1.380

Results show that having electricity access significantly contributes to improved educational attainment. Because the deviations for household clusters are bounded by counties, the last specification with year and county fixed effects is the preferred model. This model suggests that if an unelectrified household cluster gets fully connected to the grid, then the average years of schooling for that cluster will increase by 7.84 years. Factors that also positively contribute to educational attainment include having a male household head and ownership of agricultural land, while having more children under the age of 14 negatively impacts average years of schooling for the cluster.

To explore whether electricity access on its own explains the higher level of education, or whether nighttime illumination increases that marginal effect, I include normalized DMSP values in the models (see Table 8). Results from the first stage show that nighttime brightness has a significant and important impact on electricity access. In addition, the log transformation of distance to pre-existing grid has a negatively significant impact on electricity access for three out of the four specifications. Specifically, a one percent increase in distance to transmission line decreases cluster-level electricity access by about one percentage point, holding everything else constant. Most of the other independent variables perform strongly as well, and most of them have expected coefficient signs. Being in an urban cluster and having a male as household head both contribute to a higher cluster-level electrification rate. On the contrary, households with older household heads, agricultural land and livestock, and more children under 14 are less likely to be electrified, holding everything else equal.

Results from the second stage suggest that if gaining electricity access increases cluster-level average years of schooling by 3.4 years, when county and year fixed effects are controlled for (see Table 9). The coefficients for nighttime light index are largely negative and insignificant for all specifications, implying that illumination or nighttime light does not have an additional effect on educational attainment above having electricity access.

Table 9. First Stage Results under Traditional 2SLS Framework with Nighttime Light

	(1)	(2)	(3)	(4)
log(Distance)	-0.012***	-0.011***	-0.011***	-0.006
	(0.003)	(0.003)	(0.003)	(0.004)
log(Nighttime Light Index)	0.033***	0.033***	0.029***	0.024***
	(0.002)	(0.002)	(0.002)	(0.002)
Children Under 14	-0.155***	-0.148***	-0.127***	-0.124***
	(0.023)	(0.023)	(0.023)	(0.023)
Male Household Head	0.128***	0.131***	0.110***	0.058*
	(0.033)	(0.033)	(0.032)	(0.034)
Age of Household Head	-0.006***	-0.006***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Own Ag Land	-0.040**	-0.037*	-0.010	0.037
	(0.020)	(0.020)	(0.022)	(0.028)
Own Livestock	-0.350***	-0.357***	-0.370***	-0.391***
	(0.029)	(0.029)	(0.029)	(0.030)
Urban	0.118***	0.115***	0.114***	0.122***
	(0.013)	(0.013)	(0.013)	(0.013)
Number of Household Members	0.012*	-0.013*	-0.010	-0.008
	(0.007)	(0.007)	(0.007)	(0.008)
Constant	0.962***			
	(0.058)			
Year FE	N	Y	Y	Y
Region FE	N	N	Y	N
County FE	N	N	N	Y
Observations	1959	1959	1959	1959
R2	0.689	0.809	0.82	0.833
Adjusted R2	0.687	0.808	0.818	0.828
Residual Std. Error	0.191	0.191	0.186	0.181
F Statistic	478.983***	751.232***	491.404***	166.857***

Table 10. Second Stage Results under Traditional 2SLS Framework with Nighttime Light

	(1)	(2)	(3)	(4)
Electricity	3.331***	3.436***	3.409***	3.369***
	(0.161)	(0.158)	(0.156)	(0.150)
log(Nighttime Light Index)	0.002	-0.005	-0.004	-0.005
	(0.018)	(0.017)	(0.017)	(0.017)
Children Under 14	-2.010***	-2.190***	-2.391***	-2.237***
	(0.149)	(0.147)	(0.148)	(0.142)
Male Household Head	1.535***	1.465***	1.516***	1.124***
	(0.210)	(0.206)	(0.203)	(0.208)
Age of Household Head	-0.040***	-0.044***	-0.040***	-0.033***
	(0.006)	(0.006)	(0.006)	(0.006)
Own Ag Land	2.855***	2.838***	2.283***	0.993***
	(0.113)	(0.111)	(0.135)	(0.171)
Own Livestock	-1.254***	-1.017***	-1.072***	-0.871***
	(0.190)	(0.188)	(0.193)	(0.190)
Urban	0.310***	0.391***	0.354***	0.259***
	(0.085)	(0.084)	(0.083)	(0.080)
Number of Household Members	-0.235***	-0.199***	-0.096**	-0.078*
	(0.046)	(0.045)	(0.047)	(0.047)
Electricity*log(DMSP)	0.109***	0.088**	0.008	-0.01
	(0.042)	(0.041)	(0.044)	(0.044)
Constant	6.659***			
	(0.337)			
Year FE	N	Y	Y	Y
Region FE	N	N	Y	N
County FE	N	N	N	Y
Observations	1959	1959	1959	1959
R2	0.774	0.961	0.963	0.968
Adjusted R2	0.773	0.961	0.963	0.968
Residual Std. Error	1.230	1.206	1.176	1.100

A weak IV can lead to incorrect size of tests of significance and confidence intervals, and may not be consistent. To check for the strength and validity of my proposed IV, I run F-test, Wu-Hausman test and Sargan test for all the models under the traditional 2SLS framework. Because there is only one instrument in these two sets of specifications, I include both distance to grid and the log transformation of distance to grid for the Sargan test to test for the potential overidentification problem.

Table 11. Tests for IV under Traditional 2SLS Framework

Without lights	(1)	(2)	(3)	(4)
F Test	29.19***	27.95***	25.91***	5.64**
Wu-Hausman	39.70***	58.95***	40.54***	3.26*
Sargan	15.14***	12.79***	7.26*	5.38
With lights	(1)	(2)	(3)	(4)
F Test	44.25***	63.93***	30.60***	28.79***
Wu-Hausman	13.66***	24.41***	24.92***	2.93*
Sargan	17.78***	15.90***	2.15	0.56

Results show that overall, the proposed instrument performs strongly (see Table 10). Wu-Hausman test results are significantly across all specifications, implying that the OLS estimates are significantly different from the IV estimates. For the Sargan test, while the specifications without location fixed effects imply that the distance to grid may not be exogenous, the last model in the first set of specification and the last two models in the second set of specification do not find significant results, indicating that using distance as instrument is valid for these specifications.

6.3 Spatial 2SLS Models

With the spatial nature of variables of interest, I explore how the estimates are affected when spatial correlation is taken into consideration. To achieve this goal, I first need to check whether statistically significant spatial autocorrelation exists; if it does, then I need to determine which spatial weight matrix to use. The most common way to determine the most relevant spatial weight matrix is to find the one that maximizes the Moran's I coefficient, which is used to test the spatial dependence or autocorrelation between observations or location (Lee & Wong, 2001). The three types of spatial weight matrices that are commonly used by regional economists, contiguity-based, distance-based and n nearest neighbor, all have their strengths and weaknesses.

For contiguity-based matrices in particular, island effect might exist, meaning that a neighbor selected under this scheme can be comparatively far away from the cluster of interest and have different spatial characteristics. This type of matrices is usually not suited for spatial data that are not contiguous, thus is not considered for this analysis.

Since the household clusters selected for the 2008/09 survey are not the same as the ones for 2014, this dataset is not a panel data in terms of cluster. Therefore, I create two sets of weight matrices, one for the dataset for the first wave and one for the second one. Using the weight matrix manager in GeoDa, I find that the weight matrices that optimize Moran's Is for both year are k-5, or 5 nearest neighbors.

Table 12. Moran's I Coefficients

	2008	2014
Avg Edu	0.6157***	0.4251***
Number of Household Members	0.4818***	0.4791***
Urban	0.3696***	0.4005***
Male HH Head	0.2109***	0.2938***
Age of HH Head	0.2632***	0.3675***
Children Under 14	0.493***	0.4158***
Own Ag Land	0.5667***	0.6714***
Own Livestock	0.4322***	0.4913***

Based on the results, I can reject the null hypothesis of zero spatial autocorrelation for all the variables of interest at 5% significance level (Table 11). All coefficients examined are significantly positive, suggesting that the characteristics of one household cluster positively contributes to the attributes of its neighboring clusters. Based on this spatial weight matrix, I generate spatially weighted values for all these variables.

Spatially weighted educational attainment is also potentially endogenous. To cope with this issue, I add a second first stage to instrument this variable with doubly spatially weighted educational attainment, as specified in Section 4.

Table 13. Second First Stage Results under Spatial 2SLS Framework

	(1)	(2)	(3)	(4)
ww_Education	0.453***	0.395***	0.402***	0.267***
	(0.020)	(0.021)	(0.020)	(0.020)
w_Children Under 14	-1.613***	-2.356***	-3.227***	-3.022***
	(0.301)	(0.303)	(0.306)	(0.311)
w_Male Household Head	2.579***	2.396***	2.223***	1.914***
	(0.413)	(0.403)	(0.387)	(0.440)
w_Age of Household Head	-0.050***	-0.075***	-0.049***	-0.029**
	(0.012)	(0.012)	(0.012)	(0.013)
w_Own Ag Land	3.386***	3.613***	2.678***	1.766***
	(0.188)	(0.185)	(0.212)	(0.298)
w_Own Livestock	-3.296***	-2.566***	-3.184***	-3.455***
	(0.325)	(0.325)	(0.331)	(0.335)
w_Urban	0.885***	1.275***	0.834***	0.694***
	(0.161)	(0.162)	(0.158)	(0.158)
w_Number of Household Members	-0.315***	-0.191**	0.138	-0.144
	(0.087)	(0.085)	(0.089)	(0.099)
Constant	6.339***			
	(0.693)			
Year FE	N	Y	Y	Y
Region FE	N	N	Y	N
County FE	N	N	N	Y
Observations	1959	1959	1959	1959
R2	0.737	0.968	0.972	0.977
Adjusted R2	0.736	0.968	0.972	0.976
Residual Std. Error	1.390	1.357	1.269	1.168
F Statistic	682.700***	5885.783***	3973.095***	1431.461***

All the weighted values of cluster-specific characteristics included in the model and the doubly weighted average educational attainment significantly contribute to the first order of spatially weighted educational attainment (see Table 12). In particular, one year increase in the number of years of schooling for a neighboring cluster is affected by its neighboring clusters by 0.27 years, when year and county fixed effects are controlled for. Second stage results also show that electricity contributes positively to educational attainment (see Table 13). In particular, gaining electricity access for the whole household cluster would increase the cluster-level average years of schooling by over 10 years.

Table 14. Second Stage Results under Spatial 2SLS Framework

	(1)	(2)	(3)	(4)
Electricity	1.038	1.641*	5.046***	10.046***
	(0.951)	(0.876)	(1.135)	(2.072)
w_Education	0.267***	0.236***	0.102**	-0.133*
	(0.040)	(0.037)	(0.045)	(0.077)
Children Under 14	-2.279***	-2.303***	-1.993***	-1.314***
	(0.210)	(0.190)	(0.200)	(0.338)
Male Household Head	1.371***	1.282***	1.028***	0.575
	(0.228)	(0.221)	(0.228)	(0.353)
Age of Household Head	-0.049***	-0.049***	-0.027***	0.015
	(0.008)	(0.007)	(0.009)	(0.018)
Own Ag Land	1.644***	1.782***	1.994***	0.921***
	(0.221)	(0.207)	(0.201)	(0.264)
Own Livestock	-1.630***	-1.301***	-0.117	1.929**
	(0.371)	(0.360)	(0.457)	(0.863)
Urban	0.593***	0.554***	0.011	-0.739**
	(0.162)	(0.147)	(0.177)	(0.312)
Number of Household Members	-0.165***	-0.144***	-0.062	-0.031
	(0.046)	(0.044)	(0.049)	(0.073)
Constant	6.814***			
	(0.665)			
Year FE	N	Y	Y	Y
Region FE	N	N	Y	N
County FE	N	N	N	Y
Observations	1959	1959	1959	1959
R2	0.783	0.964	0.962	0.926
Adjusted R2	0.782	0.964	0.962	0.924
Residual Std. Error	1.205	1.164	1.187	1.682

To explore the mechanism of the effect of electricity access, I again include nighttime brightness in the first stage, which is the same as table 7. After correcting for the errors, results for the second stage are consistent with the set of specifications without spatial terms (see Table 14). Holding everything else constant, if a household cluster goes from having no electricity access to fully electrified, then the average years of schooling increase by about 3.3 years. Educational attainment in one household cluster is also affected by its neighboring household cluster through spillover effect, with the magnitude of about 0.17 when year and county fixed effects are controlled for. This result suggests that one year increase in years of schooling in one household cluster leads to 0.17 years increase of years of schooling for its neighboring household cluster. This effect would be through many channels, such as learning from each other and increasing motivations.

Table 15. Second Stage Results under Spatial 2SLS Framework with Nighttime Light

	(1)	(2)	(3)	(4)
Electricity	3.043***	3.127***	3.143***	3.285***
	(0.151)	(0.151)	(0.151)	(0.149)
log(Nighttime Light Index)	-0.025	-0.028*	-0.029*	-0.022
	(0.017)	(0.016)	(0.017)	(0.017)
w_Education	0.288***	0.269***	0.262***	0.172***
	(0.022)	(0.023)	(0.024)	(0.034)
Children Under 14	-1.922***	-2.040***	-2.147***	-2.134***
	(0.139)	(0.139)	(0.143)	(0.141)
Male Household Head	1.004***	0.994***	1.067***	1.077***
	(0.200)	(0.198)	(0.199)	(0.205)
Age of Household Head	-0.037***	-0.040***	-0.039***	-0.035***
	(0.006)	(0.006)	(0.006)	(0.006)
Own Ag Land	1.584***	1.654***	1.447***	0.816***
	(0.145)	(0.145)	(0.150)	(0.172)
Own Livestock	-0.793***	-0.673***	-0.754***	-0.703***
	(0.181)	(0.179)	(0.187)	(0.190)
Urban	0.301***	0.352***	0.306***	0.221***
	(0.080)	(0.079)	(0.080)	(0.079)
Number of Household Members	-0.120***	-0.105**	-0.076*	-0.073
	(0.044)	(0.043)	(0.045)	(0.046)
Electricity*log(DMSP)	-0.053	-0.057	-0.080*	-0.056
	(0.041)	(0.040)	(0.043)	(0.044)
Constant	4.962***			
	(0.341)			
Year FE	N	Y	Y	Y
Region FE	N	N	Y	N
County FE	N	N	N	Y
Observations	1959	1959	1959	1959
R2	0.804	0.966	0.966	0.97
Adjusted R2	0.803	0.965	0.966	0.969
Residual Std. Error	1.146	1.134	1.124	1.080

Tests for the IVs show that the instruments selected for these two sets of specifications are most useful for the last model when nighttime light is not included and the first and third models when nighttime light is included.

Table 16. Tests for IV under Spatial 2SLS Framework

Without lights		(1)	(2)	(3)	(4)
F Test	(w_AvgEdu)	397.30***	405.33***	370.77***	301.30***
	(Electricity)	22.70***	22.63***	17.73***	11.90***
Wu-Hausman		6.39***	4.25**	1.90	11.56***
Sargan		95.84***	89.24***	82.54***	1.03
With Lights		(1)	(2)	(3)	(4)
F Test	(w_AvgEdu)	457.97***	448.51***	323.98***	135.85***
	(Electricity)	94.56***	96.00***	71.90***	54.63***
Wu-Hausman		3.97**	3.28*	5.63**	0.78
Sargan		0.32	4.14**	2.64	0.18

CHAPTER 7: CONCLUSION AND DISCUSSION

7.1. Introduction

Economics is a subject that is concerned with allocation of scarce resources. Developing countries, in particular, are struggling with achieving economic goals with limited resources. This thesis uses microdata and satellite imagery to study the impact of electricity access on education, which is a driver of economic growth. The findings can be potentially beneficial to policymakers in Kenya and other countries that are facing similar development challenges.

7.2. Important Results and Implications

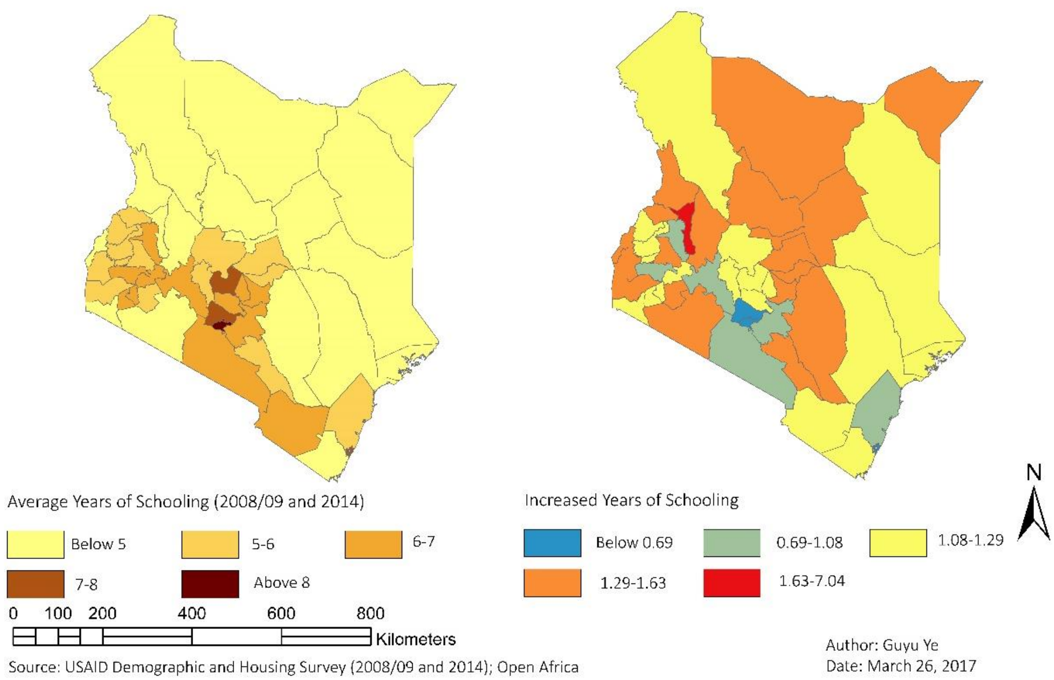
Summary statistics show that households that have electricity access have 4.08 more years of schooling on average compared to the ones that do not. In addition, on average, urban clusters have 2.86 more years of schooling compared to rural ones.

Based on the results under a traditional 2SLS framework, electricity access has a large impact on educational attainment in Kenya. In terms of mechanism, increased nighttime light intensity does not contribute to increased average years of schooling on top of overall electricity access. This finding suggests that household cluster's educational attainment benefits from having electricity access in general, not just at night. This result suggests that the effect of electricity access on education might have been carried out through other channels than illumination, such as labor saving.

Using the last model specified under the traditional 2SLS framework, I simulated four scenarios to study how investment in electrification will be delivered across the nation (see Figure 10). The first three scenarios simulate increases of cluster-level electrification rate by 10 percentage points, 20 percentage points and 50 percentage points respectively, and the last one simulates a mixed level of increase with 10 percentage point increase for urban clusters and 50 percentage point increase for rural clusters.

Figure 10 shows the average years of schooling based on DHS survey results compared with the changes of educational attainment after the fourth scenario. It is obvious to see that Eastern, Western and part of Northeastern Province will have the highest return to investment in terms of educational attainment. On the contrary, Central, Nairobi and the southern part of Rift Valley have the smallest increase in educational attainment in this scenario. The area in Eastern region from the east side of Central region has both high level of educational attainment and large return to invest.

Figure 11. Simulated Results of Increase in Educational Attainment



7.3 Limitations of the Analysis and Future Work

Although results seem very strong for most of the models, some limitations exist in the analysis, in terms of data quality as well as availability. First of all, geographic coordinates from USAID's DHS program are deviated by a random distance to protect the confidentiality of residents (USAID, n.d.). Even though deviations are kept within administrative districts, calculated distance from geo-referenced household clusters to pre-existing transmission line is biased. In terms of nighttime brightness, I extract the average values within a 2-kilometer buffer for urban clusters and within a 5-kilometer buffer for rural clusters to cope with the displacement. Yet, this simple calibration still does not produce the most accurate data.

In addition, I have to manually digitize powerlines from old maps due to the difficulty in obtaining autoCAD or Shapefile from KPLC. ArcGIS provides useful tools to convert features on a paper map into digital format, which enable me to conduct this analysis. However, the quality of the paper maps are not the best. Colors were not chosen very carefully and sometimes can be difficult to distinguish, which leads to potential error in digitizing. Furthermore, most of Kenya's power supply maps only show the transmission network, which is not the best proxy for the actual distribution network. So, with improved quality and availability of spatial data, the results of this study will be more useful and relevant.

Moreover, average years of schooling may not be the best measure of education outcome, because it does not account for learning that takes place outside of school, such as with families, via the internet and through work. In addition, students who spend the longest time in school may not be the students with highest grades; it might just be that their families are wealthy enough to keep them in school. A more direct measure for education outcome would be test scores. Moreover, cluster-level education outcomes usually do not change very quickly, because it takes time for the effect of electricity to take place; a more straightforward outcome to look at is educational attainment among school-age children, which is not provided in the dataset that I use.

Future work can be expanded to other social infrastructure, such health outcomes. Electricity is very crucial to health services, because refrigeration is crucial to the quality of vaccine and operations, and that electricity access can reduce the inhalation of smokes and fumes produced from burning traditional fuels. Similar to education, health also plays an important role in development, because improved health can increase productivity and thus sustain growth. Therefore, studies on how electricity access affects health will also provide policymakers with valuable information on how to efficiently plan for the country's energy future.

Furthermore, simulation analysis can be conducted on a more regional level with considerations of renewable energy potentials and the possibility of a decentralized electric grid. Kenya, as one of the fast-developing countries in the region, has been observing increasing adoption of off-grid solar panel, wind turbine, and geothermal. Currently, electricity is distributed to customers through a centralized national grid, which is costly to reach the poorest or geographically challenging areas. Therefore, simulating scenarios that take into consideration of possible decentralized renewable energy projects can help policymakers to explore more options on how to increase electricity access and how to improve educational attainment.

BIBLIOGRAPHY

Attigah, B., & Mayer-Tasch, L. (2013). The Impact of Electricity Access on Economic Development: A Literature Review. *PRODUSE*. Retrieved from http://www.produse.org/imglib/downloads/PRODUSE_study/PRODUSE%20Study_Literature%20Review.pdf

Barnes, D., Khandker, S., & Samad, H. (2010). Energy Access, Efficiency, and Poverty: How Many Households are Energy Poor in Bangladesh? *Policy Research Working Paper* 5332. Retrieved from http://documents.worldbank.org/curated/en/720961467998255412/pdf/WPS5332.pdf

Barnes, D., & Binswanger, H. (1986), Impact of Rural Electrification and Infrastructure on Agricultural Changes. *Economic and Political Weekly*, *21*(1), 26-34.

Barron, M., & Torero, M. (2014, December). Short Term Effects of Household Electrification: Experimental Evidence from Northern El Salvador. *MPRA Paper* No. 63782. Retrieved from https://mpra.ub.uni-muenchen.de/63782/1/MPRA_paper_63782.pdf

Barron, M., & Torero, M. (2017, March). Household Electrification and Indoor Air Pollution. Retrieved from https://www.ocf.berkeley.edu/~manuelb/Research/IAP/IAP-Mar2017.pdf

Bensch, G., Kluve, J., & Peters, J. (2011, September). *Impacts of Rural Electrification in Rwanda*. Ruhr Economic Papers #284. Retrieved April, 2017, from https://pdfs.semanticscholar.org/476f/9106f8170b544c493991e18f1a1edb9b4053.pdf

Berger, N, & Fisher, P. (2013, August). *A Well-Educated Workforce is Key to State Prosperity* (Rep.). Retrieved January, 2017, from http://www.epi.org/publication/states-education-productivity-growth-foundations/

Biryabarema, E. (2016, August). Kenya Hit by Electricity Outage, Nairobi without Power. *Reuters.* http://www.reuters.com/article/us-kenya-power-outage-idUSKCN10H05N?il=0

Blalock, G., & Veloso, F. (2007). Imports, Productivity Growth, and Supply Chain Learning, *World Development*, *35*(7), 1134-1151.

Burlig, F., & Preonas, L. (2016, October). Out of Darkness and Into the Light? Development Effects of Rural Electrification. *Energy Institute at Haas Working Paper* 268. Retrieved from https://ei.haas.berkeley.edu/research/papers/WP268.pdf

Cabraal, R., Barnes, D., & Agarwal, S. (2005). *Productive Uses of Energy for Rural Development.* Annual Review of Environment and Resources, *30*, 117-144.

Chimbaka, B. (2016). Electricity Sector Market Reforms: Getting it Right in Developing Countries - SADC. *Energy Regulation Board.* Retrieved from https://static1.squarespace.com/static/52246331e4b0a46e5f1b8ce5/t/56f138b037013b8f45f8c9bb/1458649268128/Besa+Chimbaka_Paper+on+power+sector+market+reforms.pdf

Coleman, D.Y. (2013). Kenya. *Kenya country review* [online], 1–2. Retrieved December, 2016, from: http://www.countrywatch.com

Crousillat, E., Hamilton, R., & Antmann, P. (2010, June). *Addressing the Electricity Access Gap*. Retrieved from http://siteresources.worldbank.org/EXTESC/Resources/Addressing_the_Electricity_Access_Gap.pdf

Decker, P. T., Rice, J. K., Moore, M. T., & Rollefston, M. R. (1997, March). *Education and the Economy: An Indicators Report* (Rep.). Retrieved February, 2017, from National Center for Education Statistics website: https://nces.ed.gov/pubs97/97269.pdf

Dinkelman, T. (2010, August). The Effects of Rural Electrification on Employment: New Evidence from South Africa. *Working Paper.* Retrieved from: https://www.princeton.edu/rpds/papers/dinkelman_electricity_0810.pdf

Doll, C. (2008). *CIESIN Thematic Guide to Night-time Light Remote Sensing and Its Applications*. Center for International Earth Science Information Network (CIESIN), Columbia University.

Electricity Regulatory Commission (2016). *About Us.* Retrieved from http://www.erc.go.ke

Elvidge, C., Baugh, K., Anderson, S., Sutton, P., & Ghosh, T. (2012). The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite data. *Social Geography*, *7*, 23-25. Retrieved from http://www.soc-geogr.net/7/23/2012/sg-7-23-2012.pdf

Esri. (2012). *ArcGIS 10.1 Projected Coordinate System Tables*. Retrieved from http://resources.arcgis.com/en/help/main/10.1/018z/pdf/projected_coordinate_systems.pdf

Government of Kenya. (2003, June). *Economic Recovery Strategy for Wealth and Employment Creation 2003-2007*. Retrieved from http://siteresources.worldbank.org/KENYAEXTN/Resources/ERS.pdf

Government of Kenya. (2006). *The Energy Act*. Retrieved January, 2017, from http://www.erc.go.ke/images/Regulations/energy.pdf

Hanushek, E., & Woessmann, L. (2012, July). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. *J Econ Growth*, *17*, 267–321. Retrieved from http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%2BWoessmann%202012%20JEconGrowth%2017(4).pdf

Harrison, K., Scott, A., & Hogarth, R. (2016). *Accelerating Access to Electricity in Africa with Off-Grid Solar: The impact of solar household solutions*. Overseas Development Institute. Retrieved from https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/10229.pdf

International Energy Agency (2015). *World Energy Outlook 2015*. Retrieved April, 2016, from http://www.worldenergyoutlook.org/resources/energydevelopment/energyaccessdatabase/#d.en.8609

Institute of Economic Affairs (2015). *Situational Analysis of Energy Industry, Policy and Strategy for Kenya*. Retrieved from http://dspace.africaportal.org/jspui/bitstream/123456789/34928/1/Situational-Analysis-of-Energy-Industry-Policy-and--Strategy-for-Kenya%20(1).pdf?1

Kanagawa, M., & Nakata, T. (2008, June). Assessment of Access to Electricity and the Socio-Economic Impacts in Rural Areas of Developing Countries. *Energy Policy 36*(6): 2016-2029

Kapoor, M., Kelejian, H, & Prucha, I. (2007). Panel Data Model with Spatially Correlated Error Components. *Journal of Econometrics 140* (1), 97–130.

Kenya Electricity Transmission Company (2017). *Frequently Asked Questions*. http://www.ketraco.co.ke/environment/faqs.html

Kenya Institute for Public Policy Research and Analysis (KIPPRA) (2010). A Comprehensive Study and Analysis on Energy Consumption Patterns in Kenya. *For the Energy Regulatory Commission (ERC)*. Retrieved from https://cofek.co.ke/ERCStudy_ExecSummary_02082010.pdf

Kenya Power and Lighting Company. (2016). *Last Mile Connectivity*. http://www.kplc.co.ke/content/item/1120/last-mile-connectivity

Kenya Power and Lighting Company. (2017). *History and Milestones*. Retrieve from http://www.kplc.co.ke/content/item/61/history-and-milestones

Kenya Vision 2030. (2007, October). *A Globally Competitive and Prosperous Kenya*. Retrieved from https://www.researchictafrica.net/countries/kenya/Kenya_Vision_2030_-_2007.pdf

Khandker, S., Samad, H., Ali, R., & Barnes, D. (2012, August). *Who Benefits Most from Rural Electrification? Evidence from India.* Retrieved January, 2017, from http://ageconsearch.umn.edu/bitstream/125090/2/AliR.pdf

Kirubi, C., Jacobson, A., Kammen, D., & Mills, A. (2009, July). Community-Based Electric Micro-Grids Can Contribute to Rural Development: Evidence from Kenya. *World Development 37*(7), 1208-1221. Retrieved from http://www.sciencedirect.com/science/article/pii/S0305750X08003288

Kuo, L. (2017, January). Kenya's National Electrification Campaign is Taking Less than Half the Time it Took America. *Quartz Africa.* https://qz.com/882938/kenya-is-rolling-out-its-national-electricity-program-in-half-the-time-it-took-america/

Lee, J., & Wong, D. (2001). *Statistical Analysis with ArcView GIS*. New York: John Wiley & Sons.

LeSage, J., & Pace, R. (2009). *Introduction to Spatial Econometrics.* Boca Raton, FL: CRC Press.

Min, B., Gaba, K. M., Sarr, O., & Agalassou, A. (2013). Detection of Rural Electrification in Africa using DMSP-OLS Night Lights Imagery. *International Journal of Remote Sensing, 34*(22), 2013-09-23. Retrieved from http://www-personal.umich.edu/~brianmin/MinEtAl_Detection_IJRS_2013.pdf

Ministy of Energy and Petroleum (2012, May). *National Energy Policy*. Retrieved from http://www.ketraco.co.ke/opencms/export/sites/ketraco/news/Downloads/National_Energy_Policy_Third_Draft_May_11_2012.pdf

Ministry of Energy and Petroleum (2015, June). *National Energy and Petroleum Policy*. Retrieved from http://www.erc.go.ke/images/docs/National_Energy_Petroleum_Policy_August_2015.pdf

Ministy of Energy and Petroleum (2016). *About Us*. Retrieved from http://www.energy.go.ke/index.php/about-us/background.html

National Oceanic and Atmospheric Administration. (2017). *Defense Meteorological Satellite Program (DMSP).* Retrieved from https://ngdc.noaa.gov/eog/dmsp.html

Obudho, M., Muguti, J., Bore, K, & Kakinyi, M. (2015, December). *Kenya Demographic and Health Survey 2014 Final Report*. Retrieved from http://dhsprogram.com/pubs/pdf/FR308/FR308.pdf

Ochieng, Lilian (2016). Breakthrough for Kenya's Budding Oil Production Industry. *Daily Nation.* http://www.nation.co.ke/lifestyle/smartcompany/Kenya-prepares-to-join-league-of-oil-exporters/1226-3021502-117v3hiz/index.html

Okoth, Edwin. (2016). Experts Poke Holes on Plan to take power to every home. *Daily Nation.* Retrieved from http://www.nation.co.ke/lifestyle/smartcompany/1226-3483888-miyytdz/

Olang, T, & Esteban, M. (2016). *Sustainable Renewable Energy Financing: Case Study of Kenya*. Sustainability Through Innovation in Product Life Cycle Design.

Regulus. (2017). *Electricity Cost in Kenya* [Data file]. Retrieve from https://stima.regulusweb.com/

Rutstein, S., & Rojas, G. (2006). Guide to DHS Statistics. *USAID DHS Program Publication Summary*. http://dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_29Oct2012_DHSG1.pdf

Shankar, V. (2014, January). Shaping a Sustainable Future Through Education. *The Huffington Post.* Retrieved January, 2017, from http://www.huffingtonpost.com/v-shankar/shaping-a-sustainable-fut_b_4050234.html

Songok, J. (2016, November). REA Commission Mini-hybrid Solar Power Station in Wajir South. *Kenya News Agency*. Retrieved from http://kenyanewsagency.go.ke/en/rea-commission-mini-hybrid-solar-power-station-in-wajir-south/

Squires, T., (2015, March). The Impact of Access to Electricity on Education: Evidence from Honduras. *Job Market Paper*. Retrieved from https://economics.ucr.edu/seminars_colloquia/2014-15/applied_economics/Squires_JMP_Electricity.pdf

Strauss, Valerie (2013, December). Nelson Mandela on the Power of Education. *The Washington Post.* Retrieved from https://www.washingtonpost.com/news/answer-sheet/wp/2013/12/05/nelson-mandelas-famous-quote-on-education/?utm_term=.e8b50238c228

Sustainable Energy for All. (2016, January). *Kenya's Investment Prospectus*. Retrieve from http://www.se4all.org/sites/default/files/Kenya_IP_EN_Released.pdf

Sutton, P., Roberts, D., Elvidge, C., & Meij, H. (1997). A Comparison of Nighttime Satellite Imagery and Population Density for the Continental United States. *Photogrammetric Engineering & Remote Sensing*, *63(*11), 1303-1313.

Power Africa (2015). *Investment Brief for the Electricity Sector in Kenya*. Retrieved from https://www.usaid.gov/sites/default/files/documents/1860/Kenya%20_IG_2015_05_03.pdf

Power Africa. (2016a). *Development of Kenya's Power Sector 2015-2020*. Retrieved from https://www.usaid.gov/sites/default/files/documents/1860/Kenya_Power_Sector_report.pdf

Power Africa, (2016b). *Annual Report*. Retrieved from https://www.usaid.gov/sites/default/files/documents/1860/Power_Africa_AR2016-optimized.pdf

Provost, C. (2013, March). Energy Poverty Deprives 1 Billion of Adequate Healthcare, Says Report. *The Guardian.* Retrieved from https://www.theguardian.com/global-development/2013/mar/07/energy-poverty-deprives-billion-adequate-healthcare

United Nations (2014, February 18). *Tackling Water, Sanitation, Energy Nexus Key to Sustainable Future*. Retrieved April, 2016, from http://www.un.org/apps/news/story.asp?NewsID=47165#.VtCGJvkrLIU

United States Agency for International Development. (2016). *Power Africa Kenya*. Retrieved from https://www.usaid.gov/powerafrica/kenya

United States Agency for International Development. (n.d.). *Data Collection and Use*. Retrieved from http://dhsprogram.com/data/Data-Quality-and-Use.cfm

Unites States Energy Information Administration. (2017). *State Electricity Profile* [Data File]. Retrieved from https://www.eia.gov/electricity/state/

van de Walle, D., Ravallion, M., Mendiratta, V., & Koolwal, G. (2013, June). Long-Term Impacts of Household Electrification in Rural India. *World Bank Policy Research Working Paper* 6527. Retrieved from http://documents.worldbank.org/curated/en/784414682594647 20/pdf/WPS6527.pdf

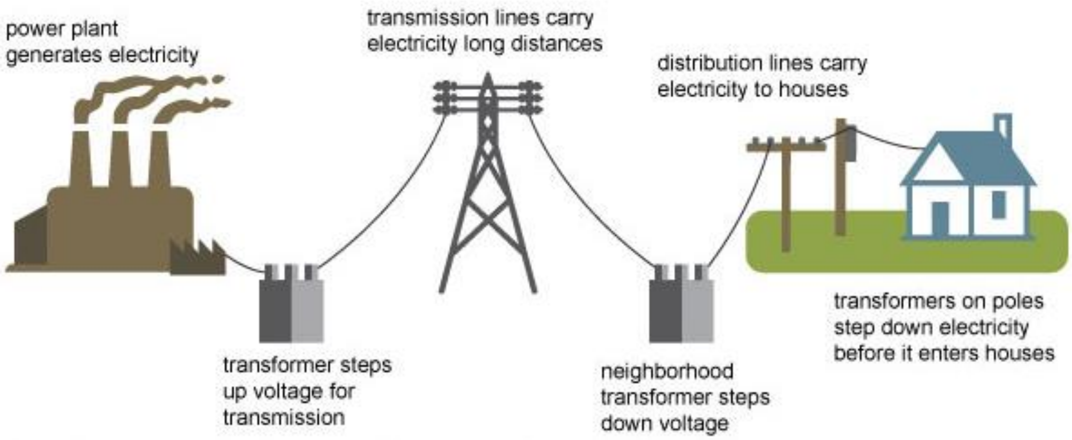
Wolfram, C., Shelef, O., and Gertler, P. (2012). How will Energy Demand Develop in the Developing World? *NBER Working Paper No*. 17747. Retrieved from http://faculty.haas.berkeley.edu/wolfram/papers/JEP%20NBER%202.pdf

World Bank (2016). *Public Participation Key to Kenya's Devolution*. Retrieved February, 2017, from http://www.worldbank.org/en/news/feature/2015/04/30/public-participation-central-to-kenyas-ambitious-devolution

World Bank. (2017). *PPP Conversion factor, GDP* [Data file]. Retrieved from http://data.worldbank.org/indicator/PA.NUS.PPP

APPENDIX: ADDITIONAL FIGURES

Figure 12. Electricity Generation, Transmission and Distribution Illustrated



Source: Adapted from National Energy Education Development Project (public domain)

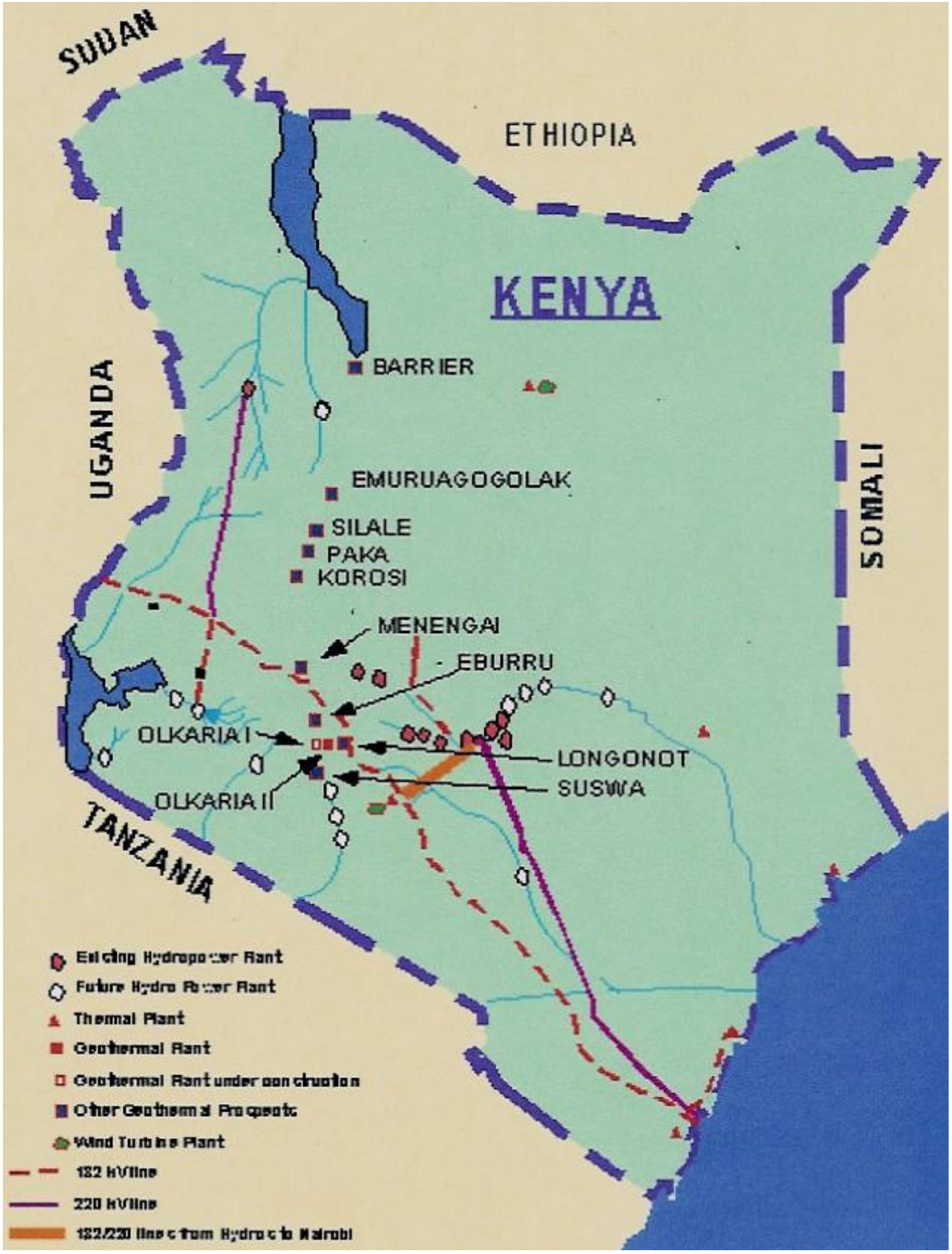
Figure 13. Nighttime Satellite Imagery in Kenya in 2009

Source: NOAA, 2009

Figure 14. Kenya's Transmission Network in 2012

Source: Government of Kenya, National Energy Policy 2014

Figure 15. Kenya’s Transmission Network in 2006



Source: Global Energy Network Institute (GENI), 2016

Figure 16. Simulations of Education Outcomes with Non-Spatial Model

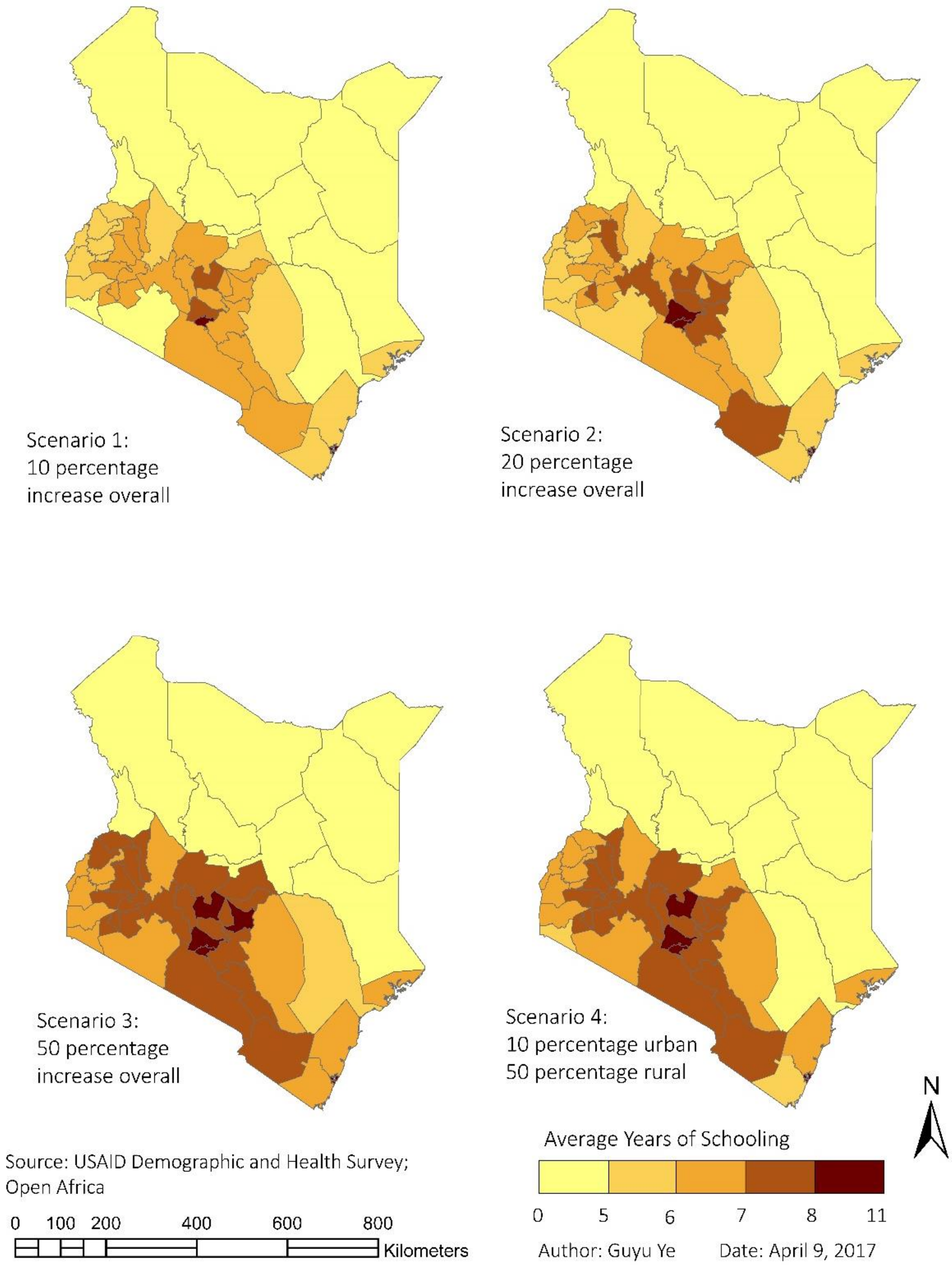


Figure 17. Simulations of Education Outcomes with Spatial Model