The Impacts of Electricity Access: Evidence from Rural India

A Thesis

submitted to
The Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Suryadeepto Nag

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

May, 2023

Supervisor: Dr. David I. Stern

Certificate

This is to certify that this dissertation entitled "The Impacts of Electricity Access: Evidence from Rural India" towards the partial fulfillment of the BS-MS dual degree program at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Suryadeepto Nag at the Crawford School of Public Policy, The Australian National University under the supervision of Dr. David I. Stern, Professor, Arndt-Corden Department of Economics, Crawford School of Public Policy, The Australian National University, during the academic year 2022-2023.

D. Stern

Dr. David I. Stern

Committee:

Dr. David I. Stern

Dr. Bejoy K. Thomas

This thesis is dedicated to the millions of children who are still growing up in darkness.

“গ্রাম নগর-মাঠ পাথার বন্দরে তৈরী হও।

কার ঘরে জ্বলেনি দীপ চিরআঁধার, তৈরী হও।

এই মিছিল সব হারার, সব পাওয়ার এই মিছিল।। ”

- সলিল চৌধুরী

Declaration

I hereby declare that the matter embodied in the report entitled "The Impacts of Electricity Access: Evidence from Rural India", are the results of the work carried out by me at the Crawford School of Public Policy, The Australian National University under the supervision of Dr. David I. Stern and the same has not been submitted elsewhere for any other degree.

Suryadeepto Nag

Suryadeepto Nag

Acknowledgments

There are several people without whom this thesis would have been limited to a rudimentary idea. Firstly, I would like to express gratitude towards my supervisor, Prof. David I. Stern for his constant guidance and indispensable insights throughout the duration of the project. I would also like to thank Debasish Das for being a mentor and a friend and encouraging me constantly throughout my time in Canberra, and the many months since. I would also like to thank Prof. Bejoy Thomas for his valuable feedback from time to time, as a member of my Thesis Advisory Committee. In addition, I would like to thank Dr. Siddhartha P. Chakrabarty, Debasish Das, and Adithyan Unni for their careful comments and suggestions on the manuscript.

Thomas Wolfe said that "We are the sum of all the moments in our lives - all that is ours is in them: we cannot escape it or conceal it." And this thesis too, has undoubtedly been influenced by the moments of my life. Therefore, I owe my thanks to all the people in my life who have shaped the way I think and do research - my mother and father, who set the process in motion, Dr. Siddhartha P. Chakrabarty and Dr. Sankarshan Basu, who introduced me to research, my friends from Bangalore who have managed to stick by me through the test of time, the friends I made in college, who did more than stick through me - Bagheera who stayed up with me when I had to work nights, and Shivani who made my days easier than they should have been, Satavisha, who walked me through every stage of the MS thesis process and made the last year as memorable as it has been.

Lastly, I would like to thank Hostel-1, where I got help from more people than I can count, the entire fraternity of students and professors at the Indian Institute of Science Education and Research Pune - which sent me down the path of studying development, and the community at The Crawford School of Public Policy, The Australian National University.

Abstract

This thesis studies the causal effects of electricity access and quality on villages and households in India. Using three panels of rural household data, we quantify the impacts of short-term (0-7 years) and long-term (7-17 years) access on household well-being. As our identification strategy, we use a propensity-score-weighted-difference-in-differences design and find that electricity access increases consumption and education in the long term, and reduces the time spent by women on fuel collection, although we find no significant impact on agricultural income, agricultural land holding, and kerosene consumption. We observe differences in the impact on rich households and poor households, and on less-developed and more-developed states, but the long-term impacts are consistently found to be greater and more significant than the short-term gains. In addition to the study on short-term and long-term access, we also investigate the village-level spillover effects of electricity. Employing a propensity-score-weighted-difference-in-differences strategy, we find positive spillover effects of village electrification on the consumption of households not connected to the grid. Subsequently, using a difference-in-differences design, we find that improvements in reliability lead to falling agricultural labor wage rates. To complement the analysis, we investigate the channels through which wages may be affected. We find that although the mechanization of agriculture increases the demand for agricultural labor, better reliability reduces women's time burden of fuel collection and allows them to provide labor to family farms, reducing the demand for hired agricultural labor.

Contents

Chapter 1

Introduction

1.1 Electrification and Impacts

The seventh United Nations Sustainable Development Goal targets achieving universal access to affordable, reliable, and modern energy services. However, according to the World Bank, 759 million people around the world did not have access to electricity in 2019. Among those with an electricity connection, a large number of households in the Global South are only nominally connected - facing problems with the capacity and reliability of the connection. Although the past few decades have seen governments around the developing world invest in electrification projects, the most recent Sustainable Development Goals Report (Sachs et al., 2022) paints a grim picture. The rate of electrification has slowed in recent years, owing to COVID-19 and the difficulty in bringing access to remote and particularly underdeveloped regions of the world. As a result, there was a 37.5% reduction in the increase in electricity access rate in the period 2018-2020, from the annual average of 0.8% in 2010-2018. As of 2020, over half of Sub-Saharan Africa and about a third of Oceania (excluding Australia and New Zealand) lacked access to electricity. Connecting such a large population of people will require large investments from governments and non-government sources. However, governments and other development institutions are often constrained by resources and often have to choose between multiple development projects. Therefore, it would be of particular interest to policymakers if electrification had direct or indirect socioeconomic benefits.

In this thesis, we explore the causal effects of electricity access and quality on well-being, using India's last-mile grid roll-out as the setting. The principal challenge in causality studies is whether all the factors can be controlled for adequately and whether the time period over which the data is collected is long enough to make proper inferences (Bruns et al., 2014). Therefore, apart from being connected to the grid, we also study the impacts of the effect of the time since which a household or village has been connected and also study the village-level impacts of reliability on agricultural labor wages. We work with a large representative data set from India, with panels spanning a period of 18 years such that it is suitable for making causal inferences, and long enough to contrast the short-term and long-term effects of electrification (if any).

Even though the direct impacts of electricity access on households and villages are frequently visited as an empirical question, there is little knowledge of the "spillover" benefits of electrification, and how electricity access can bring indirect changes in the organization of the rural economy. Similarly, while access to electricity is often the focus of impact evaluation studies, reliability or the availability of good quality electricity, in general, is less studied, despite the Global electrification expansion project moving towards improving the quality of power, as the next step to eradicating energy poverty.

The thesis is primarily structured into two major research problems. The principal objective of Chapter 2 is to investigate the causal impact of the time since connection, i.e., the duration (years) for which a household has had electricity, on several outcomes of interest relevant to household well-being, such as income, consumption, education, and domestic fossil-fuel consumption. Due to the absence of household-level data on when each household was connected, we leverage the three panels of the Human Development Profile of India (HDPI) (NCAER, 1994) and the India Household Development Survey (IHDS) (Desai et al., 2005, 2011-2012), to study households that have had a short duration and a long duration of electricity access. We study the short-term and long-term benefits of electricity access to gauge the time-associated benefits of electrification. Since we analyze the impacts on an array of different outcome variables with their own relationships with electricity access, we aspire to deduce how quickly different outcome variables respond to the treatment (being connected to the grid) and whether the time-associated benefits of electrification are different for different dimensions of development.

In Chapter 3, we study the indirect or spillover benefits of electricity access and quality. We first estimate the spillover effects of rural electrification by studying the impact on the consumption of households that were not connected to the grid themselves. Furthermore, we explore a specific facet of spillovers - the impact on casual agricultural labor and wages. The objective of the second part of this study is to explore if electricity availability has an influence on labor wages in agriculture - intrinsic to the rural economy, yet seemingly unrelated to electricity. We investigate whether the quality of electricity available in villages, measured in the average hours of daily availability of power, has any causal impacts on agricultural labor wages. Taking the analysis further, we subsequently explore the channels through which the reliability of electricity could possibly induce significant changes in the rural economy and agricultural practices that bring about changes in agricultural labor wage rates. We use post-liberalization rural India as the setting and study causal effects at the village level.

1.2 The effect of duration of connection

As a convenient and modern source of energy, electricity access is often associated with multiple benefits in development and well-being. Unfortunately, there are very few studies that investigate how long it takes for households to start experiencing said benefits of having access to electricity. As a result, it is also unknown if the impacts are most felt immediately after connection, with the benefits waning over time, whether the impacts grow with time, or whether the relationship between time and benefits takes another form.

The perceived benefits of electrification are manifold. Apart from access to electricity intrinsically being desirable, electrification could potentially impact several development indicators like income, education, health, and labor participation. These potential consequences could have significant implications for government interventions in improving electricity access and reliability, as policymakers constrained by resources are more likely to promote and carry out electrification schemes if the proposed interventions are likely to have causal impacts on the socioeconomic well-being of households. In areas still lacking access to good quality electricity, households and villages are forced to make do with primitive sources of energy like biomass and fossil fuels such as kerosene. Switching from these sources of energy to electricity could impact well-being at two levels. At the level of the village, electrification could have several benefits. It could lead to development through more efficient lighting (Fouquet, 2008), and higher productivity (Rud, 2012; Kander et al., 2014). This could bring new industries or expand existing ones. This would in turn lead to higher wages and new employment opportunities. Electrification could also accelerate infrastructure development, particularly in information and communication systems, which would further facilitate economic activities. At the household level, access to good-quality electricity could lead to cleaner, and more reliable sources of lighting, and heating. Employing electricity instead of biomass or fossil fuels in lighting could reduce incidences of cardio-respiratory diseases (Pokhrel et al., 2010; Prasad et al., 2012; Connellan, 2017). The installation of electrical appliances could reduce the labor requirements of household chores and individuals, which could increase participation in the labor market, particularly for women. Better lighting could provide an opportunity for children to study and may lead to an improvement in their performance in schools and in the long run, their employment prospects and income. Electricity could also enhance household income by aiding small businesses and installations of electric pumps and other appliances could also increase agricultural income.

There is, however, little consensus in the academic literature on whether the prevalent correlation between energy access and well-being implies a causal impact. While there is a strong correlation between electricity access and development, it is not known whether access to electricity brings about socioeconomic well-being or whether more electricity access is simply highly correlated with development. Developed countries or better-developed regions within countries are more likely to have both affluent households and better electricity access. Reverse causality is another concern, as it is also possible that wealthier households would receive electricity earlier and be able to afford better quality electricity compared to poorer households. Therefore, impact evaluation studies are crucial in determining the presence of causal impacts of electrification, or lack thereof.

Randomized Control Trials (RCT) have been a significant influence on development economics research in recent years, due to their utility in understanding cause-effect relationships of interventions (Banerjee et al., 2016), through random assignments of households into treatment and control groups. In the electricity access literature, RCT-based impact evaluation studies reveal mixed results. Among studies that evaluate the impacts of grid connections, Lee et al. (2020b) find mostly neutral impacts of electricity connections in rural Kenya on primary economic and non-economic variables, although notably, they find a statistically significant impact on the number of hours worked, in line with the seminal observational study by Dinkelman (2011), and the normalized satisfaction of the household. They also find a reduction in the use of kerosene as fuel. Another study (Barron and Torero, 2017) finds a significant reduction in kerosene expenditure and a large reduction in particulate matter concentration in Northern El Salvador. Studies based on off-grid connections (Aevarsdottir et al., 2017) also show mixed results in terms of impacts. A problem with RCTs however, is that the size of the data is often limited, and is often localized to small regions and populations, which may make it difficult to extend their results to other populations in different contexts (Pritchett and Sandefur, 2015). There are further constraints on the duration of the study, which can make it difficult to carry out RCT studies over long periods of time. In our study, we opt for a quasi-experimental approach, which allows us to investigate a longer period of 17 years and a larger sample of households from over 700 villages spread across India.

Observational or quasi-experimental studies in the literature often show positive results for the impacts of electricity access on welfare in the Global South including studies in Bangladesh (Khandker et al., 2012), India (Chakravorty et al., 2014; Khandker et al., 2014; Samad and Zhang, 2016; Van de Walle et al., 2017), South Africa (Dinkelman, 2011) and in Viet Nam (Khandker et al., 2009, 2013). Each of these studies reports positive causal impacts of electricity access, although the effects are measured to different degrees in the various studies. Observational studies themselves are not free from criticism, and in a systematic review, Bayer et al. (2020) report that observational studies report more positive results than experimental studies. While this could be due to unavoidable constraints present in experimental designs, it may also be because of liberal assumptions in observational designs and the absence of proper randomization. Lee et al. (2020a) argue that instrumental variables used in the literature so far, such as geographic cost-based instruments, may not always satisfy the exclusion criteria. In the absence of instruments that adequately satisfy the exclusion criteria, we opt for an alternative approach. We use two-stage propensity-score-weighted-difference-in-difference regressions to address the bias in the sample by weighting households to simulate a sample where the treatment (here, grid connections) was assigned approximately randomly. We then proceed with a difference-in-differences design to adequately address composition effects and arrive at a robust estimate of the impacts of electrification.

There are observational studies, however, such as in Thailand by Cook (2005), in Rwanda by Bensch et al. (2011), and in India by Burlig and Preonas (2016) that do not find strong positive impacts. These disparities in impacts could a product of the data used, but could also be a result of the studies being conducted in different countries. A geographic bias in impact evaluation studies is already known to exist and has been discussed at length in reviews by Peters and Sievert (2016) and Hamburger et al. (2019). Peters and Sievert (2016) argue that one of the reasons Africa observes weaker results than South Asia may be because Africa seldom has studies that have evaluated long-term impact. The study by Burlig and Preonas (2016) is especially pertinent in our context, as the study is based in India and investigates the effectiveness of the Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY), a massive rural electricity roll-out scheme that was launched in India in 2005, which overlaps significantly with our period of study. The authors find using nighttime data, that although the prevalence of electricity has indeed expanded, its impacts on economic welfare are limited, in the short-to-medium run. On the other hand, Van de Walle et al. (2017) study the long-term impacts of household electrification using a DID and an IV design. The authors use the India Rural Economic and Demographic Survey (REDS) which covers a period of 17 years, which is coincidentally the exact same period that we study in our study. They find that electricity brings significant positive changes in the consumption of houses that were electrified.

This dichotomy in the measured impacts between short-term and long-term studies may indicate that electricity access may not immediately trigger economic and social progress, and it is likely that it would take several years before many potential benefits become observable. Apart from some direct benefits from an electricity connection, most other benefits could show up after a "time lag", where years of an active connection would be required to show up improvements in the indicators studied. For instance, while electricity could have an immediate impact on income for household businesses that would be supplemented with electrical appliances, lighting, or heating, a potentially larger long-term impact on income might be found in households where electricity enables better schooling and education. Similarly, an improvement in the education level of decision-makers in households will likely take at least a generation to become observable. Thus, in our study, we explicitly factor in the duration of the connection to distinguish between different types of treatments - short-term connections (0-7 years), and long-term connections (7-17 years), in order to gain better comprehension regarding the impact of rural electrification on household well-being.

We base the study on post-liberalization India, with a nationally representative sample of households, each household in our study receiving access to electricity only after the 1991 market reforms. The main period of analysis involves the years between 2005 and 2011-2012, a period that overlaps with the Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY), an expansive rural electrification program launched in 2005. We analyze households that were first electrified after 1994 and before 2011-2012. According to the World Bank Global Electrification Database, less than half the population of India (49.8%) had electricity in 1994, and by 2011-2012, the number had risen to 79.9%, with more than half a billion people having been brought under the grid in this period. This aggressive roll-out of new electricity connections, predominantly in rural areas, in the World's second-most populous country (at that time), makes for a great natural experiment for an impact evaluation study.

Our empirical strategy in estimating the benefits of short-term and long-term access to electricity involves exploiting a subset of the households surveyed as a part of the IHDS in 2005 and

2011-2012, which had previously been surveyed in 1994 in the HDPI survey. Therefore, despite studying the levels of independent and dependent variables between 2005 and 2011-2012, we are able to classify households based on whether they got connected between 2005 and 2011-2012 (short-term connection) or 1994 and 2005 (long-term connection) which are analyzed as two types of treatments, along with households which had still not been connected by 2011-2012 which constitute our control group.

Since improved-electricity access is both the result and a potential driver of development, making causal inferences regarding the impacts of electrification is a challenging process. Ideally, we would like to quantify the improvements in well-being that households were to experience if electrified, over the case where they are not. However, since such counterfactuals are not observed in reality, and all households we observe are either electrified or not, calculating the effect of the treatment involves making the assumption that we can arrive at an approximate result by comparing households that have been connected to the grid against households that have not been as fortunate. While this is a problem with all empirical studies, observational studies further have to address the concern of eliminating selection bias in the assignment of households to control and (in our case, two) treatment groups. Thus, we need to correct for factors that could improve the odds of some households getting connected to the grid over others, along with correcting for bias in some households getting connected earlier than others. To do so, we use propensity-score-weighted regressions, an approach that estimates parameters, to truly reflect the effect of connecting an average household by weighting the sample to correct for selection bias.

Researchers have used the HDPI-IHDS surveys in the past to study the impacts of rural electrification in India. Chakravorty et al. (2014) use an instrumental variable approach in studying the impact of electricity access on income using the HDPI survey and the first IHDS survey round. Samad and Zhang (2016) use only the two IHDS survey rounds to study the impact of electricity access on a variety of well-being outcomes. Our study is significantly different from these works, first and foremost, in the research question. While these articles were focused on studying the impact of electricity access, and in particular the role of quality of electricity as a driver of income and well-being, our study is focused on studying the difference between short-term and long-term access, and questions about how the effects of electricity access change over time.

Furthermore, both Chakravorty et al. (2014) and Samad and Zhang (2016) make certain methodological concessions, which makes their results difficult to interpret. For instance, in Chakravorty et al.'s construction of the reliability variable, the authors use "high quality" and "low quality" as two classes of reliability along with households that were not connected at all. However, in 1994, households were asked about the frequency of outages, and in 2005 households were asked about the number of hours of electricity available daily, and the reliability variable took the value 1 in 1994 if households had power continuously (no outages), and value 0.5 if it had one or more outages, and 0 if the household was not connected. But the variable took the value 1 in 2005 if households received more than 18 hours of electricity in a day, the value of 0.5 if it received between 0 and 18 hours, and 0 if the household was not connected. The two variables for reliability - one based on the average interruption frequency, and another based on the average interruption duration are not exactly comparable. Households that could take the value 1 in one measure, could quite easily take the value 0.5 in another. Secondly, there is no known basis to assume that the wide range of reliability measures (either in terms of frequency or duration) that have been categorized as low-quality, should be given half the value as the maximum, and a design with each category being a categorical variable, may have been easier to interpret.

Our design is closer to that of Samad and Zhang (2016), although we have three major differences in our methodology. Since we are interested in studying short-term and long-term effects we use two treatment groups of households with short- and long-term access along with a control group that does not have access. Furthermore, we only consider households that had not been connected by 1994-1995 and then received access in one of two periods, if at all. Thus, our control group is restricted to households that never had electricity. In Samad and Zhang's design, the control group includes all households which did not see a change in their electricity access status, i.e. those households which never got access, together with those households which had always had access. Such a design implicitly assumes that electricity access has no long-term impacts, an assumption that is completely untested. Finally, despite both our approaches being based on difference-in-difference designs, with propensity-score-weighted regressions, the variables we use in determining propensity scores differ. Household-level variables such as affluence, and village-level variables such as the fraction of households already connected and the caste composition of villages are not considered while determining propensity scores in their study, but we incorporate these variables which changes the results significantly.

We find that access to electricity has very limited impacts for households that have had access to electricity for less than seven years, on per-capita consumption, agricultural income, agricultural land holding, education, kerosene consumption, and time spent by women in collecting fuels, and even though there are sizeable effects on some well-being outcomes such as land holding, these effects are not statistically significant. However, we find statistically significant increases in per-capita consumption, and education, and a statistically significant decrease in time spent by women collecting fuels. Based on these results, we conclude that the effects of electricity access grow over time, and the function that quantifies benefits would be a convex function of time with the slope increasing with time. To check for robustness, we repeat the analysis with households connected before 1994, as households with "very long-term access" and find that the result is that the effects grow over time for almost all variables, despite the inclusion of the third group which nearly doubles the size of the sample for most variables.

We also segregate the data into different geographical regions of the country and study if already developed regions (Maharashtra, Karnataka, the erstwhile Andhra Pradesh, Kerala, and Tamil Nadu) respond better to electricity when compared to less developed regions (Bihar, Jharkhand, Madhya Pradesh, Chattisgarh, Rajasthan, and Uttar Pradesh. We find that although the latter shows a statistically significant improvement in consumption, the increase is not reflected in other variables such as education, and kerosene consumption. Conversely, the already developed southern states do not see a statistically significant improvement in consumption but see statistically significant effects on education and kerosene consumption. However, the broader trend that long-term effects are larger and more statistically significant (if at all) is preserved nonetheless. Similarly, we segregate households into poor and rich households. The results on individual variables are more complex with the poor group performing better in terms of affluence (consumption and land-holding) while richer households perform better in the non-economic well-being outcomes like education and time spent in collecting fuel. Some variables such as land-holding respond to electricity for poor households, even though it does not for rich households or for the pooled data. Variables like consumption and land-holding also show a significant change with short-term access. Still, we find that the trend is conserved, for most variables, although the trend is reversed for agricultural income and agricultural land-holding, where there is a greater increase in the short term than in the long term, with electrified rich households reducing their land-holding both in the short-term and the long-term.

1.3 Spillover Effects

One question of particular relevance to impact evaluation of infrastructure projects is whether there are spillover effects. The spillover effects of electricity access are the potential impacts felt by households that do not have access to electricity, due to benefits received by other treated households. Spillover benefits are important to policymakers as it helps them evaluate the impact of electrification programs on households that may not have been electrified themselves, but experience the benefits of the village and other households in the village being connected to the grid. From the perspective of impact evaluation, spillovers cause problems in the estimation of causal impacts of electricity access on households, as the spillover effects result in a paucity of households in the study village that satisfactorily act as controls. Thus, in the third chapter of this thesis, we investigate the spillover effects of rural electrification on household welfare and labor wages. We study the effect of the fraction of connected households in a village on the expenditure of households not connected to the grid themselves. Furthermore, we study the impact of the reliability of electricity in the village on casual agricultural labor wages in the village.

In a household-level impact evaluation study, if one simply estimates the effects by contrasting the impacts on households that were treated and households that were not, the estimate may be biased due to spillovers, which may lead to a positive or negative bias in the estimate of causal impacts depending on whether the spillover effect is negative or positive. This bias in the estimate presents several hurdles in the estimation of causal impacts, as it could potentially lead to a violation of the Stable Unit Treatment Value Assumption (SUTVA), which assumes that the potential outcomes for any unit are unaffected by the status of treatment to any other unit. For validity, SUTVA requires that there are no interaction effects of the treatment or spillover effects. In the case of electricity access, the condition need not be satisfied. If access to electricity increases the income of a household, they may allocate larger sums of money on consumption, which could create a higher demand for goods, and the providers of the goods could see their income grow without receiving electricity themselves. Access to electricity could lead to progress in enterprises and businesses in villages, which could provide new sources of employment for households that are not connected themselves.

The violation of SUTVA and spillover benefits pose constraints for causal inferences. For instance, an instrumental variable sometimes used in the literature in household-level impact evaluation studies is the presence of village-level connections or variants of the same. The exclusion criterion in instrumental variable designs requires that the instrumental variable does not affect the outcomes of interest except through the treatment. However, if there are spillover benefits of electricity access, it implies that village-connection-based instrumental variables would all violate the exclusion criterion.

In the absence of an estimate of the spillover effects, one way to avoid biased estimates is by changing the level of the analysis. In the context of the impact evaluation of rural electrification, moving the analysis from the level of the household to the level of the village or district. In doing so, spillover effects within the village or district will not lead to biases in causal effects, and the only condition required to be satisfied is the absence of interaction effects between villages, for the validity of the assumption, which is a much easier task. However, there are some drawbacks to this approach. To begin with, they impose notable restrictions on the size of the data. Since there are several households that make up a village, the smallest unit of measurement in a village-level analysis is an average of household effects for several households. Secondly, this restricts the inferences that one can draw from the data. One would never be able to study the impact of electricity access on households at all. Instead, an empirical strategy incorporating spillover effects in the model can help estimate spillovers, and unbiased estimates of the effects of electricity access, without altering the research question or reducing the size of the sample. Another alternate could involve the quantification of the causal effects of impact and spillover benefits separately, which is one of the principal objectives of Chapter 3.

In our analysis, we use two facets of rural electrification at the village level - the fraction of households with electricity access, and the reliability of electricity. Both of the quantities are of considerable relevance to measuring the impact of electrification in the village. The former measures the extent of connecting households, and the latter measures the quality of electricity that the village receives.

We first focus our analysis on the spillover effects of the fraction of households connected. To do this, we study a panel of households from the India Development Household Survey (Desai et al., 2005, 2011-2012). We study a sample of households that did not have electricity in 2005 and households that did not have electricity in 2012. While several studies (Khandker et al., 2009; Chakravorty et al., 2016) discuss the potential of spillover effects, and some studies even control for it (Bensch et al., 2011), rarely are the spillover effects explicitly estimated (Khandker et al., 2013; Van de Walle et al., 2017). Khandker et al. (2013) include commune electrification as a variable, apart from household electrification, in their study in Viet Nam. The coefficient of the commune electrification returns the spillover effect of commune electrification. The authors found statistically significant positive impacts of commune electrification on wages and salaries, and on agricultural income. However, the impact felt by households themselves connected, and those not connected may differ quantitatively. Thus in order to measure the true spillover effects i.e., on households that were not connected to the grid themselves, we restrict our estimation to households without access to electricity and compare them based on the status of village electrification. Van de

Walle et al. (2017) use a different strategy to estimate spillover benefits. They include a linear term that includes a variable representing the number of years since the village was first connected, and the coefficient of this term is used to quantify the effect of connecting the village. The authors find positive impacts of the years since connection in the village, especially on households that were not themselves connected to the grid. This strategy assumes that the spillover effects of electrification grow with time. However, the increase in spillovers (if any), for villages connected longer may not necessarily be due to a growing duration since connection, and may also be due to more households progressively getting connected over the years leading to increased benefits. To disentangle the effect of time and the effect of the number of households connected, we study both variables independently in our design.

In order to estimate robust results, corrected for selection bias in assignment to treatment, we use a generalized propensity score-weighted-difference-in-differences design. We find a statistically significant positive spillover effect on both the logs and levels of per-capita household consumption in households without electricity access. Contrary to the results of Van de Walle et al. (2017), we find no significant spillover effect of the number of years that a village has been connected for. Interestingly, we observe a statistically significant negative impact of reliability on per-capita consumption, implying that villages with better reliability have relatively lower per-capita consumption in households without access, although the precise estimate in this particular analysis may be biased as the sample is weighted to balance the assignment to treatment in terms of the number of households electrified.

Furthermore, our analysis of consumption, although informative about the overall effect on affluence, does not lend insights into the channels through which village electrification causes spillovers. It also does not explain why reliability is found to have a negative impact on consumption. Therefore, to get deeper insights into spillover effects, we shift our focus to reliability and its effect on wages, agricultural wages in particular.

Since rural India is still rather agrarian, agricultural laborers make up a majority of the labor force in Indian villages. Studies in the past have assumed that the spillover benefits of village electrification in India would be minimal due to most households being self-employed in agriculture (Chakravorty et al., 2016). However, the data on village-specific wage rates of casual agricultural labor paint a different picture (Table 1.1). From the table, we can see that the change in the wage rate, after controlling for inflation, is consistently lower among villages that see an improvement in reliability. Although, the observation may merely be a correlation and it is important to study if

		Means (standard deviation)					
		2004-2005		2011-2012		Δ	
		Positive Treatment	Negative Treatment	Positive Treatment	Negative Treatment	Positive Treatment	Negative Treatment
Wage Rate for Women (2012 Rs. per day)	Kharif (summer/monsoon season)	86.73	85.94	134.2	138.9	41.23	46.53
		(39.8)	(42.51)	(59.26)	(70.89)	(48.73)	(50.46)
	Rabi (winter season)	84.56	85.63	132.93	138.86	40.02	48.01
		(40.91)	(42.399)	(56.43)	(68.46)	(48.19)	(52.45)
	Annual Average	86.71	85.98	133.29	138.93	40.44	47.31
		(40.2)	(41.72)	(57.2)	(69.97)	(46.99)	(49.97)
Wage Rate for Men (2012 Rs. per day)	Kharif (summer/monsoon season)	116.99	119.66	168.26	183.48	51.24	61.93
		(51.51)	(52.81)	(68.21)	(83.76)	(57.66)	(63.38)
	Rabi (winter season)	115.95	117.95	166.76	184.65	49.48	63.51
		(53.04)	(51.8)	(66.04)	(83.15)	(56.72)	(61.76)
	Annual Average	116.5	119.55	167.2	183.32	50.52	62.81
		(51.45)	(52.21)	(66.76)	(83.07)	(56.14)	(61.62)
Wage Rate Average (2012 Rs. per day)	Kharif (summer/monsoon season)	105.37	105.46	151.52	161.73	45.53	54.75
		(46.86)	(46.57)	(61.01)	(74.74)	(51.71)	(56.00)
	Rabi (winter season)	105.23	104.74	150.13	162.96	42.8	55.58
		(49.01)	(46.66)	(58.52)	(73.85)	(51.1)	(54.81)
	Annual Average	104.93	105.42	150.5	161.63	44.8	55.45
		(47.08)	(46.46)	(59.37)	(73.87)	(50.16)	(54.64)

Table 1.1: Casual agricultural labor wage rate statistics for positively and negatively treated villages. “Positive Treatment” villages refer to those villages which see an improvement in the average number of hours of electricity received in a day, “Negative Treatment” villages refer to those which do not, these include the control villages which observe no change in reliability. Includes 1281 Indian villages. The wages have been standardized to the 2012 Rupee by multiplying the 2005 rates by the national consumer price index (reported by the World Bank). Averages are unweighted. The data are unweighted. Source: IHDS I, and IHDS II surveys.

reliability is truly a causal factor, after controlling for other characteristics and village-level fixed effects.

Among lower and lower-middle-income countries, India has been among the forerunners in the expansion of energy access, and studying the causal effects of reliability in India has consequences for several countries of the Global South expected to ramp up their power generation and distribution to meet electrification targets in the coming years. Electricity, and its quality, apart from being a marker of development could affect rural life considerably, which is intrinsically linked to the largely agrarian rural economy in India. Changes in the quality of electricity received by firms and households in the villages could affect rural life, practices, and organization and through them, the demand and supply of labor in agriculture, thereby affecting wages.

Although agriculture, especially in the manner in which it is practiced in India, seems unrelated to electricity, with limited agricultural mechanization, there is little reason to assume that agricultural practices will remain unaffected by changes in infrastructure and energy access. Electricity itself is a marker for development, and an essential resource in the 21^{st} century, and may contribute to changes in life and work in villages. Rather, by exploring the causal effects of electricity availability on agricultural labor wages, we can add to our limited understanding of the role of how electricity in agriculture. Nevertheless, agricultural mechanization is an obvious and indispensable aspect of agriculture that is likely to be impacted by the quality of electricity. In a seminal empirical study, Barnes and Binswanger (1986) found that early grid roll-out in the mid-late twentieth century led to increased agricultural productivity through private investments in mechanized irrigation in rural India. More recent studies (Badiani-Magnusson and Jessoe, 2018), as well, demonstrate that electricity subsidies lead to lower groundwater levels which the authors attribute to increased adoption and usage of irrigation pumps. With respect to the impact on wages, agricultural mechanization could potentially supplant human labor, even though the adoption of electrical machinery in agriculture has been somewhat limited in the Indian context.

The question of the causal impacts of reliability on agricultural labor wages is also an important question in labor economics, related to the literature on the nexus of technology and labor. A timeless question in labor economics is whether technological advancements bring about changes in labor wages (Brown and Campbell, 2002). Technological changes can affect labor wages in several ways. Changes in technology can influence the ways in which firms work and can affect productivity. Increased productivity may lead to better spoils for workers. However, in keeping up with technology, workers may need to update their skill mix. Studies find that older workers may receive lower wages following technological transitions (Allen, 2001). However, technological advancements need not necessarily affect only firms and establishments and may affect lifestyles and market structures themselves. In our work, the reliability of electricity is the technological marker, and casual agricultural labor is the focal occupation. By studying the impact on wages and the channels of impact, we can make inferences about the effect of the reliability of electricity on the overall manner in which agriculture is practiced.

There are few studies that look at the economic and social impacts of reliability. This includes studies on businesses (Gibson and Olivia, 2010; Allcott et al., 2016) and households. Among studies in rural settings, Gibson and Olivia (2010) use data from 4000 households in Indonesia to study the effect of the quality of infrastructure and energy on non-farm enterprises, Rao (2013) uses a cross-section of data from India to estimate the impact of electricity reliability on household enterprise income, Chakravorty et al. (2014) study the effect of reliability on non-farm income in India, with data from 1994 and 2005, and Samad and Zhang (2016) study the impact of reliability on a variety of economic and social variables in India, using one panel that was studied by Chakravorty et al., and another panel from 2011-2012. Both these studies find a positive impact of reliability on the variables considered, although Samad and Zhang find a neutral effect on agricultural income. Dang and La (2019) conduct a similar study with data from Viet Nam.

Two of these papers are of particular relevance to our study. Chakravorty et al. (2014) and Samad and Zhang (2016) are both studies in India over long periods, with panels over 7-8 years apart. Incidentally, both studies share a panel. While Chakravorty et al. use one panel from the Human Development Profile of India (NCAER, 1994) and another from the India Household Development Survey (Desai et al., 2005), Samad and Zhang use two panels of the India Household Development Survey - 2004-2005 Desai et al. (2005) and 2011-2012 Desai et al. (2011-2012). Chakravorty et al. use an instrumental variable approach to study the impact of access and reliability on non-farm income. For this, the authors assume the density of transmission lines in the district as the instrumental variable. Contrary to expectations, the authors find a larger effect in the instrumental variable approach, rather than the ordinary least squares, implying that the impact of electrification is greater after controlling for income and other household characteristics, which the authors explain by saying India's electrification program was progressive. However, the methodology, although elegant, may not be completely rigorous. This is because the data in HDPI measures does not provide reliability in terms of the number of hours, and instead uses the frequency of disruptions as a variable. Since the survey does not ask households about the average duration of the outages, questions loom over the parity between the variables of the two surveys. Furthermore, the dummy variables for low, medium, and high reliability are 0, 1, and 2 respectively, but there is little reason to assume that the effect of high reliability will be exactly twice that of medium reliability, and it may have been more rigorous to include two separate dummy variables for medium and high reliability. Samad and Zhang use a two-stage propensity-score-weighted fixed-effects regression model to quantify the impact of reliability. The authors use a continuous variable for reliability as opposed to the dummy variables used by Chakravorty et al. and once again find a greater effect in the second stage of the regression when compared to the first, which corresponds to the findings of Chakravorty et al., although the former estimate weaker impacts in comparison. In another study, Pepino et al. (2021) study the impact of the voltage of electricity available on the income of households in the Philippines. The study works with qualitative categorical variables on whether the households experience low or fluctuating voltages, along with variable quantifying reliability.

Distinct in intention from these studies, the objective of this chapter is not to investigate the impact of reliability on households. Instead, we restrict our primary focus to village-level effects. Substantiating the village-level effects with the causal mechanisms, however, demands an analysis at the level of households to understand changes in how agriculture is practiced by agricultural households and their expenditure on hired casual labor. Therefore, we supplement our village-level analysis with an analysis of agricultural households to study the demand for labor.

Establishing a causal relationship between changes in the quality of electricity access and development is challenging, and an accurate estimate of the impacts needs sufficient controlling of compositional factors that could influence labor wages. In the rural electrification impact evaluation literature alone, there is little consensus on whether electricity has causal impacts on household welfare Bayer et al. (2020). A major concern is reverse causality, where it is not the electrification of villages that causes them to be better off, but rather it may be the better-developed villages that get access to electricity. Similarly, in exploring the impact of reliability on wages, if one found a positive impact of reliability on wages, it would not be straightforward to infer whether the quality of electricity supercharged affluence and led to spillover benefits, or whether it was the more affluent villages, with higher wage rates, that saw an improvement in the quality of electricity. Alternately, if one finds that villages that saw a larger improvement in reliability, had lower wages, it need not necessarily imply that the quality of electricity drives down wages. Such an observation could be a result of poorer villages receiving better quality electricity late. More affluent villages, with higher wage rates, may have already had good quality electricity access, with little room for improvement in quality. Apart from reverse causality, estimating the impacts of infrastructure projects also has other hurdles, as improvements in infrastructure or technology, may coincide with other development in the region, which makes it difficult to attribute effects to the variable of interest alone. To address these issues, we employ a difference-in-differences (DID) design and control for changes in other village-level and household-level characteristics that could affect wage rates or other variables that we wish to study. The DID approach helps control for village-level fixed effects that do not change with time, and controlling for other characteristics helps isolate the changes in wage rates to reliability alone.

Our analysis is lent further aid by the rich IHDS survey data sets, which have data on several variables both at the level of households and villages. The data, representative at the national level, includes essential variables such as whether a village has electricity, since when a village has electricity, the fraction of households in the village that have access to electricity, and the average number of hours in a day that a village has power, allowing us to attribute the effects to the aspect of electricity availability that we are primarily interested in - reliability. The surveys also have data on various other characteristics such as the status of infrastructure in the village, proximity to banks and markets, the number of schools in the village, etc., making it easier to control for a large number of confounding variables. Furthermore, our analysis is at the level of the village, ruling out spillover benefits from the treatment in other villages.

India makes for a good setting for our study. By 2005, most Indian villages had already been connected to the grid, and according to the IHDS data, about 70% rural households had access to electricity. To electrify the remaining villages, the Government of India launched the Rajiv Gandhi Grameen Vidyutikaran Yojana. However, several villages, although connected to the grid, did not receive power for sufficient hours, and the quality of access remained poor. Therefore, India began devoting focus to increasing power generation and improving the quality of power available to villages. This also makes our study extremely relevant to large parts of the developing world where a majority of households are still connected nominally, and continue to face disruptions to the availability of electricity. According to data from the World Bank, in 2005, India also made up the largest rural population of the World with villages making up 71% of the national population. At this time, agriculture, forestry, and fisheries made up over 15% of the Indian GDP (over 100 billion 2021 USD in size), compared to the global average of 3.2% share of the World GDP, making India an ideal site for investigating changes in agricultural practices and wages.

Our primary results indicate that improvements in reliability have a statistically significant negative impact on agricultural wages. This effect is reflected in the casual agricultural labor wage rates for both men and women, although the impact on women's wages is found to be smaller. Surprisingly, this effect of wages is found only for agricultural wages, and we find no statistically significant impact on casual non-agricultural labor wage rates and domestic labor wage rates. This makes the possibility of an increase in unskilled labor supply, due to an electricity-facilitated increase in labor market participation somewhat unlikely, implying that the phenomenon was restricted to agriculture. However, we find a statistically significant reduction in the time spent by women in fuel collection, when reliability improves, which may increase labor supply. On analyzing the hiring practices in agricultural households, we found a reduction in the households' expenditure on hired agricultural labor, and a slightly less significant reduction in the number of hired man-days of labor, with an increase in the hours of electricity access, indicating that an increase in reliability led to a decrease in demand for hired farm labor.

To explain the negative impact of reliability on agricultural labor demand, we set up a two-part statistical test. We postulate that in order to establish that reliability affected labor demand through a particular channel, the variables relevant to the channel must display a statistically significant response to a change in reliability, i.e., in the first part, we estimate if an improvement in reliability causes a change in a variable that could affect labor demand. The second part involves establishing that a change induced in the channel, resulted in a statistically significant change in the expenditure of agricultural households on hired labor. In essence, the two tests first test whether a change in reliability causes a significant change to a potential causal variable, and subsequently test whether accounting for a change in the causal variable as an explanatory variable renders the effect of reliability on household expenditure on labor insignificant. To illustrate this, consider the example of agricultural mechanization. For us to be able to satisfactorily establish mechanization as a channel, we would first have to show that an improvement in reliability leads to increased agricultural mechanization. Then, in the second part, we would have to show that if agricultural mechanization is controlled for, then the effect of reliability vanishes, which would then imply that reliability affects demand through agricultural mechanization.

We considered three channels. First, we considered the possibility of an electricity-mediated transition of the rural economy, away from agriculture, which would result in lower demand for agricultural labor demand. We further hypothesize that improvements in reliability may allow agricultural households to make better use of electrical farm machineries such as electric pumps and electric threshers, which could potentially displace human labor and drive down demand. Thirdly, we hypothesized that if electricity availability gave women in agricultural households more time on their hands, they could supply labor to their household farm.

On testing the possible hypotheses, we found that improvements in reliability do not result in households moving away from agriculture to other businesses. Nor did they result in increased ownership of irrigation pumps. We also found that agricultural households that had irrigation pumps, spent more on hired labor, when controlled for other household and village-level characteristics, perhaps to operate the machinery. Instead, we find that the reduction in demand could potentially be explained by an increased supply of voluntary labor in family farms, particularly by the households' women, who had a reduced time burden of fuel collection, due to better reliability enabling a transition away from traditional fuel use which demanded long hours in fuel collection.

However, it is difficult to confirm if the reduction in the time spent by women in fuel collection necessarily is due to switching to electricity as a fuel, or if households can afford to hire labor for fuel collection. Our results in Chapter 2 indicate that households may not substitute traditional fuels with electricity, because we find no impact on the consumption of kerosene, implying that households may be fuel stacking. Furthermore, it is difficult to entirely rule out the possibility of an increasing supply of labor. We deduce that the fall in wages may not be due to supply as there aren't significant effects on non-agricultural labor wages. But this assumes that different forms of labor are fungible, which may not be a valid assumption. While we find a negative impact of reliability on time spent by women in fuel collection, it is difficult to check whether this prompts women to participate in the labor market or supply labor to family farms, or a third alternative.

Overall we contribute to multiple literatures. The first addresses the causal effects of advancements in energy and infrastructure in villages, particularly in the context of the impacts of the quality of electricity . The second explores the nexus between electricity and agriculture and shows that changes in the quality of electricity can affect agricultural practices and organization in a village. Additionally, we add to existing knowledge on the impact of technological advancements on labor wages, for the specific case of agriculture, and show that electrical agricultural equipment does not replace traditional manual agricultural labor, and instead creates more demand for it, possibly to work on the machines. Finally, we build on and add to what we already know of the impact of electricity access on the lives of women, reducing the time burden of fuel collection, which in the case of agricultural households, they can allocate to working on family farms.

The chapter is organized as follows. In Section 3.1 we describe the data. We look at the per-capita consumption in households that have and lack electricity. We also look at the status of electrification and reliability along with other characteristics and wage rates across the two time periods in 2005 and 2012. We use the data to familiarize ourselves with the relevant variables and also use it to lend support to assumptions essential to our research design, such as the random assignment of changes in reliability. We then describe our empirical strategy in detail in Section 3.2, and use it to estimate the impacts of village electrification on unconnected households, and on agricultural labor wages, the results of which are presented in Section 3.3. Finally, in Section 3.4 we explore several possible channels in detail and test various hypotheses that could help explain the effect of reliability on wages.

1.4 The History of Electrification in India

In this section, we overview the history of rural electrification in India. India represents between a sixth and a fifth of the World's population. In the developing world, its share is even larger. India has had an aggressive electrification program and the number of electrified villages and households has grown significantly in the last decade of the twentieth century and the first decade of the twenty-first century, a period that we use in our study. In this period, there are a large number of households that were electrified over a short period of time, and a sizeable number of households that still remain unelectrified. This period also came after the economic shock following the dissolution of the Soviet Union, and the decision to liberalize India's economy, and was a period of profound economic growth in the country. The HDPI-IHDS data covers the period from 1994-2012 and has data on electricity access, and reliability, at the level of households and villages, along with other socioeconomic variables of considerable importance, making it a suitable data set for impact evaluation.

According to a report by the Central Electricity Authority, under the Ministry of Power of the Government of India, no village in India had access to electricity at the time of its Independence in 1947. By 1950, the number had grown to include over 3000 villages, although most of these were still nominally connected, as the installed capacity was only about 1700 MW compared to 382 thousand MW as of 2021.

For most of India's history since India's independence from British rule, and the formation of the republic in 1950, development has been primarily planned and overseen by the Five-Year Plans. In the initial phase of India's development post-independence, the Planning Commission emphasized the development of industries, and therefore, electricity access in villages, and by extension, households were still largely limited. Only after the famines of the 1960s, did the Government set up the Rural Electrification Corporation in 1969, to facilitate rural electrification, primarily through providing financial assistance. It was only in the early 1970s that the number of electrified villages in India crossed 100,000 out of about 600,000 villages in total.

Due to agricultural problems with famines and droughts, the Plans began shifting their attention towards electrifying villages to facilitate the installation and use of electric pumps to aid with irrigation, and by the mid-1980s, over half of the villages in India had been connected to the grid. However, these figures need to be taken with a grain of salt as the number of households in these villages which actually had access to electricity was still dismal. Although a large number of villages were connected to the grid in the period from the 1970s to the 1990s, the number of villages connected plateaued. On the other hand, the installed capacity, and the per-capita consumption of electricity continued to grow exponentially. Based on India's National Sample Surveys, researchers found that only 18 percent of rural households had been connected to the grid in 1972, while over 70 percent had been connected by the end of the next two decades (Van de Walle et al., 2017).

Most villages were electrified by the turn of the century, and there remained only about 100,000 villages with a population of at least 100 people each that were still unelectrified. To connect these villages, the Government of India launched the Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY) or the Rajiv Gandhi Rural Electrification Initiative in 2005. RGGVY targeted electrifying all rural households and developed schemes to provide electricity to poor households free of cost. However, there are conflicting reports on the effectiveness of the program. While government sources and a report from the World Bank (Pargal and Banerjee, 2014) indicate that there was a significant increase in the electrification rate, particularly among rural consumers, other studies such as Burlig and Preonas (2016) use night-time lights data find that although there was a significant increase, it was limited and not as expansive as it may have been thought to be. This might indicate that while electricity access may have been brought to a larger number of households, many of these households may not have electricity of sufficient quality or may not have access to resources to be able to utilize the grid connection adequately to achieve the intended benefits. The final Plan (2012-2017) aimed at electrifying all villages in India, and the Government claimed to have achieved that objective by April 2018, and by 2019, only 18,734 households, remained to be electrified, all in the state of Chattisgarh.

Historically, rural electrification in India has involved two components, the electrification of villages, and subsequently the electrification of households within the villages. It has often been the case that villages have access to electricity but households do not. Therefore, in our analysis, when we correct for selection bias in assignment to treatment, we factor in both village-level and household-level characteristics. Studying only one of the two results in biases and errors. This is because just controlling for village characteristics that motivate earlier electrification discounts the fact that villages continued having unelectrified households for years or even decades after the village was electrified. Controlling only for household-level characteristics like affluence, on the other hand, would be dismissing that no matter how well-off a household is, they cannot get connected to the grid if electricity has not been brought to their village.

Chapter 2

The Time-Associated Benefits of Rural Electrification on Household Welfare

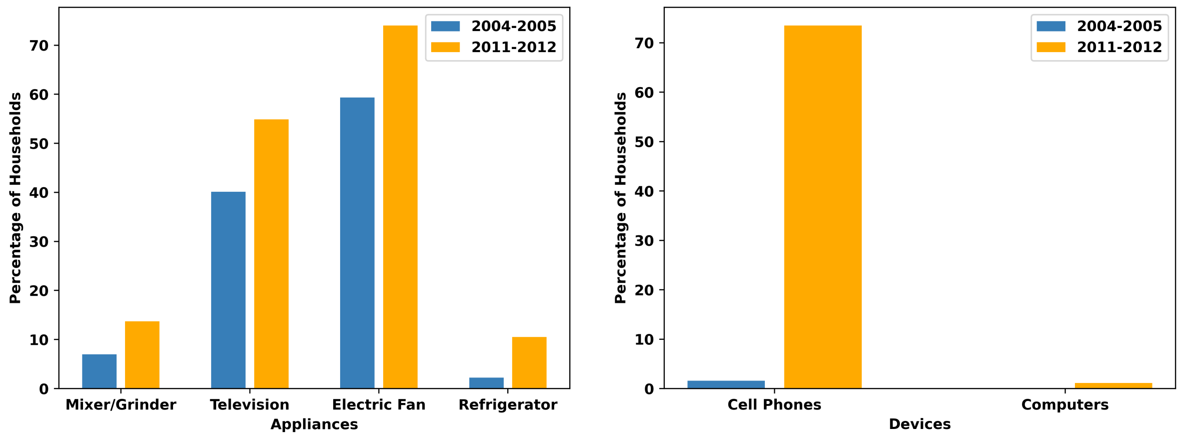
2.1 Arguments for Time-Associated Benefits

The time-associated benefits of village electrification have been explored in empirical studies in the past. Lewis and Severnini (2020) studied the benefits of short and long-run access to electricity in rural American counties that were electrified in the mid-twentieth century. They found that there were significant benefits to long-run connections. For instance, they found that counties that were electrified earlier had significantly larger improvements to median dwelling value, farmland value, retail and manufacture payroll per worker, and farm revenue per worker, over those that were electrified later. Van de Walle et al. (2017) studied the impact of how long villages have been connected in the context of spillover benefits to households in villages connected to the grid without household-level access. They found that the time since the village was first connected had a positive impact on total consumption for households that were not connected themselves. In contrast, the impact of long-term household-level connections has received considerably less interest. Considerations of the time-associated effects of electrification have been left out of the wider literature on the benefits of electrification to household well-being. For instance, in Samad and Zhang (2016), the authors use a difference-in-difference methodology where getting connected to the grid is assumed to be the treatment. This implicitly implies that households that already had electricity prior to the period of analysis are not being treated and are a part of the control group, which need not be an accurate representation of the control group.

The electrification of households could bring improvements to the lives of members in multiple ways. One hypothesis is that a majority of the benefits of household electrification are reaped indirectly through a great availability of time for members of the household, due to a reduction in, if not a total elimination of, the time burden of undertaking various domestic tasks that are rendered redundant with the advent of electricity, such as the demand for gathering alternate sources of fuel, such as firewood. Empirical studies find a connection to the grid results in a significant reduction in the time spent on biomass collection (Samad and Zhang, 2016), and an increased participation in the labor force (Dinkelman, 2011). If a convenient energy source at one's disposal indeed reduces the time spent in unpaid domestic work and increases participation in the labor force, the benefits of electrification are likely to grow with time if participation in the labor force for a longer period has benefits for the individual. Since the surplus time could also be used to augment existing household incomes with other productive activities such as businesses, growth in the business over time would translate to greater returns of electricity access in the long term. Similarly, access to electricity may further productivity by increasing the time available for productivity by increasing the duration of hours in the light. This can increase the time spent studying after school (Samad and Zhang, 2016; Aguirre, 2017), which can be beneficial for children's education.

Apart from the benefits of education which grow with years of schooling, some benefits, such as the effect of schooling on wages through enhanced employment, only come into effect after several years, perhaps more than a decade. Thus the short-term benefits due to lighting on education would only involve improvements in learning and performance, while the long-term benefits would involve the compounded benefits of learning and performance, along with the exclusively long-term benefits of better employment and consequently augmented income.

A third pathway for the benefits of electrification growing over time arises from poor households not having adequate resources at their disposal to make the best of the electricity connection that they now have. A large fraction of the benefits of the positive technological change will be unrealized if households do not have access to the resources to make use of the technology. For instance, a farmer who cannot afford an electric pump, will not be able to make full use of the electricity connection, which could have otherwise further reduced her labor hours in irrigating her field. Therefore, several households will have to wait considerably long until they can afford and buy the necessary appliances and machines to fully optimize their benefits from electricity, at any given time. This can be observed in the ownership of electrical appliances by rural Indian house-



(a) Electrical Appliances

(b) Technological Devices

Figure 2.1: Ownership of household electrical appliances and technological devices by rural households ($N = 2121$) in India which were electrified between 1994-1995 and 2004-2005. Between 2004-2005 and 2011-2012, the ownership of mixers/grinders increased from 7% to 13.7%, the ownership of televisions went from 40.1% to 54.9%, the ownership of electric fans went from 59.3% to 74.0%, and the ownership of refrigerators increased from 2.2% to 10.5%. Between 2004-2005 and 2011-2012, the ownership of cell phones increased from 1.6% to 73.5%. While there were no households that owned a computer in 2004-2005, 1.1% of the households had computers by 2011-2012. Data are unweighted. Source: HDPI, IHDS I, and IHDS II surveys.

holds that were connected to the grid, prior to 2005 (Figure 2.1a). These households, although already connected to the grid, saw their ownership of appliances grow gradually, with the ownership in 2012 being considerably greater than in 2005, thereby allowing for greater productivity and well-being. The growth in ownership is more remarkable for more expensive appliances such as televisions, mixers/grinders, and refrigerators, compared to electric fans. This could imply that affordability may have a significant role to play in the lag between getting electrified and being able to reap its full benefits.

Apart from affordability, there is also the problem of accessibility. For example, computers and cell phones have been in existence in certain affluent urban households, but they were extremely scarce in rural Indian households in the early twenty-first century, even among those that had been electrified. Advancements in telecommunication infrastructure in India have made cell phones extremely commonplace in households that have electricity, allowing households to make better use of their electrical connections. Similarly, computer ownership has increased as well, although still considerably scarce. This is visible in Figure 2.1b where there is a monumental rise in the ownership of cell phones because of better access. Similarly, not a single household in the sample owned a computer in 2005, despite already having been connected to the grid, while over 1% of the households owned computers by 2012. Therefore, despite being connected to the grid, and having the resources to set up computers at home and use rechargeable cell phones, the lack of access to these devices contributed significantly to opportunity costs, restricting households from making the most of their connection immediately following the grid connection.

The above arguments suggest that although a household's access to electricity may be a shock, benefiting from the connection may be a more gradual and continuous process, with different outcomes possibly being influenced differently, at different paces. Therefore, in our analysis, we work with multiple outcome variables, each of which may have a different relationship with electricity access and thus respond differently. These include variables such as consumption, which could start showing benefits faster as a result of labor participation or increased domestic production, and variables like the education level of the household's decision maker, which one would expect to have a considerably slower response (if any) to electricity access.

2.2 A Potential Resolution to the Experimental-Observational Disparity?

In this section, we take a small digression from our main analysis of the time-associated benefits of electrification. In the previous section, we discussed that studies that investigated the relationship between the duration of connection and household well-being were scarce in the literature. However, through a short meta-analysis, we attempt to draw inferences on the role of time in the frequency of measured positive impacts in impact evaluation studies.

One matter of significant concern in the impact evaluation literature is that experimental studies measure fewer positive results than observational studies. Since randomized controlled trials elude the problem of selection bias in the assignment to treatment and control groups better than any other contemporary research methodology, the disparity in the frequency of positive results could hint at observational studies tending to overestimate results. The disparity has generally been attributed to liberal assumptions in the empirical strategies of observational studies (Lee et al., 2020a). However, another possible reason that may cause experimental studies to find fewer positive impacts may be because experimental studies are typically constrained to be shorter in duration than observational studies. In line with the arguments made in section 2.1, we conjecture that controlling for the duration of connection could potentially resolve (some of) the differences in the outcomes found in observational and experimental studies, as experimental studies are constrained to be much shorter in duration, while observational studies have the advantage of exploiting data spanning several years and multiple panels. If confirmed, this could resolve a considerably important conundrum in the impact evaluation literature while simultaneously lending empirical support to the time-associated differences in the causal effects of electrification.

To test our conjecture, we compare the results of different impact evaluation studies and the duration of the connection in each case. To avoid biases and facilitate reproducibility, we use articles from the systematic review by Bayer et al. (2020) on impact evaluations between 2000 and 2020. Of the 31 studies shortlisted in it, we choose those studies which have data on the duration of the connection. Experimental studies, randomized control trials in particular, typically have information on the duration of the experiment, which serves as a direct measure for the duration of the connection. There are seven such studies. Among observational studies, we select the eight studies which use difference-in-differences (DID) methods, and define the duration of the connection, as half the time between the panels (median time), since there is seldom information regarding the exact time at which the households got electrified. Apart from these, we also use a study that does not use experimental or DID methods, but had data for the duration of the connection, taking the tally of our sample to nine observational studies and 16 studies in total. We observe that among the 16 studies selected, the seven experimental studies had an average duration of connection of 1.61 years, while the nine observational studies had an average duration of connection of 4.77 years, which is nearly three times as large. Furthermore, it can be observed in Figure 2.2 that while studies, where the duration of connection was low, reported both positive and neutral results, studies with long connected times reported only positive results, with the lowest average impacts seen in the sample of studies, increasing as the duration increases. It can also be observed that for studies with a duration shorter than 5 years, positive impacts in observational studies are also considerably rarer.

Among the shortlisted studies, we consider the same five dependent variables used by Bayer et al. (2020), namely, total income/expenditure, savings, energy expenditure, business creation, and education. The authors cite their rationale for choosing these variables as they have been frequently assessed at the household level. We then denote positive impacts as 1, neutral impacts as 0, and negative impacts as -1. Lastly, we average over all the measured outcomes in each study, to avoid biasing our analysis by discriminating between studies and projects and giving more weight to

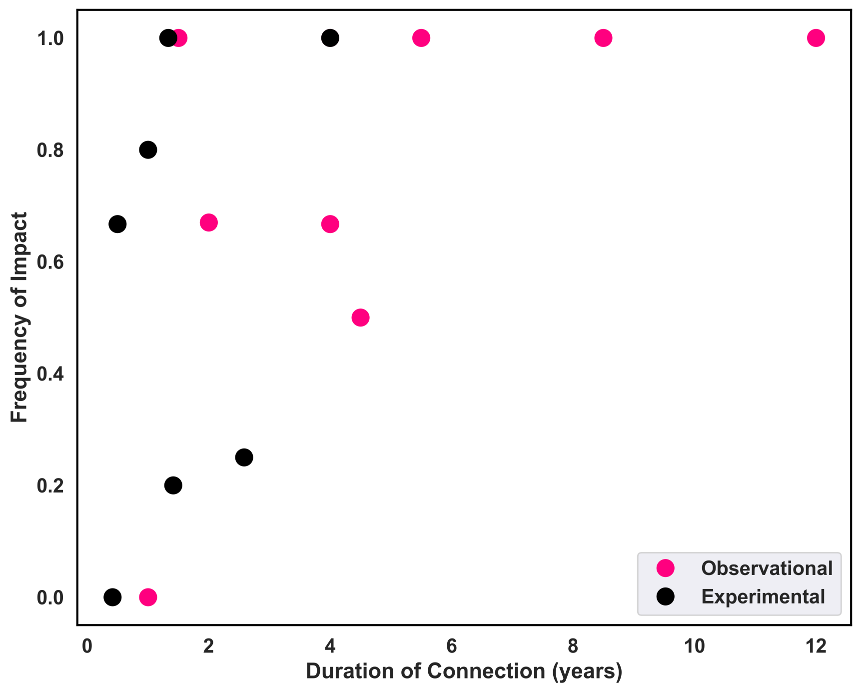


Figure 2.2: The variation in the Frequency of Impact for studies across different Durations of Connection. The impact is 1 for a variable if a significant positive impact was observed on any of the chosen variables, -1 for a negative impact, and 0 for a neutral result. These values are then averaged on a scale from -1 to 1, giving us what we call the Frequency of Impact. For studies with non-negative impacts, this is the equivalent of the frequency of positive impacts. Source: Bayer et al. (2020)

studies that measured more outcomes. We call the averaged quantity the Frequency of Impact, although it is subtly different from the frequency of measuring a non-neutral impact. This is because we have drawn a distinction between positive and negative impacts. For studies with only non-negative impacts, this quantity is the same as the frequency of a non-neutral impact.

To appropriately measure the impact of the duration of time, the methodology (experimental/observational), and the type of connection (grid/off-grid), we employ an ordinary least squares (OLS) regression with the Frequency of Impact as the dependent variable and the logarithm of the duration of the connection, the methodology of the study (observational or experimental), and the type of connection (grid or off-grid) as the dependent variable. The categorical variable for the methodology takes 1 for an observational study and 0 for an experimental one, while the categorical variable for the Type of connection takes 1 for a grid connection and 0 for an off-grid connection.

Outcome: Frequency of Impact	OLS (1)	(2)	(3)	(4)
Intercept	0.4976***	0.5596***	0.5557***	0.5389***
	(0.1078)	(0.1409)	(0.1533)	(0.1464)
Logarithm of Duration of Connection (Years)	0.4905**			0.5755*
	(0.1998)			(0.2876)
Methodology (Observational = 1, Experimental = 0)		0.2009		-0.0128
		(0.1879)		(0.2353)
Type of Connection (Grid = 1, Off-grid = 0)			0.1863	-0.1029
			(0.1939)	(0.2486)

Table 2.1: Effects of the Duration of connection (log), Methodology and Type of connection on the impact measured in Impact evaluation studies, on a sample of 16 impact evaluation studies. The outcome variable measures the Frequency of Impact, which is the average of the sign of impact measured for different outcome variables studied in each impact evaluation study. Each cell presents the coefficient (and standard error) measured in the respective regression.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

We regress each variable alone and all three together. In the final regression involving all three variables, we find in the regression, that the coefficient of the (log) duration of connection is the largest, followed by the intercept, and the coefficients for the methodology and the type of connection are much smaller. Since there are only 16 data points in our sample, we expect the significance to be low. The time variable has a significant impact at the 5% level when regressed alone and at the 10% level when regressed along with the other variables. We do not find methodology or the type of connection to have a significant contribution. This implies that most of the variation seen in the impacts may not be due to the methodology of the study or the type of connection, but rather the duration of the connection and the disparities observed in outcomes between studies that use different types of connection or different methods of analyses are because of a high correlation with the duration of the connection. Interestingly, although not significant, observational studies in isolation seem to have a positive effect on the frequency of results, when controlled for the duration of the connection, the effect even becomes weakly negative.

However, since the number of observations used for the regression is rather small, the results are not meant to completely rule out any impact that methodology or the type of connection may have but to rather check whether the duration of the connection, a variable that has not found much favor in the literature so far, could be an important factor in impact evaluation studies. Nevertheless, the paucity of studies investigating the time-associated benefits of electrification on social and economic development presents a major gap in the literature, which remains to be studied.

2.3 Data and Sample

Reverting back to our main study on the time-associated benefits of electrification, in this section we overview the data and the setting of our problem. Our data comes from a panel of households from three survey waves. The first wave of the sample comes from the Human Development Profile of India (NCAER, 1994), conducted in 1994-1995. The second wave comes from the India Household Development Survey - I (Desai et al., 2005) from 2004-2005, and the third comes from the India Household Development Survey - II (Desai et al., 2011-2012). The original HDPI survey covered a random sample of 33,230 households located in 16 major states, 195 districts, and 1,765 villages. Meanwhile, the IHDS-I survey covered 41,554 households from 384 districts, 1,503 villages, and 971 urban blocks. IHDS-II covers the same villages and urban blocks but interviewed 42,152 households. Some households which had been interviewed in the previous wave could not be contacted in later periods, given that the last round of surveys was 17-18 years after the first. Some other households had split by then.

We restrict our analysis to the set of rural households present in all three surveys, that had not split. This leads us to our final set of 9233 households from 727 villages, with a mean of 12.70 households per village and a standard deviation of 7.05. Since this sample was generated randomly from the original HDPI survey, this sample can be assumed to be representative at the level of the country, even if not at the level of smaller units such as districts and villages. Our analysis covers various household-level and village-level characteristics which we use in our estimation of the impacts of short-term and long-term electrification.

2.3.1 Sample Characteristics

The panel of villages surveyed in the IHDS surveys consists of 1406 villages, but the common panel of households from the HDPI survey in 1994-1995 and the IHDS surveys from 2004-2005 and 2011-2012 are located in 727 villages. A major concern in establishing causality in electricity access and welfare arises from the possibility, that improvements in electricity access often coincide with other development in infrastructure and institutions in the period. This potential cor-

	Means (standard deviation)		Correlation with % of households electrified	
	2004-2005	2011-2012	2004-2005	2011-2012
Presence of metalled roads†	0.67	0.87	0.2694***	-0.3550***
	(0.47)	(0.33)		
Distance to nearest bank branch office or credit cooperative (km)	4.69	5.06	-0.0822**	-0.0529
	(5.46)	(5.08)		
Distance to the closest general market shop (km)	2.23	2.57	-0.0516	-0.0420
	(4.02)	(5.03)		
Presence of Development Group or NGO †	0.14	0.13	-0.0095	-0.1032***
	(0.35)	(0.34)		
Number of Primary Healthcare Centers	0.13	0.11	0.0475	0.0720*
	0.40	0.31		
Number of Government Primary Schools	1.73	1.70	-0.0408	-0.1099***
	(1.62)	(1.60)		
Number of Private Primary Schools	0.64	0.73	0.1496***	-0.0022
	(1.32)	(1.50)		
Number of Government Middle Schools	0.66	0.87	0.1641***	-0.0396
	(0.62)	(0.72)		
Number of Private Middle Schools	0.31	0.51	0.1366***	0.0650*
	(0.73)	(1.40)		
Number of Government Secondary Schools	0.32	0.41	0.1464***	0.0648*
	(0.51)	(0.66)		
Number of Private Secondary Schools	0.18	0.26	0.1593***	0.1204***
	(0.51)	(0.74)		
Number of Government Higher Secondary Schools	0.13	0.17	0.1657***	0.0717*
	(0.34)	(0.42)		
Number of Private Higher Secondary Schools [1]	0.11	0.16	0.0483	0.0756**
	(1.52)	(0.55)		

Table 2.2: Descriptive statistics for village characteristics 2004-2005 and 2011-2012. Includes 727 villages. The data are unweighted. Source: HDPI, IHDS I, and IHDS II surveys. The correlation reported is Pearson's correlation coefficient, and significance has been calculated with the two-sided alternate hypothesis.

† Dummy variable which takes 1 for "yes" and 0 for "no".

[1] The drop in standard deviation is due to one village that reports 40 private higher secondary schools in 2005.

*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

relation could be a result of governments undertaking multiple development activities in regions, one of which might be an improvement in electricity access. Therefore, it is crucial to have information regarding the state of development of other forms of infrastructure and access so that they can be controlled for. In this regard, the IHDS surveys encompass an excellent array of variables that describe the presence of roads, schools, primary healthcare centers, development groups, and the proximity to banks and markets.

We can observe that there is a noticeable improvement in the presence of paved roads, and the number of both government and private schools in villages. However, there is little change in terms of the presence of bank branch offices/credit cooperatives, general market (kirana) shops,

	Means (standard deviation)			
	2004-2005		2011-2012	
	Electricity	No Electricity	Electricity	No Electricity
Number of Adult Men (21+)	1.73	1.46	1.59	1.25
	(1.01)	(0.87)	(0.94)	(0.79)
Number of Adult Women (21+)	1.69	1.43	1.66	1.33
	(0.85)	(0.73)	(0.82)	(0.67)
Number of Adolescent Boys (15-21)	0.40	0.40	0.33	0.32
	(0.65)	(0.64)	(0.59)	(0.57)
Number of Adolescent Girls (15-21)	0.41	0.39	0.33	0.32
	(0.66)	(0.63)	(0.61)	(0.57)
Number of Boy Children (<15)	0.88	1.00	0.62	0.70
	(1.10)	(1.16)	(0.92)	(1.02)
Number of Girl Children (<15)	0.45	0.26	0.44	0.26
	(0.50)	(0.44)	(0.50)	(0.44)
Presence of Water Source inside the House †	0.45	0.26	0.44	0.26
	(0.50)	(0.44)	(0.50)	(0.44)
Presence of Flush Toilets †	0.17	0.03	0.30	0.06
	(0.38)	(0.16)	(0.46)	(0.24)
Presence of Separate Kitchens †	0.61	0.37	0.55	0.24
	(0.49)	(0.48)	(0.50)	(0.43)

Table 2.3: Descriptive statistics for household characteristics 2004-2005 and 2011-2012. Includes 9163 households. The data are unweighted. Source: HDPI, IHDS I, and IHDS II surveys.
† Dummy variable which takes 1 for "yes" and 0 for "no".

and development groups/NGOs. In fact, these characteristics display a marginal decline between 2005 and 2012. Table 2.2 shows the village characteristics for 727 villages from rural India which make our sample. The observation that despite an increase in the rate of electricity access (Table 2.5 and 2.4) several other variables such as the proximity to banks, credit cooperatives, markets, development groups, and primary healthcare centers have not shown improvements, makes this period suitable for our analysis, as they rule out the possibilities of overstating the benefits of electrification by attributing improvements caused due to these auxiliary variables to electricity. We also observe good variation, with high standard deviations (often several times the means) in the variables that do show improvements (primarily the number of schools of different levels), allowing our analysis to leverage this variation in better attributing the benefits of electricity access.

Table 2.3 shows the household characteristics, which describe the number of adult men and women, the presence of water sources inside houses, separate kitchens, and the presence of flush toilets. We can observe that households that do not have electricity, do consistently worse in terms of household infrastructure in both rounds of the survey. Furthermore, the table makes an interesting illustration that households without electricity, (and with poorer household infrastructure on average) tend to be smaller in size, reflected in the smaller number of men and women (adults and children) in these households. This may be because receiving electricity access may be a function of the ability to pay, and larger households would have more disposable income to afford to get connected and acquire the material resources to utilize the connections.

In Table 2.2 and Table 2.3, we have excluded descriptive statistics related to electrification which we present under Subsection 2.3.2, and descriptions of household well-being, which includes variables related to income, agricultural income, expenditure, food security, education, and fossil fuel consumption of households, which we describe in Subsection 2.3.3.

2.3.2 Electricity Statistics - Village and Households

Table 2.4 presents the means, standard deviations, and fraction of households, as appropriate, for different characteristics of household electrification for 2004-2005 and 2011-2012, and Table 2.5 presents the characteristics of village electrification. Among the 9,233 households, 4,746 households (51.40%) already had electricity connections prior to the survey in 1994-1995. Due to differences in variables between the HDPI survey and the IHDS survey, we primarily deal only with the IHDS surveys and only use data on electricity connections from the HDPI survey.

By 2004-2005, 6,256 households (68.27%) out of 9,163 households had electricity (there was no data for 70 households), and 7,766 households (84.39%) out of 9,202 households were connected by 2011-2012, leaving 1436 households (15.61%) unelectrified at the end of the period. Despite the fact that the number of electrified households in 2004-2005 and 2011-2012 were under 70% and 85%, respectively, we find that over 93% of the villages that these households belong to had been connected by 2004-2005 and close to 99% of the villages had grid access by 2011-2012. This implies that even though a village has access to electricity, there may be several households in the village that may not have access to electricity. Furthermore, considering that in over a fifth of the villages in both surveys, every household had electricity, the numbers for several villages are likely well below the mean. This fact will become important in our further analysis when we choose variables that determine whether a household receives electricity or not.

We can observe that the mean reliability and the standard deviation changed little between 2005 and 2012, apart from a small drop in the mean. However, despite the marginal drop in the mean

	Means (standard deviation)			
	1994-1995	2004-2005	2011-2012	Δ
	(1)	(2)	(3)	(3)-(2)
Electrification Access†	0.51	0.68	0.84	0.16
	(0.50)	(0.47)	(0.36)	(0.46)
Reliability (Hours of Access in a Day)[1]		14.79	14.08	-0.45 [3]
		(6.68)	(6.69)	(7.50)
Pay for electricity†[1]		0.86	0.81	-0.03 [3]
		(0.34)	(0.39)	(0.45)
Pay to the company†[1]		0.79	0.72	-0.03 [3]
		(0.41)	(0.45)	(0.52)
Monthly Expenditure on Electricity [2] (Rs.)		348.84	259.60	-56.61 [3]
		(486.03)	(364.82)	(505.90)

Table 2.4: Descriptive statistics for household electrification 1994-1995, 2004-2005 and 2011-2012. Includes 9,233 rural unsplit households. The data are unweighted. Source: HDPI, IHDS I, and IHDS II surveys.
[1] Among those connected, [2] Among those who pay for the connection, [3] Households already connected by 2004-2005. † Dummy variable which takes 1 for "yes" and 0 for "no"

reliability, the standard deviation in the change of reliability for households connected by 2005 is unusually large, illustrating that some households with very poor reliability could have seen their reliability improve considerably, and households with good reliability may have seen a decline in quality.

On comparing electrical reliability at the level of households and villages between Table 2.4 and Table 2.5, we observe that although the numbers for 2011-2012, both mean and standard deviation are similar, there is a small difference among the 2004-2005 versions of the measures, likely due to measurement errors. On comparing the full distribution of reliability (Figure 2.3), we observe that the distributions for households and villages look rather similar. This resemblance in the distributions is critical as it indicates that variations in reliability, and possibly some other components of electrical connections such as the capacity available, happen largely at the level of the village and could be exogenous to household characteristics. Since reliability is the only variable that is measured both at the level of the household and the village, we cannot test this for other components such as frequency of disruptions or capacity, for which we do not have data.

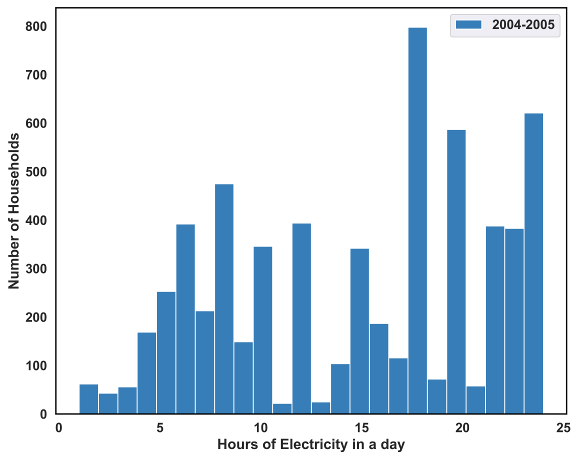
Interestingly, the average household expenditure on electricity drops considerably (29.17%), which could imply that poorer households had received access to electricity. It is also evident that fewer households were paying for their electricity connection to electricity distribution companies,

	Means (standard deviation)		
	2004-2005	2011-2012	Δ
	(1)	(2)	(2) - (1)
Village has access†	0.93	0.99	0.06
	(0.25)	(0.10)	(0.24)
Electrified Households in Village (%)	68.63	79.25	10.66
	(34.13)	(26.80)	(30.66)
Village fully electrified†	0.22	0.23	0.01
	(0.41)	(0.42)	(0.47)
Years since First Connected	25.37	31.16	4.89
	(15.54)	(15.84)	(14.87)
Reliability (Hours of Access in a Day)	13.02	13.68	0.73
	(7.05)	(6.68)	(7.56)

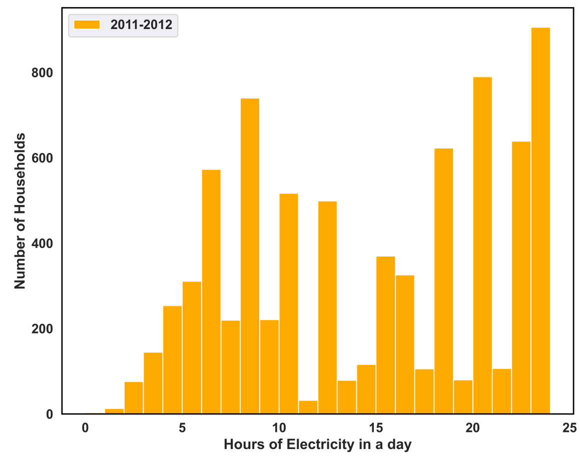
Table 2.5: Descriptive statistics for village electrification 2004-2005 and 2011-2012. Includes 727 villages. The data are unweighted. Source: HDPI, IHDS I, and IHDS II surveys.
† Dummy variable which takes 1 for "yes" and 0 for "no".

		Means (standard deviation)		
		Electrified Households in Village (%)		
		2004-2005	2011-2012	Δ
		(1)	(2)	(2) - (1)
Population of Village	More than 5000	68.12	76.57	8.45
		(33.19)	(28.00)	(32.21)
	1001-5000	70.30	79.02	8.86
		(32.97)	(27.31)	(29.13)
	Less than 1000	65.18	81.79	16.50
		(36.87)	(24.39)	(32.41)

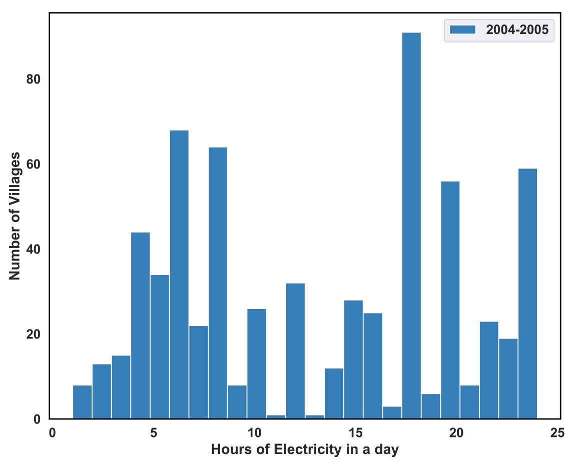
Table 2.6: Variation in village electrification with size 2004-2005 and 2011-2012. Includes 725 villages. The data on sizes are from IHDS I, which uses the 2001 census (village). The data are unweighted. Source: HDPI, IHDS I, and IHDS II surveys.



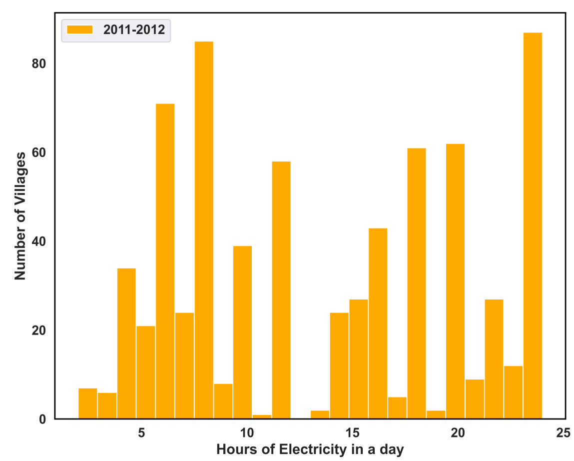
(a)



(b)



(c)



(d)

Figure 2.3: Distributions of electrical reliability for households (2.3a) 2004-2005, (2.3b) 2011-2012, and villages (2.3c) 2004-2005, and (2.3d) 2011-2012, for unsplit rural households and villages common to three surveys. The data are unweighted. Source: HDPI, IHDS I, and IHDS II surveys.

and a larger fraction of households were connected who did not receive bills or were connected by government schemes. Furthermore, since the average decrease in payers among those who were already connected to the grid is lower than the decrease overall, a large fraction of new connections may have been connected by government schemes, which makes this period, ideal for impact evaluation.

Table 2.6 also suggests that the size of the village (in terms of population) may have had a role to play in which villages were connected. While there is no observable trend in the relationship between whether a village was already connected by 2005 and the size of the village, the percentage increase in electrified households in small villages (with a population of less than 1000 in the 2001 census) was almost twice as much as the increase in bigger villages. This becomes crucial in identifying the factors which could have influenced whether a household was more likely to receive treatment in this period, and thus help address selection bias in our estimates in subsequent sections.

2.3.3 Characterizing Household Well-being

In this study, we attempt to study the impact of electrification on household well-being. Well-being is not restricted to income and wealth alone, and there are several components of well-being apart from income and expenditure that are relevant to the differential lives and livelihoods of households (Sumner, 2004). From Amartya Sen's influential work on the role of capabilities (Sen, 1985, 1995, 1997) to the United Nations Sustainable Development Goals, the notion of economic as well as non-economic well-being and the importance of addressing multidimensional poverty has become commonplace in the social sciences. To this end, we do not study well-being as households attaining the 'means' to their desired resources alone, but whether households do in fact achieve the 'ends.' Access to electricity could play a pivotal role in facilitating the non-economic aspects of well-being. For instance, a child in a household without electricity, may not have the opportunity to study after dark, even if they are well-off. If the child does study in the dark, they would have to do so under fossil-fuel-based lighting such as kerosene lamps, which may expose the student to illnesses of the eyes and the respiratory tract. Access to electricity could thus enhance the opportunity for households to realize several well-being goals that they may not have been able to achieve with economic progress alone.

In order to quantify well-being, outcomes can be studied individually or together as an index. We choose to study the impact on variables independently as opposed to studying a single index, as an index may be sensitive to changes in only one of many variables, and it may not be an appropriate estimate for impacts on well-being or multidimensional poverty (Ravallion, 2011). Furthermore, since we are interested in the time-associated benefits of electricity access, studying variables separately allows us to identify which variables respond faster and slower to electrification. The main outcome variables in our study are household-level indicators such as consumption (log), agricultural income (log), agricultural land ownership (log), years of education of the highest educated adult, domestic kerosene consumption, and the time spent collecting fossil fuels. These variables cover several economic and non-economic dimensions of well-being.

Consumption (Expenditure) is a well-known descriptor of income. Even though the IHDS has data on the family's income in the year preceding the survey, this may not be a particularly accurate measurement. This is because the survey allows both positive and negative responses for income. A negative income would be in the context of losses incurred in the year or in terms of debt, which may be anomalous, and not all households may report debt or losses as negative income, i.e., in the context of business, whether a household describes income as profits or revenue may vary. Furthermore, since we have two panels spread over seven years, we cannot ignore the possibility of an anomalously good or bad year for the household in terms of income. Consumption, on the other hand, is a more consistent measure and more invariant to anomalous events and short-term fluctuations. Furthermore, due to consumption smoothing, consumption is not only a measure of the household's economic well-being in the year of the survey but also of the years prior to it. To study the impact on agriculture, we study agricultural income. Although agricultural income is associated with the same problems as income, there is no better alternative in this case. India is one of the largest cultivators in the World, and the impact of rural electrification on agriculture is both in terms of an improvement in agricultural productivity of the household and in terms of diversification of livelihoods. We also study the impact on agricultural land ownership. In rural India, and other agrarian economies, the ownership of agricultural land, apart from being a key dimension of natural capital, is a good proxy measure of household wealth, and thus, it is one of our three outcome variables for economic-well being.

Among non-economic indicators of well-being, we study the effect of electrification on education, fossil fuel consumption, and the time spent collecting biomass. To study the impact on education, we study the years of schooling of the most educated adult. It is a variable that is extremely unlikely to respond to electrification in the short term, as the beneficiaries of the effects of electricity access on schooling are likely to be young students and thus it may take years before

	Means (standard deviation)			
	2004-2005		2011-2012	
	Electricity	No Electricity	Electricity	No Electricity
(Log) Per-capita Consumption (Rs.)	9.99	9.53	10.18	9.76
	(0.63)	(0.52)	(0.62)	(0.56)
(Log) Agricultural Income (Rs.)	10.08	9.21	10.09	9.12
	(1.53)	(1.44)	(1.58)	(1.40)
(Log) Agricultural Land Holding (acres)	1.64	1.68	1.63	1.34
	(1.27)	(1.49)	(1.35)	(1.35)
Schooling of the Highest-educated Adult (years)	7.28	4.08	7.62	3.80
	(4.82)	(4.37)	(4.94)	(4.39)
Kerosene Consumption (Litres)	3.33	3.64	2.74	3.13
	(3.01)	(2.12)	(2.29)	(1.54)
Time spent by Women in Collecting Fuel (mins per week)	293.25	264.68	238.39	312.17
	(296.96)	(270.50)	(387.17)	(424.77)

Table 2.7: Descriptive statistics for household well-being characteristics 2004-2005 and 2011-2012. Includes 9163 households. The data are unweighted. Source: HDPI, IHDS I, and IHDS II surveys.

these students become adults, making it a particularly interesting variable to explore in the analysis of time-associated benefits. The study of domestic fossil fuel (in our case, kerosene) consumption addresses a different aspect of well-being - related primarily to healthcare. However, it also answers a question related to energy consumption and fossil fuel dependency. If electricity were to replace domestic fossil fuels for the purposes of heating and lighting, the carbon footprint of the household may reduce if the electricity is produced using clean and renewable sources. However, it is entirely possible that energy consumption through electricity may not displace domestic fossil fuel consumption, and may only augment it in the household's energy consumption. Apart from domestic fossil fuel consumption, households also burn biomass for heating, cooking, etc. and electricity could potentially displace their usage. This would not only result in healthcare benefits for the household but may potentially provide its members with more hours saved from the burden of collecting biomass, which could result in greater hours of leisure, higher participation in labor markets, or better time allocation to other domestic activities.

Table 2.7 presents the main outcome variables used in our analysis to study, and their means and standard deviations. We observe a difference in the levels of these variables between households with and without electricity. This difference is particularly pronounced in 2011-2012. Agricultural land holding which does not show much difference between households that have and don't have access to electricity in 2005 also differs considerably between the two groups in 2012. In addition to differences between electrified and unelectrified households within each period, there are also changes between periods. There is typically a small increase in the levels of variables between the two surveys, particularly among households that have electricity, or a decrease in the cases when a decrease is desirable such as in kerosene consumption or time spent in collecting fuels, with the only exceptions being agricultural land holding which shows a marginal decreases. On the contrary, the corresponding comparison for households without electricity reveals that households without electricity were typically worse off in 2012 than in 2005, with the exceptions of total expenditure and kerosene consumption which increased and decreased, respectively. This may be because only the poorest still didn't have access to electricity by 2012.

The reason for the stark difference between the two households, and particularly the reason for the deplorable circumstances of unelectrified households in 2012, may hint towards poorer or less developed households getting electrified last, which results in the disparity observed in Table 2.7. Nevertheless, it is also possible that some of this disparity is due to the benefits of being electrified. In the subsequent sections, we describe our methodology and results for quantifying the causal impacts of short-term and long-term electricity access on household well-being. In the tables above, we see that there are considerable differences among households in multiple variables. For instance, in 2012, households without electricity access had much lower land holdings, but these households were also less likely to have flush toilets and separate kitchens. Therefore, these composition effects need to be controlled while studying the outcomes, and we cannot simply look at electricity access. In order to study the composition effects, we employ a difference-in-difference approach, with a large set of control variables, to attribute the changes in the levels of well-being appropriately to electricity access. The empirical strategy in the estimation of benefits is described in detail in the following section.

2.4 Empirical Strategy

Since we are interested in the estimation of time-associated benefits of electrification, our analysis requires the knowledge of when a household was connected or how long a household has had access to electricity. While the two panels of the IHDS data have information on when villages were first electrified, there is no data on when individual households were first connected to the grid. In order to get this information, we leverage the households common to the HDPI and IHDS surveys. These surveys comprise over 9000 households from over 700 villages in India, interviewed thrice over a period of 17-18 years. Since all three surveys ask households whether they have electricity, we know the interval during which households were connected, i.e., between which two survey rounds (at least for households that were not yet connected by the HDPI survey). We consider those households which did not have electricity in 1995. These households can be classified into three groups - those who did not have electricity in 1995 and continued to not have electricity in 2005 and 2012, those who did not have electricity in 1995 but got connected before 2005, and those households who did not have electricity in 1995 or 2005 but got connected by 2012. These three groups make our control group, long-term treatment, and short-term treatment groups, respectively. Our aim is to study whether the treatment groups perform differently compared to the control group, and how each of the treatment groups performs vis-a-vis the other.

At this juncture, it is important to explain why we only consider households that have not been electrified by 1995. This is different from previous impact evaluation studies such as the one by Samad and Zhang (2016) which grouped households that had electricity prior to the period of consideration with the control group. This is because our premise is centered around the possibility that a long-term electricity connection may have persisting causal effects on household well-being. If we cannot rule out the possibility of long-term benefits, then grouping households that have had electricity for decades with the control group, i.e., households that do not have electricity may lead to incorrect estimates of the impact measured. In other words, we postulate that simply having electricity can lead to causal impacts, and it is not only the shock of getting connected, which could bring benefits to the household. Even heuristically, the essence of impact evaluation studies is to ask whether connecting households to the grid brings any causal benefits that the household would not have received, had it not been connected. Thus, the counterfactual of a household being connected would be a similar (ideally, identical) household that did not receive the connection, rather than a household that has always been connected. If households that have had electricity for long periods are made part of the control group, the research question changes slightly, as then we are asking what is the causal effect of a household not having been connected a few years ago, compared to a household that was.

Let y_{it} be the outcome variable (such as the logarithm of total expenditure), for household i at time t. We construct two indicator variables T_L and T_S for long-term and short-term treatment, where the former takes 1 if the household had access to electricity in 2005, and 0 otherwise, while the latter is 1 if the household had access in 2012, and 0 otherwise. Therefore, the control group (those who still did not have access in 2012) takes 0 for both indicator variables. Let X_{ijt} denote the vector of community- (village) level and household-level characteristics for household i, in village j at time t. Hypothetically, if all households were equally likely to be connected in each period or not be connected at all, then we could estimate the effects using the following linear model:

$$y_{ijt} = \alpha_0 + \alpha_t + \beta_L T_{L,ijt} + \beta_S T_{S,ijt} + \gamma.X_{ijt} + \mu_i + \nu_j + \epsilon_{ijt} \tag{2.1}$$

where α_t is the time-fixed effect felt by all households at time t, β_L and β_S are the coefficients of effects for having access to electricity, respectively, γ is the vector of coefficients for the control variables, μ_i are the household-specific effects of household i, ν_j are the village specific effects of village j, and ϵ_{ijt} is the idiosyncratic error term. By taking the difference of the levels of y_{ijt} across the two panels ($\Delta y_{ijt} = y_{ijt} - y_{ijt-1}$), we can eliminate α_0, μ_i, and ν_j.

$$\Delta y_{ijt} = (\alpha_t - \alpha_{t-1}) + \beta_L \Delta T_{L,ijt} + \beta_S \Delta T_{S,ijt} + \gamma.\Delta X_{ijt} + \Delta\epsilon_{ijt} \tag{2.2}$$

Note that by differencing T_L and T_S, the variables now constructed represent whether a household was connected in the short term or the long term, and not whether a household has access in the period. Specifically, this means that $T_{L,ij2012}$ and $T_{S,ij2012}$ would both be 1 for households that were connected between 1995 and 2005. However, after differencing, $\Delta T_{S,ij2012}$ would be 1 only for households that were connected between 2005 and 2012, and 0 for households connected between 1995 and 2005, and only $\Delta T_{L,ij2012}$ would be 1 for these households. Thus β_L and β_S measure the average treatment effects (ATE) for long-term and short-term connections, respectively. Although for this to be true, we exclude a small set of households that had electricity in 2005 but lost connection by 2012. This exclusion is in line with our research question since we are interested in studying the effect of connecting households, and there is no real basis for assuming the effect of losing access will be the exact opposite of getting access. Since we can measure $\Delta T_{L,ijt}, \Delta T_{S,ijt}$, and ΔX_{ijt}, we can arrive at an estimate of β_L and β_S using the difference-in-differences (DID) regression. However, these would be the ATEs only under the assumption that each household was equally likely to be assorted into the two treatment and control groups, which is an improbable assumption.

Thus by using a simple DID regression, we would fall prey to endogeneity arising from selection bias in assignment to treatment. For instance, more developed villages may be more likely to be connected to the grid earlier. Similarly, at the household level, wealthier or better-educated households may be connected preferentially because of their economic and social capital. To find the true ATE, we should control for the selection bias in assignment to treatment. In a randomized control trial, we would have been able to adjust for this by selecting a truly random sample of households to be treated. But since we work with a natural experiment, we are required to adjust the sample retrospectively. For this, we use propensity-score-weighted-regressions, using the generalized propensity score (Imbens, 2000).

An alternate method to using propensity score weighted regressions would have been to use an instrumental variable. Common instrumental variables have been geographic cost-based instruments (Dinkelman, 2011), and power source proximity-based instruments (Van de Walle et al., 2017; Thomas et al., 2020; Handayani et al., 2023). Bensch et al. (2011) and Lee et al. (2020a) argue that geographic cost-based instruments may violate the exclusion restriction necessary in instrumental variable designs, particularly because the land gradient may be correlated with other factors like the development of roads and other infrastructure, which may themselves impact economic growth and well-being of households in the village. Power source proximity-based instruments assume that households or villages close to the power source may be more likely to be connected earlier due to lower costs. A proximity-based cost instrument in terms of household adoption choice has also been used by Khandker et al. (2009, 2012). A problem with such measures is that households that are not electrified could also benefit from the spillover effects of proximity to a power source, such as benefits through employment in power plants or through spillover effects of other houses in the village or neighborhood being electrified. Another possible instrument might be a dummy variable for whether the village is connected, under the assumption that villages are connected randomly. Such a variable would be correlated with whether a household is connected and would not be a "weak instrument". However, such a measure also assumes that there are no spillover benefits. Van de Walle et al. (2017) find using panel data from India that households that are themselves not connected do not benefit from village electrification unless they are themselves connected. However, using data for what fraction of a village is connected, we find that the electrification of other houses in a village has a positive impact on the consumption of households that are themselves not connected .

The propensity score is a measure that quantifies the probability of a household being assigned to the treatment that it received. We would like to balance our sample such that every household was equally likely to receive treatment. And thus by weighting each observation by the inverse of the propensity score, we can arrive at a sample corrected for selection bias. We can illustrate this using a simple example. Suppose there is only one type of treatment, receiving access to electricity, and suppose there are two kinds of households - rich households and poor households, with the sample having an equal number of both types of households (for simplicity). Suppose we found that rich households were twice as likely to get connected, compared to poor households.

Therefore, the proportion of rich households in the treatment group would be double what it is in the population. This would lead to incorrect ATE estimates. To correct for this, we weight each observation with the inverse of its propensity score, i.e., since rich households are twice as likely to receive the treatment we assign a weight of $1/2$ to the observations of rich households. This way, the sample would be corrected for selection bias, and it would be the equivalent of there being an equal number of rich and poor households in the treatment and control groups, despite the actual sample being balanced.

Therefore, our empirical strategy reduces to a two-stage problem. In the first stage, we estimate the propensity scores for each household being assigned to either control, short-term treatment, or long-term treatment. For this, we use a multinomial logistic regression. For determinants of which type of treatment (or lack thereof) a household receives, we use attributes from the 2005 panel. While some households (those with long-term treatment) had already been assigned to treatment by this time, propensity scores can still be calculated. This is because these scores need not be the actual metrics used by governments and agencies in assigning households to treatments. Instead, it is the actual variation in initial observable characteristics between households. We use two setups for calculating propensity scores. In one setup, we assume there is a selection bias at the level of the village, but assignment to households within villages is random. For this, we use village-level characteristics such as the population of the village, caste composition of the village, the fraction of households in the village that have electricity, the number of years since the village was first connected to the grid, the presence of paved roads, proximity to bank branch offices/credit cooperatives, general market ships, presence of NGOs/development organizations, the number of primary health care centers, and the number of primary, middle, secondary and higher secondary schools - both private and government. In the second approach, we assume endogeneity at the level of village assignment and household assignment. Therefore, we use the village-level determinants listed above but also use household-level determinants - the number of adult men, the number of adult women, the presence of a water source inside the house, the presence of flush toilets, the presence of separate kitchens, and the total consumption (log) of the household.

We then use these propensity scores to perform a weighted least squares regression in the second stage. In the second stage, we use equation 2.2 to perform a propensity-score-weighted regression to calculate the average treatment effects. The outcome variables we use are consumption (log), agricultural income (log), agricultural land ownership (log), years of education of the highest educated adult, number of meals eaten in a day, domestic kerosene consumption, and the time spent collecting fossil fuels. And the control variables used in the regression are the number of hours of electricity available to a household in a day, the number of adult men in the household, the number of adult women in the household, whether the household has a water source, flush toilet, and separate kitchen, the population of the village, the presence of paved roads, proximity to bank branch offices/credit cooperatives, general market ships, presence of NGOs/development organizations, the number of primary health care centers, and the number of primary, middle, secondary and higher secondary schools - both private and government. Section 2.5 presents the results of our analysis.

2.5 Results

2.5.1 Estimates of Short-Term and Long-term Impacts

The results of the regressions are presented in Tables 2.8-2.13. Each table presents the results of three regressions. The three regressions are the simple DID regression, along with two propensity score-weighted-DID regressions, where the first regression uses propensity scores calculated using only village-level determinants, and the latter estimates propensity scores using both village and household characteristics as determinants. Our preferred estimate is the last one. The tables present the estimated coefficients and the robust standard errors clustered at the level of the village for the two types of treatment variables - short-term electricity access, and long-term electricity access, along with the change in the reliability of electricity, which despite being a control variable is useful to present, as it is a component of the electricity available to the household.

For most of our variables, there are considerable differences in the estimates, which highlight the importance of correcting for selection bias in assignment to treatment. The difference, in most cases, is particularly remarkable between the simple DID and the propensity score-weighted-DID regressions. There is also a considerable difference in the estimates of the regressions weighted by village-level propensity scores and village and household-level propensity scores, which suggests that household-level characteristics play a crucial role in determining which households within a village are more likely to be connected to the grid.

We find that electricity access has both short-term and long-term benefits on consumption. Interestingly, while the simple DID finds the short-term effect to be statistically significant and similar in size to the long-term effect on per-capita consumption, after correcting for selection bias

	Δ Logarithm of per-capita Consumption (Rs.)		
	Simple DID regression	p-Weighted DID regression	
		village-level p-score	village & household-level p-score
$N = 4026$	(1)	(2)	(3)
Intercept	0.2124 ***	0.2446***	0.1843***
	(0.0360)	(0.0417)	(0.0444)
Short-term Electricity Access	0.0833**	0.0612	0.0675
	(0.0403)	(0.0457)	(0.0526)
Long-term Electricity Access	0.0875**	0.0509	0.1678***
	(0.0354)	(0.0393)	(0.0439)

Table 2.8: Time-associated causal effects of electrification on consumption. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

	Δ Logarithm of Total Household Agricultural Income (Rs.)		
	Simple DID regression	p-Weighted DID regression	
		village-level p-score	village & household-level p-score
$N = 2031$	(1)	(2)	(3)
Intercept	-0.0315	0.2118	0.2592
	(0.0949)	(0.1942)	(0.2252)
Short-term Electricity Access	0.2817**	0.0182	0.0510
	(0.1181)	(0.1881)	(0.1991)
Long-term Electricity Access	0.1199	-0.0988	-0.0982
	(0.0928)	(0.1579)	(0.1788)

Table 2.9: Time-associated causal effects of electrification on agricultural income. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

	Δ Logarithm of Total Agricultural Land Held by the Household (acres)		
	Simple DID regression	p-Weighted DID regression	
		village-level p-score	village & household-level p-score
$N = 2119$	(1)	(2)	(3)
Intercept	-0.2536	-0.2451	-0.2459
	(0.1658)	(0.1543)	(0.1569)
Short-term Electricity Access	0.1409	0.1544	0.1418
	(0.1526)	(0.1543)	(0.1542)
Long-term Electricity Access	0.1507	0.0233	0.0890
	(0.1690)	(0.1677)	(0.1722)

Table 2.10: Time-associated causal effects of electrification on agricultural land holdings. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.

*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

	Δ Years of Education of the Highest Educated Adult in the Household		
	Simple DID regression	p-Weighted DID regression	
		village-level p-score	village & household-level p-score
$N = 4018$	(1)	(2)	(3)
Intercept	0.6138***	0.5550***	0.5184***
	(0.1353)	(0.1457)	(0.1811)
Short-term Electricity Access	0.2830	0.2306	0.3074
	(0.1900)	(0.1975)	(0.2158)
Long-term Electricity Access	0.4104***	0.4655***	0.4819**
	(0.1585)	(0.1681)	(0.1970)

Table 2.11: Time-associated causal effects of electrification on education. Robust standard errors clustered at the village level. The control variables are reliability, the partial sum of reliability, households electrified in the village, the partial sum of households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.

*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

	Δ Kerosene Consumption per month (litres)		
	Simple DID regression	p-Weighted DID regression	
		village-level p-score	village & household-level p-score
$N = 4024$	(1)	(2)	(3)
Intercept	-0.1814	-0.1200	-0.2927**
	(0.1259)	(0.1308)	(0.1453)
Short-term Electricity Access	-0.0592	-0.0397	0.0022
	(0.1541)	(0.1582)	(0.1665)
Long-term Electricity Access	-0.1754	-0.2068	0.0821
	(0.1329)	(0.1283)	(0.1414)

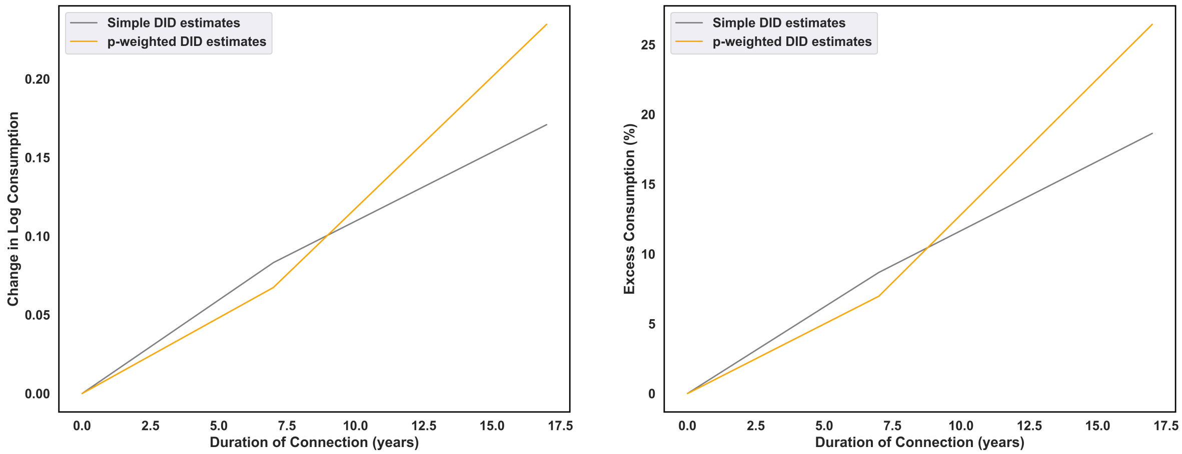
Table 2.12: Time-associated causal effects of electrification on kerosene consumption. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

	Δ Time Spent by Women in Collecting Fuels (minutes/week)		
	Simple DID regression	p-Weighted DID regression	
		village-level p-score	village & household-level p-score
$N = 1805$	(1)	(2)	(3)
Intercept	94.1467***	81.1999**	92.8710**
	(35.2429)	(37.7452)	(39.4325)
Short-term Electricity Access	-56.4753	-21.1919	-35.3700
	(47.0229)	(46.7489)	(51.2091)
Long-term Electricity Access	-118.5650***	-100.5928***	-115.5688***
	(36.4354)	(37.9802)	(40.5683)

Table 2.13: Time-associated causal effects of electrification on the time spent by women in collecting firewood. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

at the level of the village and household, we find that long-term electricity access has an impact (significant at the 1% level) almost two and a half times as large as the short-term impact which is not statistically significant. This implies that the consumption benefits of electrification against time may be a convex function and continue to grow for a considerable number of years (Fig 2.4) after a household receives a connection. A major reason for our final regression showing larger long-term benefits may be that poorer households are less likely to get connected to the grid earlier, and therefore, they are weighted more heavily in the second weighted regression. Poor households are also likely to be slower to extract the full benefits from electricity as they may take longer to be able to afford the necessary appliances and capital to make optimal use of the electricity. Consistent with this postulate, we do not find a large significant impact of long-term connection when household income is not a determinant in estimating propensity scores, which is the case in our first weighted DID regression. This particular result contrasts our findings with other studies in India using the HDPI-IHDS data such as those by Chakravorty et al. (2014) and Samad and Zhang (2016) both of whom find a larger effect in short-term gains in affluence (Chakravorty et al. (2014) use income, not consumption) after controlling for endogeneity, although neither of these two studies looks at long-term effects.

In contrast, our other two economic indicators do not show any statistically significant improvement. While agricultural income shows short-term gains with electricity access in the simple DID, neither agricultural income nor agricultural land-holding shows any statistically significant change with electricity access when weighted by the inverse of the propensity scores, either in the long-term or the short-term, in line with the findings of Samad and Zhang (2016). However, it is noteworthy to mention that agricultural income does show a considerable decrease in the long run (nearly 10%) but is also accompanied by a large standard error which may imply that there are variations in how agricultural income changes for households. Similarly, land-holding shows a considerable increase (over 14%) in the short term but the variation is too high to make this effect statistically significant. The fact that consumption grows, but agricultural income does not may also hint that electricity access enables households to diversify their earnings and pursue economic activities apart from agriculture. Among non-economic indicators of well-being, the education of the highest educated individual shows a significant improvement in the long-term, but no short-term improvement, which can be intuitively explained by the fact that households that received electricity less than seven years before the survey date, experienced the benefits of electricity access on schooling only for that long a period, and since full schooling takes longer than a decade, the long-term benefits are likely to be much more pronounced. While there seems to be a significant reduction in kerosene consumption in the long term after correcting sampling



(a) (b)

Figure 2.4: Short-term and long-term growth of per-capita consumption (a) Changes in Log consumption (b) Excess consumption over the control group. We observe that while the simple DID estimates predict a concave growth in consumption due to electricity access with the effects waning over time, the propensity score-weighted-DID estimates predict a convex growth function with benefits increasing with the duration of the connection. The coefficients used are from Table 2.8.

bias due to village characteristics, the effect disappears after correcting for household characteristics. This is an important result as it signifies that electrical connections may not displace domestic kerosene consumption and households may stack fuel and increase their energy consumption after being connected to the grid. This may be because while households can use electricity for less power-demanding activities such as lighting, they may not have connections strong enough for more energy-intensive activities, and may resort to using fossil fuels. Households also may not have the necessary appliances to displace kerosene with electricity. For instance, households may not have electrical immersion rods for heating water and may be using stoves in their absence, thus demanding fuels such as kerosene or petroleum gas. However, getting access to electricity does reduce the time spent by women in collecting fuels significantly (significant at the 1% level), which means despite not reducing the consumption of kerosene, households were most likely reducing the consumption of other fuels such as firewood, collecting which is, in general, a time-consuming activity, relieving women of a significant burden, allowing greater time available for labor participation, leisure, etc.

At first glance, some of our results may seem at odds with those of Samad and Zhang (2016), even though we use the same survey, and the same period, and we both use propensity-score-weighted-fixed-effects regressions to control for selection bias. For instance, Samad and Zhang (2016) find significant short-term improvements in consumption, larger than the simple fixed effects results, while we do not. They also find a considerable reduction in kerosene consumption, which we again do not find after controlling for household-level characteristics. However, there are several differences between our samples and methods that may lead to these outcomes. Firstly, Samad and Zhang do not study both short and long-term access as we do and thereby are not restricted by the set of households that are common between three panels, allowing them a sample of over twenty thousand households as opposed to our sample which is shy of ten thousand. Furthermore, as discussed earlier, they also include households that were already connected to the grid prior to the first round of the survey in the control group, while we only consider households that had not yet been connected by the time of the first round of the survey. Thirdly, they do not use affluence (either income or consumption) as a determinant while calculating propensity scores, while we do (we use the logarithm of household consumption at the time of the first round of the survey). In fact, our results sometimes agree when we only correct using village characteristics. This can be seen in Table 2.12 where we observe a considerable reduction in kerosene consumption in the long term when only village-level characteristics have been used in the estimation of propensity scores (although this is not statistically significant) and the effects only disappear after household characteristics (including affluence) are used in estimating propensity scores. This is because, in our first stage regression , we find that the logarithm of the consumption of the household is extremely statistically significant, and thus using it while controlling for selection bias removes the effect. This may imply that households reduced kerosene consumption, not because they were connected to electricity, but because they may have switched to more expensive fuels such as LPG, or in the event that they were connected, they could afford expensive electrical appliances to be able to substitute kerosene. Similarly in the case of consumption 2.8 when affluence is not corrected for, the difference between short-term benefits and long-term benefits is greater than in the simple DID case, with short-term being larger and more significant (Column (2) of Table 2.8), and since long-term connections (which make up more than half their sample) are clubbed with the control group in their study, the improvements in their estimated short-term benefits are likely to be more pronounced. This incoherence between our results and those of Samad and Zhang (2016) may also be explained by controlling for affluence in the first stage. As discussed in Section 2.1, poorer households are more likely to take longer to realize the full impacts of electrification, as they may not be able to afford all necessary electrical appliances instantly. Since poorer households are less likely to get connected than rich households, they are given larger weights in the regression, and this leads to an improvement in the long-term impact, as poor households are more likely to benefit in the long term compared to the short-term.

Interestingly, in the case of years of schooling, the propensity-score-weighted regressions estimate higher improvements than the simple DID, akin to consumption. The fact that some variables do better after accounting for selection bias implies that the households which benefit the most after electrification are the ones that are, unfortunately, least likely to be electrified. This has important policy consequences and indicates that progressive electrification programs which target poorer households may yield larger improvements in household well-being and human development, in the absence of which and a continuation of the contemporary paradigm of electrification, the biggest potential beneficiaries will be deprived of the opportunity to enhance their lives and livelihoods.

2.5.2 Heterogeneous Effects: Regional Variations in Impacts of Electrification

In the period of our analysis, India was home to the second largest number of people in the world, with the seventh largest geographical spread, with over 30 states and union territories, each with its own natural resources, vegetation, climate, state of urbanization, and cultures. Therefore, despite observing a consistent pattern of larger and more significant long-term impacts at the national level across all three variables which show any impact at all, at the national level, the impacts and thus the association between the time since the connection and the impacts felt may be different in different geographical regions, and the national estimates may not necessarily be representative of the benefits received in different parts of the country.

In order to study whether the impacts of electrification on the short-term as well as the long-term are felt uniformly across the country, we focus on two groups of states. First, we use the large states of central and northern India acronymized "BMRU", for Bihar, Madhya Pradesh, Rajasthan, and Uttar Pradesh, referring to the states as they were called before some of them were split into two states in the year 2000. Despite the split, all states except Uttarakhand (formed out of Uttar Pradesh) are relatively large and continued to be laggards in terms of several development indicators (Ghosh, 2006). Thus we use the set of Bihar, Chattisgarh, Jharkhand, Madhya Pradesh, Rajasthan, and Uttar Pradesh as one region of study. In contrast, we study states in Southern India - Karnataka, Kerala, Tamil Nadu, and the erstwhile united Andhra Pradesh, along with Maharashtra, who are leaders in economic and non-economic well-being. Apart from similarities in the extent of development in these states, the states in each group also largely enjoy geographical and cultural proximity, which may play into the impacts of electrification.

The notion that different regions within a country may benefit differently from electricity connections is based on the broader literature on impact evaluation. At the Global level, studies in South Asia find much more promising results than those in Sub-Saharan Africa. One of the early studies discussing the weaker results of impact evaluation in Africa was by Peters and Sievert (2016), who argued that while studies in other parts of the World, particularly South Asia yielded promising results in impact evaluation studies of electricity access, the association between electricity access and classical poverty indicators may not be transferable to the African continent due to a variety of factors such as a dearth of access to markets and a shortage of labor demand. Since then, systematic reviews by Hamburger et al. (2019) and Bayer et al. (2020) have only confirmed the disparities in estimated impacts. Since the so-called "BMRU" states differ from the Southern States considerably in terms of economic and non-economic well-being, they may face similar problems such as those argued in the context of Africa. In addition to economic impacts, other impacts such as those on schooling may be hindered by poorer education infrastructure, or those on landholding may be affected by the kind of land-owning and leasing regulations of the state.

To study the variations in impacts between the two regions, we select the sub-samples by filtering according to the states. The sample of "BMRU" states is larger for two reasons. Firstly, the combined population of these states is larger than that of the Southern states, and thus the HDPI-IHDS surveys have more households from these states. Additionally, the Southern States have a greater fraction of households connected to the grid by the time of the 1994-1995 survey, and since we only consider households that hadn't received access by then, we are left with a smaller number of households in South India. In terms of outcome variables, we restrict our analysis to consumption, education, and kerosene consumption. This is because these outcome variables have data for a larger number of households, unlike agricultural income, agricultural land-holding, and time spent by women in collecting fuel. Considering those outcome variables which already had only around 2000 households (give or take a couple of hundred households), would lead to a very small sample of South Indian households (less than 500).

For the three variables thus chosen, the entire analysis is performed, as discussed in Section 2.4, for each of the sub-samples, i.e., we calculate the propensity scores for the short-term, long-term, and no connections for each household in the sub-sample, and regress the outcome variables in the propensity-score-weighted-least-squares-regression, against treatment and control variables.

		Δ Logarithm of per-capita Household Consumption (Rs.)	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
"BMRU" States	Intercept	0.2423***	0.2195***
		(0.0554)	(0.0582)
	Short-term Electricity Access	0.0954	0.1232*
		(0.0558)	(0.0658)
	Long-term Electricity Access	0.1519***	0.2109***
$N = 1822$		(0.0522)	(0.0580)
Southern States	Intercept	0.3289 **	0.0687
		(0.1234)	(0.1557)
	Short-term Electricity Access	0.1059	0.1263
		(0.1472)	(0.1929)
	Long-term Electricity Access	-0.0413	0.2690
$N = 544$		(0.1322)	(0.1702)

Table 2.14: Heterogeneous Effects: Differences in effects of electrification on per-capita consumption between "BMRU" states and Southern Indian States. "BMRU" states include Bihar, Jharkhand, Madhya Pradesh, Chattisgarh, Rajasthan, and Uttar Pradesh. Southern Indian States include Maharashtra, Karnataka, the erstwhile united Andhra Pradesh, Kerala, and Tamil Nadu. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

Tables 2.14-2.16 present the results for consumption, education, and kerosene consumption. From the results, we can observe that there are differences between the "BMRU" states and the South in terms of the impacts. In terms of the impact on per-capita consumption, the "BMRU" states show a large and statistically significant impact - 12.32% improvement on the short-term, significant at the 10% level, and 21.09% improvement on the long-term, significant at a 1% level. At the same time, the impacts in South India do not seem to be statistically significant, despite the estimated coefficients being rather similar to those estimated in the "BMRU" region - a 12.63% increase in the short-term and a 26.90% increase in the long-term. However, due to the small sample size, it is difficult to find statistically significant impacts in South India. Despite the large coefficients, the standard errors of the effects measured in South India are roughly three times as large as the effects in BMRU, although one would expect them to be about 1.8 times as large purely based

		Δ Schooling of the Highest Educated Adult (years)	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
"BMRU" States	Intercept	0.5095***	0.3927*
		(0.1895)	(0.2370)
	Short-term Electricity Access	-0.1816	0.0246
		(0.2623)	(0.3016)
	Long-term Electricity Access	0.3975*	0.3968
$N = 1817$		(0.2355)	(0.2832)
Southern States	Intercept	0.3298	0.1547
		(0.4126)	(0.4503)
	Short-term Electricity Access	1.3250**	0.5819
		(0.5486)	(0.6330)
	Long-term Electricity Access	1.0298**	1.0607**
$N = 543$		(0.4562)	(0.5255)

Table 2.15: Heterogeneous Effects: Differences in effects of electrification on the years of education of the highest-educated adult between "BMRU" states and Southern Indian States. "BMRU" states include Bihar, Jharkhand, Madhya Pradesh, Chattisgarh, Rajasthan, and Uttar Pradesh. Southern Indian States include Maharashtra, Karnataka, the erstwhile united Andhra Pradesh, Kerala, and Tamil Nadu. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

on sample size, the results are more mixed and not all households benefit from the electrification equally. The tripling of standard errors implies that the effect of electricity access on consumption may be "fuzzy".

The trends reverse for the impacts on the years of schooling of the highest educated adult, where the "BMRU" states show no impact in the short term, and in the long term, show an increase of 0.39 years, but this increase isn't statistically significant. On the contrary, the Southern states show an increase of 0.58 years in the short term (although not statistically significant) and an increase of over a full year of schooling in the long term, which is significant at the 5% level. The trends in the impact of electricity access on kerosene consumption mimic those on education. Households

		Δ Kerosene Consumption per month (litres)	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
	Intercept	-0.3190*	-0.3806*
		(0.1906)	(0.2004)
"BMRU" States	Short-term Electricity Access	0.0762	0.0906
		(0.1989)	(0.2328)
	Long-term Electricity Access	0.0429	0.1122
$N = 1822$		(0.1859)	(0.1864)
	Intercept	0.0035	0.4632
		(0.4013)	(0.5618)
Southern States	Short-term Electricity Access	0.0150	-0.7965
		(0.3958)	(0.6397)
	Long-term Electricity Access	-0.8931**	-1.2587***
$N = 543$		(0.3698)	(0.4497)

Table 2.16: Heterogeneous Effects: Differences in effects of electrification on the domestic consumption of kerosene between "BMRU" states and Southern Indian States. "BMRU" states include Bihar, Jharkhand, Madhya Pradesh, Chattisgarh, Rajasthan, and Uttar Pradesh. Southern Indian States include Maharashtra, Karnataka, the erstwhile united Andhra Pradesh, Kerala, and Tamil Nadu. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

in the "BMRU" states experience virtually no reduction in kerosene consumption. Meanwhile, the Southern States show a reduction of nearly 800 milliliters per month in the short term (although not statistically significant) and a reduction of close to 1.26 liters per month in the long term which is significant at the 1% level. Unlike the case with consumption, the standard errors roughly double for education, and increase two and a half fold, with a four-fold reduction in the size of the sample, which is more consistent with expectations.

The case of kerosene consumption highlights an important phenomenon - Even though we do not observe any significant impact at the national level (Table 2.12, there are statistically significant impacts felt by certain regions in the long term. It is also important to note that despite seeing a significant improvement in per-capita expenditure, households in the "BMRU" region do not observe corresponding improvements in education or a reduction in kerosene consumption. The smaller impact on impact on education may be due to poorer schooling infrastructure, or a lack of confidence in schooling when compared to South Indian States, who are forerunners of education in the country. On the contrary, the South Indian case demonstrates that despite there not being a significant impact on affluence (estimated through consumption), households can observe considerable improvements in schooling, with grid connections. The differences in the case of kerosene consumption are harder to explain. A resolution to this conundrum may depend on other factors such as the capacity of electricity, or voltage fluctuations, which are not captured in this survey, as differences in the strengths of connections may allow for a limited fuel displacement potential in some regions.

It is crucial to observe, however, that the broad relationship between short-term impact and long-term impact that was observed at the national level, is preserved for all three variables, in both regions. Among outcomes that show significant changes (consumption in "BMRU", education, and kerosene consumption in the South), it is always the long-term access group that observes the more significant impact, or sometimes is the only group that observes any significant impact. Even when the impacts are not significant, the coefficients of the long-term access group are observed to be larger, for instance in the case of consumption for Southern States, or education in the "BMRU" states. The only exception is with kerosene consumption in the "BMRU" states, where both the treatment groups observe a marginally larger kerosene consumption than those who were never connected, although the coefficients measured are virtually 0.

2.5.3 Heterogeneous Effects: Variations in Impacts of Electrification with Affluence

The question of potential gains from household electrification is likely to be answered in significantly different ways for affluent and deprived households. In Section 2.1, we postulated that due to affordability-related concerns, poor households could take longer to realize the benefits of electricity. Therefore, we could potentially observe richer households experiencing benefits much faster than poor households. However, there are other possibilities of how affluence could be interlinked with the prospect of benefiting from connections. Richer households, due to greater capital, are likely to have higher levels of income, land holding, education, etc., and may stand little to gain in these avenues after electrification.

		Δ Logarithm of per-capita Household Consumption (Rs.)	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
Poor Households	Intercept	0.3578***	0.3448***
		(0.0407)	(0.0459)
	Short-term Electricity Access	0.1522***	0.1750***
		(0.0481)	(0.0554)
	Long-term Electricity Access	0.2227***	0.2303***
$N = 2013$		(0.0407)	(0.0463)
Rich Households	Intercept	-0.0421	0.0250
		(0.0465)	(0.0741)
	Short-term Electricity Access	0.0745	-0.0036
		(0.0492)	(0.0843)
	Long-term Electricity Access	0.1587***	0.1229
$N = 2013$		(0.0478)	(0.0771)

Table 2.17: Heterogeneous Effects: Differences in effects of electrification on per-capita consumption between poor and rich households. Poor households are defined as those who have below-median total household consumption in 2005, and rich households are defined as those who have above-median total household consumption in 2005. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

To study if the short and long-term impacts of electrification affect affluent and poor households in a similar or different manner, we follow a similar approach to what we did in the previous subsection. We split the sample of households into "rich" and "poor" households based on the total household consumption as of 2005. Poor households are defined as those who consumed less than the median total household consumption in 2005, and correspondingly those households which consumed more than or equal to the median are dubbed rich households. The propensity scores are then estimated for each sub-sample for households that received long-term, short-term, and no access, and the weighted regressions are carried out. The results of the regressions are presented in Tables 2.17-2.22.

We observe diverse results in terms of impacts across the six variables, with different variables

		Δ Logarithm of household agricultural income (Rs.)	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
	Intercept	0.0370	-0.0230
		(0.1328)	(0.1460)
Poor Households	Short-term Electricity Access	0.3040*	0.2380
		(0.1820)	(0.2111)
	Long-term Electricity Access	0.0597	0.0912
$N = 1015$		(0.1406)	(0.1727)
	Intercept	-0.1734	0.2238
		(0.1278)	(0.2378)
Rich Households	Short-term Electricity Access	0.3313**	0.0098
		(0.1546)	(0.2390)
	Long-term Electricity Access	0.2900**	-0.0900
$N = 1016$		(0.1328)	(0.1921)

Table 2.18: Heterogeneous Effects: Differences in effects of electrification on household agricultural income between poor and rich households. Poor households are defined as those who have below-median total household consumption in 2005, and rich households are defined as those who have above-median total household consumption in 2005. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

responding differently for rich and poor households. For per-capita consumption (Table 2.17), we observe the trends seen in most regressions in the paper, with the long-term impact being larger and more significant than the short-term impact. Interestingly, the impact is felt more strikingly for poor households who observe a 17.50% increase in consumption with short-term access, and a 23.03% increase with long-term access, both significant at the 1% level. Rich households, on the other hand, show virtually no impact in the short term and show a 12.29% increase in consumption in the long term, which is not statistically significant. Since rich households were affluent even in the absence of grid connections, their existing economic activities may already have been giving them high returns, and switching to electricity-based businesses or production may not have instantly given returns much better than their existing businesses, and thus these households only

		Δ Logarithm of agricultural land held (Acres)	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
	Intercept	-0.2395	-0.2545
		(0.2060)	(0.1721)
Poor Households	Short-term Electricity Access	0.2512	0.4059*
		(0.2168)	(0.2132)
	Long-term Electricity Access	0.3739*	0.3663*
$N = 1059$		(0.2072)	(0.1887)
	Intercept	-0.2532	-0.2318
		(0.1844)	(0.1912)
Rich Households	Short-term Electricity Access	0.0603	-0.1382
		(0.1640)	(0.1931)
	Long-term Electricity Access	0.0410	-0.1426
$N = 1060$		(0.2075)	(0.2412)

Table 2.19: Heterogeneous Effects: Differences in effects of electrification on agricultural land holding between poor and rich households. Poor households are defined as those who have below-median total household consumption in 2005, and rich households are defined as those who have above-median total household consumption in 2005. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

show improvements in the long-term. Furthermore, some rich households may have also been generating their own electricity prior to being connected to the grid. Poor households, on the other hand, reaped larger benefits because their incomes had a larger potential to grow, and electricity access may have rescued them from their previously meager earnings.

Electricity access has no significant impact on agricultural income either for rich or poor households (Table 2.18). However, while the full sample showed that there is no impact on agricultural income in the short term, but a negative (although not significant) impact in the long term, disaggregating the sample into rich and poor households shows that there are short-term impacts as well for poor households, although not statistically significant. On average, poor households showed

		Δ Years of Schooling of Highest Educated Adult	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
Poor Households	Intercept	0.6572***	0.5972***
		(0.1721)	(0.1763)
	Short-term Electricity Access	0.4231*	0.3765
		(0.2559)	(0.2921)
	Long-term Electricity Access	0.4393**	0.3981*
$N = 2009$		(0.2034)	(0.2214)
Rich Households	Intercept	0.5174**	0.3639
		(0.2129)	(0.2855)
	Short-term Electricity Access	0.1715	0.1790
		(0.2902)	(0.3237)
	Long-term Electricity Access	0.4338*	0.5364*
$N = 2009$		(0.2460)	(0.3012)

Table 2.20: Heterogeneous Effects: Differences in effects of electrification on the schooling of the highest educated adult between poor and rich households. Poor households are defined as those who have below-median total household consumption in 2005, and rich households are defined as those who have above-median total household consumption in 2005. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

a 23.80% increase in agricultural income in the short term and only a 9.12% increase in the long term. Rich households responded quite differently and displayed a negligible change in the short term and a large 23.90% reduction in the long term, although this reduction was not statistically significant due to the large standard errors. An interesting commonality between rich and poor households is that both the long-term access groups see a smaller increase or a larger decrease in agricultural income. This may suggest that with the advent of electricity connections, households were moving away from agriculture and either diversifying to other modes of production or moving out of agriculture entirely. However, for rich households, this chain seems to occur mainly in the long term, which may explain why rich households only saw an improvement in consumption in the long term.

		Δ Kerosene Consumption per month (litres)	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
Poor Households	Intercept	0.0033	0.0127
		(0.1347)	(0.1528)
	Short-term Electricity Access	-0.1287	-0.0284
		(0.1744)	(0.1977)
	Long-term Electricity Access	0.1257	0.0993
$N = 2012$		(0.1349)	(0.1487)
Rich Households	Intercept	-0.5675***	-0.5989***
		(0.1802)	(0.1808)
	Short-term Electricity Access	0.0882	0.0729
		(0.2267)	(0.2522)
	Long-term Electricity Access	-0.0653	0.0530
$N = 2012$		(0.1906)	(0.1935)

Table 2.21: Heterogeneous Effects: Differences in effects of electrification on household kerosene consumption between poor and rich households. Poor households are defined as those who have below-median total household consumption in 2005, and rich households are defined as those who have above-median total household consumption in 2005. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

The results for agricultural land holding largely show similar trends as those for agricultural income, with poor households increasing their land holding while richer households reduced theirs. However, poor households showed a mammoth increase in land holding both in the short term and the long term, showing a 40.59% increase in land holding, in the short term, and a 36.63% increase in land holding in the long term, both significant at the 10% level. Rich households, although did not show statistically significant reductions, did display an average reduction of about 14% for both households with short-term and long-term access.

The impact on education (Table 2.20) is the most homogeneous, with positive impacts on both poor households (0.3981 years) and rich households (0.5364 years) in the long term, significant at

		Δ Time spent by women in collecting fuel (minutes per week)	
		Simple DID regression	p-Weighted DID regression
		(1)	(2)
Poor Households	Intercept	62.0868	73.3159
		(43.6159)	(47.5513)
	Short-term Electricity Access	-28.9664	-85.3941
		(75.3501)	(87.2903)
	Long-term Electricity Access	-70.5233	-89.0701*
$N = 902$		(47.8816)	(52.6618)
Rich Households	Intercept	146.9374***	129.3738**
		(48.0117)	(60.4043)
	Short-term Electricity Access	-97.6996*	-17.2312
		(56.9752)	(72.2787)
	Long-term Electricity Access	-178.9180***	-145.6118**
$N = 903$		(53.3286)	(67.3887)

Table 2.22: Heterogeneous Effects: Differences in effects of electrification on the time spent by women in collecting fuel between poor and rich households. Poor households are defined as those who have below-median total household consumption in 2005, and rich households are defined as those who have above-median total household consumption in 2005. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

the 10% level. Although not statistically significant, poor households also show a sizeable short-term impact of 0.3765 years. Kerosene consumption (Table 2.21) is more or less unaffected by electricity access both in the case of rich and poor households. The time spent by women in fuel collection is the only outcome variable that is significantly affected by electricity access for rich households but not as much for poor households (Table 2.22). While women in poor households are relieved of the burden of fuel collection by close to an hour and a half per week in poor households, both in the short and long term, the short-term effect is heterogeneous within the group leading to large standard errors and is thus not statistically significant. The long-term effect is significant at the 10% level. Rich households, on the other hand, observe a small statistically insignificant impact (17 minutes per week) in the short term, but a large statistically significant (at the 5% level) long-term impact of a reduction of close to two and a half hours per week spent in fuel collection by women. An explanation for poor households seeing time-homogenous effects, and rich households seeing a large increase with time may be due to the kind of fuel households used, to begin with. For instance, rich households may have been using collected, as well as bought fuel, and access to electricity may have helped them displace the bought fuel in the short run and collected fuel in the long run. On the contrary, poor households who were less likely to use bought fuel substituted collected fuel from the short run itself. It is also possible that rich households hire labor to collect fuel and have more free time, something that poor households are unlikely to be able to do. It is also possible, that with the advent of electricity, rich households decide to get other appliances to enhance their quality of life, such as possibly buying a gas stove, which would then reduce the demand for biomass collection.

2.6 Robustness Check: Very Long-term Connections

Our broad findings indicate that the impacts of electrification typically grow over time, and to demonstrate this, we used a difference-in-differences set up with two types of treatments - those who have had electricity access in the short term, and those who have had electricity access in the long term. Short-term, here, is a period of up to seven years, and long-term has been defined as greater than seven years and less than seventeen years. The rationale behind this categorization of connections based on the duration is rooted in our panels which are from 1995, 2005, and 2012. After removing households that were electrified before 1995, we are left with three types of households based on the nature of the connection - households that had not been connected by 2012, households that had been connected between 2005 and 2012, and households that were connected between 1995 and 2005 and retained that connection till 2012.

To check whether our results for the nature of the impacts of electrification changing over time are robust, we introduce the set of households that had already been electrified before 1995 and continue to be connected in 2012. These households can be assumed to have a “very long-term connection.” These households had been excluded from the main analysis because the period in which these households were actually connected is unknown which makes these households extremely diverse and heterogeneous in terms of how long they have had access to electricity. Nevertheless, this group makes for a good test of the robustness of our main results. If the results were to change completely upon the inclusion of this group, such as showing that impacts are higher in the short-term and lesser in the long term, then one would be less confident about our

	Δ Logarithm of per-capita Household Consumption (Rs.)	
	Simple DID regression	p-Weighted DID regression
$N = 7892$	(1)	(2)
Intercept	0.2217***	0.1907***
	(0.0340)	(0.0563)
Short-term Electricity Access	0.0998***	-0.0238
	(0.0355)	(0.0814)
Long-term Electricity Access	0.0874**	0.0597
	(0.0358)	(0.0584)
Very Long-term Electricity Access	0.0146	0.1467***
	(0.0354)	(0.0557)

Table 2.23: Robustness test: Time-associated causal effects of electrification on consumption, including households with very long connections. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, and whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

main results. Ideally, the inclusion of a fourth category of households, should not change the results much qualitatively and should preserve the broad trends observed.

It is important to note that over half the households common to the three surveys were already connected in 1994 (Table 2.4), and thus the inclusion of these households increases the size of our samples considerably. However, it is noteworthy that the variation in the lengths of time for which the households in the "very long-term access" group have been connected does make these results somewhat harder to interpret due to the lack of granularity, for instance, if the impacts of electrification prove to be non-monotonic after a certain period. Thus, these are presented as a check for robustness and not the main result. Another possible caveat of this approach is that since households that were connected between 1995 and 2005 may have already benefited from electricity access, some households connected just prior to 1995 may have characteristics similar to them, and characteristics dissimilar to those households which were electrified much earlier, making it difficult to estimate propensity scores, which may have an impact on the magnitude of the coefficients measured.

The empirical strategy in this section follows a modified version of equation 2.2, with the new

	Δ Logarithm of the Total Agricultural Income of the Household (Rs.)	
	Simple DID regression	p-Weighted DID regression
$N = 4269$	(1)	(2)
Intercept	-0.0787	0.1675
	(0.0865)	(0.1730)
Short-term Electricity Access	0.2451**	0.1048
	(0.1042)	(0.1824)
Long-term Electricity Access	0.1492	0.0727
	(0.0931)	(0.1601)
Very Long-term Electricity Access	0.2217**	-0.0071
	(0.0889)	(0.1411)

Table 2.24: Robustness test: Time-associated causal effects of electrification on agricultural income, including households with very long connections. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

	Δ Logarithm of the Total Agricultural Land Holding of the Household (Acres)	
	Simple DID regression	p-Weighted DID regression
$N = 4596$	(1)	(2)
Intercept	-0.1840	-0.1265
	(0.1554)	(0.1513)
Short-term Electricity Access	0.2589*	0.0648
	(0.1350)	(0.1373)
Long-term Electricity Access	0.1798	0.0895
	(0.1645)	(0.1706)
Very Long-term Electricity Access	0.1769	0.1362
	(0.1603)	(0.1650)

Table 2.25: Robustness test: Time-associated causal effects of electrification on agricultural land ownership, including households with very long connections. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

	Δ Years of Education of Highest Educated Adult in the Household	
	Simple DID regression	p-Weighted DID regression
$N = 7881$	(1)	(2)
Intercept	0.5512***	0.3593*
	(0.1249)	(0.2150)
Short-term Electricity Access	0.2825*	0.5507**
	(0.1647)	(0.2635)
Long-term Electricity Access	0.4336***	0.7163***
	(0.1562)	(0.2360)
Very Long-term Electricity Access	0.4096***	0.5406***
	(0.1323)	(0.2086)

Table 2.26: Robustness test: Time-associated causal effects of electrification on the schooling of the highest educated adult, including households with very long connections. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

treatment variable $\Delta T_{VL,ijt}$.

$$\Delta y_{ijt} = (\alpha_t - \alpha_{t-1}) + \beta_{VL}\Delta T_{VL,ijt} + \beta_L \Delta T_{L,ijt} + \beta_S \Delta T_{S,ijt} + \gamma.\Delta X_{ijt} + \Delta\epsilon_{ijt} \tag{2.3}$$

Here $\Delta T_{VL,ijt}$ is a dummy variable that takes 1 if the household received a connection before 1995. Accordingly, the propensity scores are estimated for the four categories, and subsequently, the weighted regressions are carried out. Tables 2.23-2.28 show the results for the six outcome variables.

There is considerable parity between the two sets of results in which outcome variables show statistically significant responses. Per-capita consumption, education, and the time spent by women in collecting fuel show statistically significant impacts, while agricultural income, agricultural land-holding, and kerosene consumption do not. There are some differences, however, between our main analysis and the robustness tests. The inclusion of very-long term connections reduces the estimated magnitudes of the coefficients of the short-term and long-term impacts on consumption, rendering them statistically insignificant and estimating the short-term impact as virtually 0. The coefficients measured for impacts on agricultural income are also smaller than those measured

	Δ Kerosene Consumption (litres per month)	
	Simple DID regression	p-Weighted DID regression
$N = 7890$	(1)	(2)
Intercept	-0.1384	-0.3198*
	(0.1227)	(0.1850)
Short-term Electricity Access	-0.0517	-0.0034
	(0.1518)	(0.2051)
Long-term Electricity Access	-0.1701	-0.0309
	(0.1335)	(0.2035)
Very Long-term Electricity Access	-0.5405***	-0.0295
	(0.1342)	(0.1808)

Table 2.27: Robustness test: Time-associated causal effects of electrification on household kerosene consumption, including households with very long connections. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

in Section 2.5, and the impact measured for time spent in fuel collection is also slightly smaller.

In contrast, the impact on the schooling of the highest-educated adult is magnified by the inclusion of households connected before 1995. This particular outcome is unique in multiple ways. Apart from being magnified, this is also the only outcome variable where the impact is statistically significant for all three types of treatments - short-term, long-term, and very long-term, all of them significant at least at the 5% level. This is also the only variable that does not show a monotonic increase or decrease, with the impact being the maximum for long-term connections, and the estimated effects for short-term access and very long-term access are similar in size. The marginally smaller effect on households with very long-term connections may be because these households have already reaped benefits over a long period, and there may be little left to improve, compared to households that hadn't been connected until much more recently.

Overall, we can observe that the trends observed in Section 2.5 are broadly reproduced in this section for most variables. Both the main results and the robustness check agree that the impact on per-capita consumption increases with time, the impact on agricultural income decreases with

$N = 2928$	Δ Time Spent by Women in Collecting Fuel (minutes per week) Simple DID regression (1)	p-Weighted DID regression (2)
Intercept	95.2886***	64.1895
	(33.4424)	(48.4581)
Short-term Electricity Access	-44.8394	18.5201
	(39.9268)	(51.6107)
Long-term Electricity Access	-115.0545***	-89.6976
	(35.8753)	(56.8679)
Very Long-term Electricity Access	-136.0877***	-94.9951*
	(35.1322)	(50.0845)

Table 2.28: Robustness test: Time-associated causal effects of electrification on the time spent by women in collecting fuel, including households with very long connections. Robust standard errors clustered at the village level. The control variables are reliability, households electrified in the village, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

time, becoming negative in the long run (here, very long run), the impact on education increases from short-term to long term, Kerosene consumption shows negligible changes and the reduction in the time spent by women in collecting fuel increases with time. The only variable which shows an opposite trend is agricultural land-holding, but the impacts measured were not statistically significant, to begin with, and are not significant, even after the inclusion of households with very long-term access. This parity between the two sets of results implies that our results are fairly robust, with some minor concessions.

Chapter 3

The Effect of Reliable Electricity on Casual Agricultural Labor Wages

3.1 Stylized Facts

For our analysis, we use data from two waves of the India Household Development Survey (Desai et al., 2005, 2011-2012). The first round of the survey was conducted in 2004-2005, and the second round was conducted in 2011-2012. The IHDS-I survey covered 41,554 households, and IHDS-II covered 42,152 households in total from 384 districts, 1,503 villages, and 971 urban blocks. Since we are primarily interested in village-level quantities like labor wage rates, we restrict our preliminary analysis to the village questionnaires and data. In all, one can construct a panel of 1406 villages from the two survey rounds. We further use household-level data to explain the phenomena we observe in the village-level analysis. Although there were differences in some of the questions asked in each survey, most of the variables of our interest were present in both rounds.

The IHDS surveys make for an excellent data set to study the impact of electrical reliability on agricultural labor wages. Primarily, the survey has detailed data on several insightful dimensions of village-level electricity. The survey tells us whether a village has electricity, what fraction of households in the village have electricity when the village was first connected to the grid, and the average number of hours in a day that a village receives power. Similarly, the surveys also have data on the casual agricultural wage rates for sowing, and harvesting, for both men and women and for both major cropping seasons in India - Kharif (summer/monsoon) and Rabi(winter). Apart

	Means (standard deviation)				Correlation of
	2004-2005		2011-2012		∆Reliability with
	Positive Treatment	Negative Treatment	Positive Treatment	Negative Treatment	Initial levels (2005)
Percentage of households with Electricity access (%)	69.98	71.50	80.36	79.90	0.00
	(32.83)	(30.19)	(25.99)	(24.86)	
Presence of Metalled Roads†	0.62	0.73	0.86	0.90	-0.09**
	(0.49)	(0.45)	(0.35)	(0.31)	
Distance to the nearest bank branch/credit cooperative (km)	5.27	4.03	5.44	4.54	0.10***
	(6.05)	(4.61)	(5.99)	(4.95)	
Distance to the closest market (km)	2.37	2.38	2.43	2.78	0.00
	(4.69)	(4.33)	(5.26)	(5.50)	
Presence of NGO/Development Organization†	0.11	0.15	0.14	0.14	-0.07*
	(0.31)	(0.36)	(0.35)	(0.35)	
Presence of Primary Healthcare Center†	0.11	0.17	0.12	0.15	-0.09**
	(0.32)	(0.44)	(0.33)	(0.36)	
Number of Government Primary Schools†	1.54	1.95	1.67	1.52	-0.17***
	(1.34)	(1.89)	(1.54)	(1.78)	
Number of Private Primary Schools†	0.60	0.83	0.71	0.85	-0.07**
	(1.16)	(1.85)	(1.28)	(1.48)	
Number of Government Middle Schools†	0.70	0.74	0.88	0.89	-0.09**
	(0.68)	(0.74)	(0.73)	(0.72)	
Number of Private Middle Schools†	0.31	0.40	0.47	0.51	-0.04
	(0.79)	(1.12)	(1.10)	(1.10)	
Number of Government Secondary Schools†	0.28	0.34	0.38	0.42	-0.05
	(0.51)	(0.53)	(0.66)	(0.61)	
Number of Private Secondary Schools†	0.13	0.29	0.23	0.32	-0.09***
	(0.46)	(1.00)	(0.66)	(0.75)	
Number of Government Higher Secondary Schools†	0.11	0.14	0.18	0.17	-0.03
	(0.35)	(0.37)	(0.44)	(0.43)	
Number of Private Higher Secondary Schools†	0.05	0.08	0.14	0.19	-0.07*
	(0.27)	(0.35)	(0.53)	(0.58)	

Table 3.1: Descriptive statistics for treatment and control Villages 2004-2005 and 2011-2012. "Positive Treatment" villages refer to those villages which see an improvement in the average number of hours of electricity received in a day, and "Negative Treatment" villages refer to those which do not (These include control villages where there is no change). Correlation refers to Pearson's correlation coefficient, with the two-sided alternate hypothesis. Includes 1254 villages. The data are unweighted. Source: IHDS I, and IHDS II surveys.
† Dummy variable which takes 1 for "yes" and 0 for "no".
*Significant at the 5% level, **Significant at the 1% level, ***Significant at the 0.1% level

from data on agricultural wage rates, there is also data on the wage rates of casual non-agricultural laborers, domestic laborers, and construction laborers, which allows us to compare the trends between different groups. At the level of the household, the survey has data on consumption, which is our primary outcome variable in studying the spillover effects on households, the status of the households connection, which helps us construct a sample of households that have not been connected to the grid, and various other control variables that may influence consumption expenditure in a household. Since the surveys also have household-level data, we can further investigate the causal factors of the trends in agricultural wages that we observe, at a more microscopic level, by studying changes in agricultural households.

From Table 2.3 and 2.7, we can observe that there are significant differences in the characteristics of households that have electricity and households that do not, and therefore, the potential to benefit from interactions with other connected households, may be different for households with

	Means (standard deviation)					
	Reliability of Electricity Access (hours per day)			Households with Access (%)		
	2004-2005	2011-2012	Δ Reliability	2004-2005	2011-2012	Δ Households with Access
	(1)	(2)	(2)-(1)	(3)	(4)	(4) - (3)
Himachal Pradesh	13.50	22.54	9.04	98.04	99.58	1.54
	(8.92)	(1.97)	(9.25)	(7.53)	(1.64)	(7.30)
Punjab	11.34	19.51	8.17	95.59	95.95	0.36
	(5.90)	(4.04)	(7.77)	(10.45)	(8.00)	(13.24)
Uttarakhand	14.71	16.41	1.71	84.41	91.71	7.29
	(5.27)	(3.24)	(7.00)	(13.81)	(10.31)	(15.37)
Haryana	8.64	7.60	-1.04	90.19	93.71	3.52
	(4.45)	(4.31)	(5.11)	(16.65)	(14.00)	(17.55)
Rajasthan	7.42	10.06	2.64	58.96	81.07	22.11
	(4.54)	(4.79)	(5.68)	(33.54)	(20.55)	(35.28)
Uttar Pradesh	7.52	7.40	-0.12	44.69	44.82	0.13
	(3.67)	(2.80)	(4.29)	(32.45)	(28.42)	(36.33)
Bihar	3.74	6.77	3.02	30.05	38.12	8.07
	(4.45)	(4.25)	(5.57)	(27.29)	(25.11)	(34.65)
Arunachal Pradesh	19.00	24.00	5.00	95.83	100.00	4.17
	(4.40)	(0.00)	(4.40)	(6.07)	(0.00)	(6.07)
Nagaland	10.20	8.20	-2.00	58.20	97.60	39.40
	(4.40)	(2.48)	(5.48)	(30.81)	(3.88)	(32.63)
Manipur	5.00	7.00	2.00	70.00	100.00	30.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Tripura	17.20	21.00	3.80	48.00	86.80	38.80
	(8.63)	(1.26)	(9.17)	(34.29)	(15.20)	(34.64)
Meghalaya	20.00	18.67	-1.33	80.83	95.33	14.50
	(5.92)	(5.85)	(9.48)	(15.39)	(6.90)	(18.93)
Assam	6.37	8.68	2.32	35.00	65.84	30.84
	(4.77)	(5.14)	(6.42)	(30.82)	(24.04)	(30.85)
West Bengal	17.25	20.27	3.02	40.64	69.41	28.76
	(6.17)	(2.87)	(7.21)	(32.76)	(22.74)	(33.98)
Jharkhand	10.67	10.71	0.05	55.48	78.71	23.24
	(7.21)	(5.72)	(7.84)	(33.38)	(23.41)	(31.55)
Odisha	17.52	12.94	-4.58	34.38	63.47	29.09
	(6.64)	(5.95)	(8.35)	(27.22)	(25.39)	(34.62)
Chattisgarh	15.82	17.43	1.61	61.45	84.59	23.14
	(6.73)	(5.05)	(7.32)	(30.65)	(20.69)	(26.46)
Madhya Pradesh	6.47	8.51	2.04	80.20	76.75	-3.45
	(3.55)	(2.91)	(4.10)	(23.38)	(24.71)	(30.05)
Gujarat	18.15	23.66	5.52	89.85	90.77	0.92
	(5.90)	(1.95)	(5.86)	(20.52)	(17.62)	(28.13)
Maharashtra	16.45	15.93	-0.53	79.24	89.31	10.07
	(3.87)	(4.91)	(5.98)	(18.07)	(11.18)	(20.13)
Andhra Pradesh	15.78	13.69	-2.09	84.09	90.99	6.90
	(5.02)	(4.11)	(6.41)	(17.71)	(9.96)	(20.78)
Karnataka	11.23	11.15	-0.08	81.47	89.12	7.65
	(5.89)	(5.10)	(7.13)	(15.89)	(13.22)	(19.07)
Kerala	22.22	23.44	1.22	83.67	95.89	12.22
	(2.35)	(0.68)	(2.48)	(8.87)	(3.93)	(10.83)
Tamil Nadu	22.11	11.89	-10.22	88.20	92.17	3.98
	(4.21)	(4.41)	(5.69)	(9.78)	(12.50)	(15.04)

Table 3.2: Electrical reliability and the fraction of electrified households by state, in 2004-2005 and 2011-2012. Includes 1220 villages. The data are unweighted. Source: IHDS I, and IHDS II surveys.

Villages from the present day Union Territories of Jammu & Kashmir, Ladakh, Daman & Diu and Dadra & Nagar Haveli have been omitted in this table.

electricity access and households without. In Van de Walle et al. (2017), the authors find that the impact of spillovers on consumption is considerably different for households with electricity access and households without. This motivates us to study the spillover effects of electricity on households not connected. The average spillover effect is studied as a control variable in Chapter 2.

At the crux of our analysis of wages, is studying whether a causal link exists between the reliability of electricity and agricultural labor wage rates. Since we work with observed data, rather than an experiment, whether a village experienced improvements or declines in reliability was not determined by randomization, and while there is little chance of reverse causality between wage rates and reliability, it is possible that better-developed villages receive better quality power, and be more prosperous at the same time, which could bias the composition of the treatment group, and one would arrive at a biased estimate of the effect on wages. Therefore, it is important to check if there are differences in the general characteristics of villages that saw an improvement in the quality of electricity and those that did not.

Table 3.1 presents the mean and standard deviations village characteristics of villages that receive more reliable electricity access ("Positive Treatment" group) and those that do not ("Negative Treatment" group) along with Pearson's correlation coefficient of the change in reliability with the 2005 levels of the variables. Overall, there are only marginal differences between most characteristics of treatment and control group villages, but it seems as though control villages are slightly more developed than treatment villages, albeit with large standard deviations. This may be because better-developed villages may already have reliable electricity on average and there may be less scope for improvement. But given, that the correlation coefficients are either small or completely insignificant. The only variable that shows a correlation coefficient greater than 0.1 is the number of Government Primary Schools. This may suggest that while villages that received improvements in reliability were marginally less developed than those that did not, the differences were too small sufficiently affect assignment to treatment.

Interestingly, most characteristics show a larger improvement or smaller decline in the period under consideration for the treatment group, when compared to the control group. This adds weight to our earlier postulate that the means are slightly lower for the treatment group because these villages were less developed. Therefore, along with improvements in electrical reliability, they also saw improvements in other development parameters. This may make it difficult to attribute the differences to electrical reliability alone, and it is important to control for several other variables,

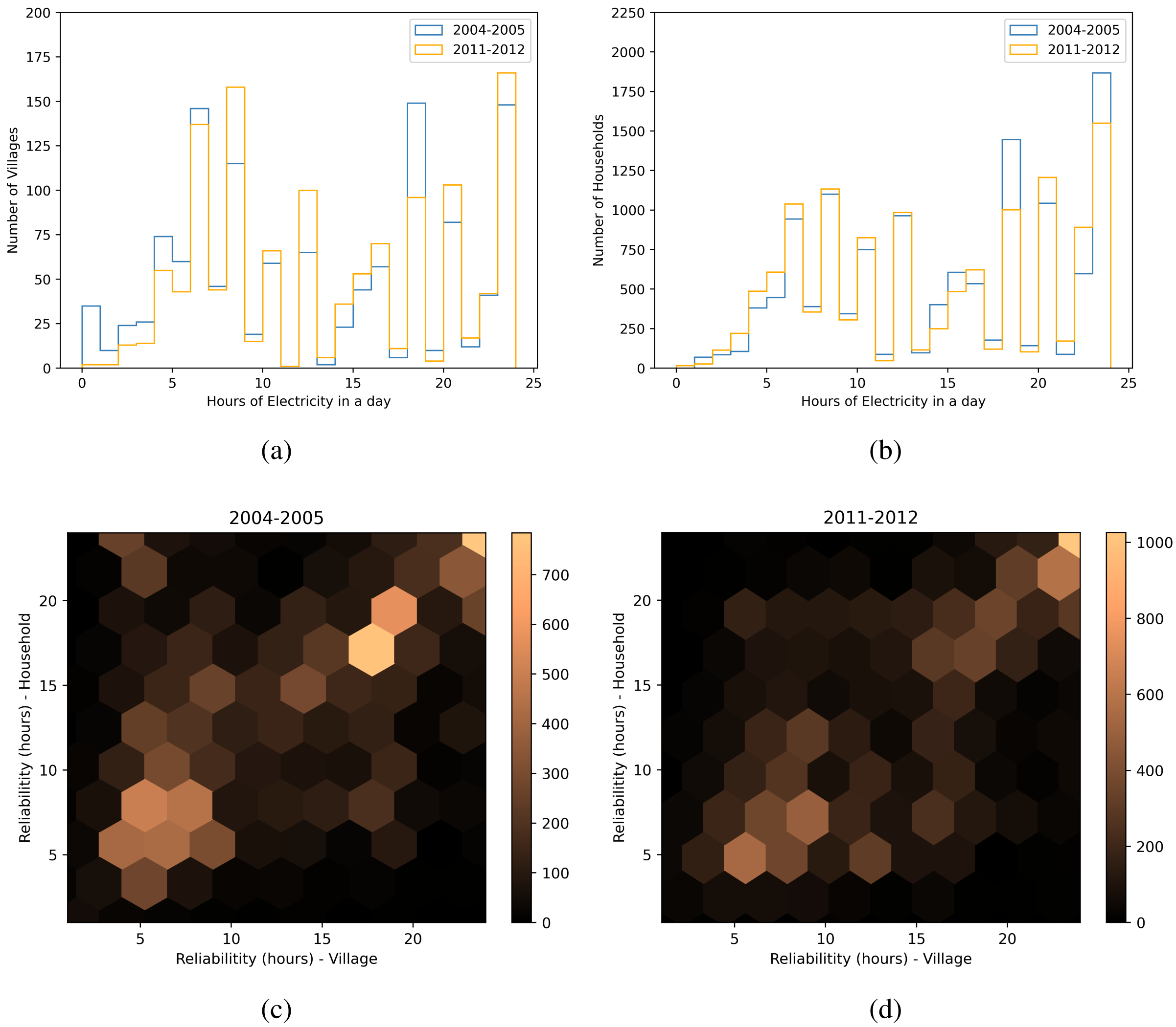


Figure 3.1: Distributions of electrical reliability for (3.1a) Villages, and (3.1b) households, for rural households electrified by 2004-2005, and density plots showing the household reliability against village reliability (3.1c) 2004-2005 (Pearson's corr. coeff. = 0.5491, $p = 0.000$), and (3.1d) 2011-2012 (Pearson's corr. coeff. = 0.6628, $p = 0.000$). The data are unweighted. Source: IHDS I, and IHDS II surveys.

improvements in which, could also causally impact labor wage rates. Apart from observable differences, there would also be time-invariant unobservable characteristics that are unique to villages. Since such variables cannot be controlled for, we use a difference-in-difference design to eliminate the time-invariant village-level unobservable characteristics, that could affect labor wage rates.

Table 3.2 shows the state-wise levels and changes for the hours of electricity available and the fraction of households with electricity access. There is significant variation, even at the level of the states, and although most states see an improvement in the reliability of electricity, some states see a decline, with some declines amounting to large figures with Odisha seeing a reduction of nearly five hours per day, and Tamil Nadu facing a ten-hour decline. The effect of the 2011-2012 power generation shock is also noticeable as 9 out of 24 states actually see a reduction in the quality of electricity, and the induced variation is visible as the standard deviation in the change in reliability is large, often considerably larger than the standard deviation for the levels. This signifies that there are large variations even within the state, implying that there may be random variations in the reliability of electricity that a village receives, based on unobservable factors. We also observe that most states see an increase in the average fraction of households in villages that have access. But once again, there is considerable variation in the changes, implying that it is also possible for villages to have fewer households with electricity access in the later period. There is also little correspondence between the changes in reliability and the changes in households with access. Since we hypothesized that the effect on labor wages could spread indirectly, the fraction of households electrified would be an important variable to control for, in order to isolate the impact of reliability.

Figure 3.1 shows us the reliability of electricity access for villages and households. We can notice from figure 3.1a shows a histogram of the distribution of reliability for the villages in the sample, and figure 3.1b shows the corresponding distribution for households, and we can observe that there is considerable parity between the plots for village-level measurements of reliability, and household-level ones. The correlation between household-level reliability and village-level reliability is 0.5491 and 0.6628, in 2004-05 and 2011-12, respectively, both coefficients being extremely statistically significant. Nevertheless, figure 3.1c and figure 3.1d show that despite the strong positive correlation, there are large variations as well, with points scattered almost all over the parameter space. The parity between village and household-level reliability data convinces us that the responses to the survey are reliable. The variation thereafter may be due to other factors such as the distribution of power within the villages. Another advantage of having data on reliability at the level of the household is that it allows us to explore the mechanisms through which village-level reliability could affect labor wages through individual households.

Revisiting Table 1.1, one can observe that there is little difference between wage rates in the summer/monsoon, and those in winter, i.e. wages are not seasonal. This may indicate that the changes are not specific to the type of crop, and is rather a broader change in the dynamics of agricultural labor and supply which affects labor wage rates throughout the year. This also simplifies the analysis as it allows us to construct a variable that represents labor wage rates for a village by averaging the wage rates of the kharif and rabi seasons. Moreover, since there is negligible variation in the means and variances of kharif and rabi wage rates, in the event that a village has data for only one of the two wage rates, it can be used as the sole representative of the wage rate in the village. On the contrary, there is a significant gender wage gap in the wage rates, with the rates for men being much larger at the onset (about 30-40%) of the natural experiment, and despite a reduction, still sizeable in 2012 (about 25-35%). Apart from a wage difference, there are also differences in how men's and women's wage rates differ between treated and controlled villages. While the men's wage rates differ between the two groups by approximately 10%, the corresponding disparity in women's rates is only about 5%. These differences in men's and women's wage rates are noteworthy as it necessitates a dedicated heterogeneity analysis for men's and women's wage rates as they may not exhibit identical trends.

To add to our previous observations from Table 3.1, there is no discernable difference in the 2005 levels of wage rates between those villages where reliability improves (treatment) and those where it does not (control). This is an important prerequisite in a difference-in-differences analysis as the initial levels for the variables of interest are extremely similar both for the treatment and control group villages and therefore any differences observed in the 2012 levels can be attributed to changes during the period, rather than preexisting differences in initial levels. There are, however, observable differences in the 2012 levels, which is of particular interest to us. The data reveals that even though there were no distinguishable differences in the 2005 wage rates between villages that received betterment in electricity quality and villages that did not, the wage rates of villages that saw an improvement in reliability were lower on average in 2012, due to an observably smaller (about 5-10%) increase in wage rates in treated villages - a trend consistent across gender and season. This lends more motivation towards performing a formal analysis of the causal effect of reliability on agricultural labor wage rates and exploring the mechanisms through which electricity quality affects unskilled labor wages.

3.2 Methodology

Our preliminary analysis is focused on whether there are spillover effects of electricity access on consumption. Suppose the outcome for treated households is $Y(1)$ and the expected outcome on untreated households is $Y(0)$, then one may expect the average treatment effect (ATE) to be

$$ATE = E[Y(1) - Y(0)] \tag{3.1}$$

However, $Y(0)$ would include benefits that untreated households receive through spillovers, if any. Suppose the spillover benefits on untreated households is denoted as $Y_S(0)$, then the unbiased ATE is given by

$$ATE = E[Y(1) - Y(0) + Y_S(0)] \tag{3.2}$$

To evaluate whether such spillover benefits exist, we design an empirical strategy to study the impacts of village electrification on untreated households. For this purpose, we consider the set of untreated households in the data and study how consumption in the household, a proxy for affluence, varies with the fraction of households electrified in the village. Since treated households may receive smaller spillover benefits than untreated households, from other households being electrified, we consider only untreated households in our analysis. We use household and village-level data from IHDS and use a two-stage propensity score-weighted-Difference-in-Differences design to estimate the spillover effects. Our sample includes some pure controls, i.e., unelectrified households in villages that have not been connected to the grid, and various degrees of treatment, which correspond to different fractions of households in the village that are connected to the grid.

We use the IHDS survey rounds of 2004-2005 and 2011-2012 and perform our DID analysis using the changes in the levels of the outcome, treatment, and control variables over the two periods. The outcome variable we use is the per-capita consumption expenditure of households, and to ensure that our results are robust and not anomalous, we study the impact on both changes in the levels and logs of consumption. We assume that spillover benefits primarily operate through other households. Therefore, we use the fraction of households with access to electricity as our main treatment variable. However, households that are nominally connected may not receive large benefits themselves and are unlikely to lead to spillover effects on other households, making controlling for the quality of electricity important. Therefore, we also use the average hours that a village has power each day, as an independent variable. We also include the pre-treatment levels of the fraction of households electrified and the reliability to control for the overall levels of the variables rather than just changes in the levels.

To be able to attribute the effects to electricity alone, we use several household and village-level control variables. To address compositional effects and the potential misattribution of effects of other development characteristics to electricity, we control for the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers. Therefore, we can model the first differences in per-capita consumption (Δy_{ijt}) of unelectrified households using the following equation, where E_{jt} is the fraction of households connected in the village j at time t, Q_{jt} is the reliability of electricity in the village j at time t, and T_{jt} is the number of years that village j has been connected for, measured at time t, γ is a vector of coefficients for control variables, and ϵ_{ijt} is the idiosyncratic error term. Note that we also include the lagged fraction of households connected, and a lagged reliability term, to study the overall effects of the variable and not just the increase in levels.

$$\Delta y_{ijt} = \alpha + \beta_E \Delta E_{jt} + \beta_{EL} E_{jt-1} + \beta_Q \Delta Q_{jt} + \beta_{QL} Q_{jt-1} + \beta_T T_{jt} + \gamma . \Delta X_{ijt} + \Delta \epsilon_{ijt} \quad (3.3)$$

However, controlling for other variables need not be sufficient to arrive at an unbiased estimate of the causal effects. A simple DID design assumes that the assignment of households to treatment is random. However, this assumption need not be entirely valid. There could be several factors that could lead to a rapid increase in the fraction of households with access to electricity resulting in a selection bias. In order to address the selection bias, we weight households with the inverse of their generalized propensity scores to remove selection bias due to observables. We use a multinomial regression for the first stage estimation of the propensity scores. The observables that we use are the pre-treatment levels of the log of the total household consumption, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers, the fraction of brahmins, non-brahmin forward castes, other backward castes, scheduled castes, and scheduled tribes in the village.

Subsequently, we study the reliability of electricity, measured in the average number of hours of electricity received in a day has an effect on casual agricultural labor wage rates. The IHDS surveys of 2004-2005, and 2011-2012 have data on casual agricultural wage rates at the village level, for men and women, summer and winter crops, and different agricultural labor jobs such as plowing, sowing, harvesting, etc. In order to construct an outcome variable for the wage rates, we consider only the sowing and harvesting wages, because these two groups are consistent across survey rounds. Plowing wage rates were also present in the first round of the survey but we exclude plowing wage rates from our outcome variable, as plowing was relegated to the category of "Other agricultural work", a category which may include wages for other work such as winnowing or threshing. Therefore, to maintain parity in the variables considered between the two rounds of the survey, we take an average of the sowing wage rates and the harvesting wage rates. In the presence of only one of the two rates, it is assumed representative of wage rates in the village. These wage rates are then categorized according to season and then averaged with equal weight given to each season. In the event that there is data only for one season, the available data is used to extrapolate for other seasons. The data is then averaged across seasons to find the wage rates for men and women separately, which are then averaged (in the same method described above) to find our final variable of interest - casual agricultural labor wage rates, measured in 2012 Rs. per day. We standardize all monetary values in 2012 Rs., and we use the national consumer price index for India as reported by the World Bank to bring the 2004-2005 data to the standard units. We use both the levels and logs of the wage rates as outcome variables.

As for the empirical strategy, an ideal design would have involved a random assignment of different quality power supplies to different villages. But there are several factors that determine the quality of power that a village receives, and manipulation of the quality received at the village level is both unfeasible for such a large geographical region and population and perhaps unethical as well. Instead, since Table 3.1 and Table 1.1 show very little association between different village-level characteristics and wage rates, the changes in reliability can be attributed almost entirely to unobservables (such as geography, proximity to power stations, political determinants, etc.), and the effect of unobservables that determine assignment to treatment can be assumed to be instances of natural randomization.

Consider the equation

$$y_{ijt} = \alpha_0 + \alpha_t + \beta Q_{ijt} + \gamma.X_{ijt} + \mu_i + \nu_j + \epsilon_{ijt} \tag{3.4}$$

where α_t is the time-fixed effect felt by all households at time t, β is the coefficient that measures the effect of reliability Q_{ijt}, γ is the vector of coefficients for village characteristics, which make up the control variables, μ_i is the time-invariant village-specific effect of village i, ν_j is the time-invariant district-specific effects of village j, and ϵ_{ijt} is the idiosyncratic error term. We assume that there are no village- and district-level time-variant effects. However, since we are studying the effect of reliability at the level of the village, and not merely for households, an important consideration also involves asking how many households have access to electricity, regardless of how reliable it is. This is because the wage rates of unskilled labor are unlikely to be affected by the electrification status of the laborer's household, and more likely to be affected by the connection at their employer's farm/firm. Therefore, a greater fraction of households with electricity will allow the effect to permeate better through the systems and networks of the village. Therefore, apart from reliability, which is our main variable of interest, we also control for the fraction of electrified households, say E_{ijt}. Thus, we can rewrite equation 3.4 as

$$y_{ijt} = \alpha_0 + \alpha_t + \beta_1 Q_{ijt} + \beta_2 E_{ijt} + \gamma.X_{ijt} + \mu_i + \nu_j + \eta_{it} + \theta_{jt} + \epsilon_{ijt} \tag{3.5}$$

However, by observing equation 3.5 one can note that while, Q_{ijt}, E_{ijt}, and X_{ijt} are observable and measured, μ_i and ν_j are not. A part of this problem can be fixed by rewriting the equation in terms of its first difference with time, i.e.,

$$\Delta y_{ijt} = (\alpha_t - \alpha_{t-1}) + \beta_1 \Delta Q_{ijt} + \beta_2 \Delta E_{ijt} + \gamma.\Delta X_{ijt} + \Delta\epsilon_{ijt} \tag{3.6}$$

where the notation ΔZ_{ijt} represents $Z_{ijt} - Z_{ijt-1}$ for any variable Z. As one can notice, the time-invariant unobservable effects terms have been done away with, as a result of the differencing.

The ideal check for a parallel trends assumption would be if we could have visualized what the data would look like had there been no intervention, i.e., if the treatment group and the control group showed similar changes in the levels of agricultural labor wage rates, we would know that the assignment to treatment was unbiased. This is often ruled out by researchers by observing the trends in the outcome variable when there are multiple pre-treatment periods. However, since we are constrained by two survey rounds, we cannot compare differences in the treatment and control groups in the pre-treatment period. Instead, we rely on one observation from the data and reasoning based on our understanding of rural India. Firstly, on observing the data in Table 1.1 we can see that the wage rates in treatment and control group villages are indistinguishable from each other. This, of course, does not necessarily imply that the growth of wage rates in fixed periods was equal, but it does show that on average, the 2005 wage rates were similar in villages that were to be treated and those that weren't.

The biggest argument for the parallel trends assumption, however, is that the reliability of electricity supply and agricultural labor wages are rather disconnected, and are affected by distinct sets of factors most of which do not affect the other. Reliability is mostly dependent on factors outside the village or district, such as power generation and transmission, while agricultural wage rates are mostly tied to agricultural practices and production which are mostly specific to the village. Therefore, the possibility that both changes in reliability and wages are the result of a certain omitted unobservable variable, is extremely low.

Therefore, using equation 3.6 we can estimate the impact of changes in reliability on wages. But the models stated in equation 3.5 and 3.6 make an assumption, that a fixed change in the reliability alone is responsible for changes in the wages. This implicitly assumes that a fixed increase or decrease in villages receiving poor-quality electricity in the initial round would have the same impact on wages as the same increase in a village that already received good-quality electricity, i.e., an increase in quality from 18 hours to 24 hours a day, has the same impact as an increase from 4 hours to 10 hours a day, which may not be an accurate model. Similarly, a village that sees a 20% improvement from 75% to 90%, need not be affected in the same way as a village that sees the same improvement from 5% to 25%. Therefore we include the pre-treatment levels of the fraction of electrified households and the reliability to account for lagged effects. Therefore,

$$\Delta y_{ijt} = \Delta\alpha_t + \beta_1\Delta Q_{ijt} + \beta_2\Delta E_{ijt} + \beta_3 Q_{ijt-1} + \beta_4 E_{ijt-1} + \gamma.\Delta X_{ijt} + \Delta\epsilon_{ijt} \tag{3.7}$$

This is the final form of the equation that we intend to fit and with knowledge of ΔQ_{ijt}, ΔE_{ijt}, Q_{ijt-1}, and E_{ijt-1}, we can find all the coefficients.

3.3 Results

3.3.1 The Spillover Effects of Electrification on Consumption

Table 3.3 presents the results on the impact of rural electrification on untreated households. We find that there is a statistically significant positive impact of the change in the fraction of households connected on per-capita consumption in households without connections. Employing the propensity score-weighted-DID regression, we find that for every 1% increase in the fraction of connected households, there is a 0.18% increase in per-capita consumption expenditure of the household, significant at the 5% level. We also find a statistically significant positive impact of the pre-treatment level of household electrification in the village. The pre-treatment levels are relevant because one needs to study the effect of the overall fraction of connected households as well, apart from the simple increase in households connected. Our results reveal that every 1% more households connected to the grid at the start of the study period results in a 0.21% increase in per-capita consumption in untreated households, significant at the 5% level. In contrast, reliability has a negative spillover effect. Although we find that the average effect of a change in reliability at the level of the village has a neutral effect on per-capita consumption, the pre-treatment level of reliability has a negative impact. For every additional hour of electricity that the village received at the start of the study period, there was a -0.8% decrease in per-capita consumption, although the impact on logs is significant only at the 10% level, while the impact on levels is significant at the 5% level.

The coefficients for the effects on the levels of per-capita household income largely agree with the effects on the logs. Among the minor differences is an increased statistical significance (at the 1% level) of the impact of the change in the fraction of households connected in the village, with a 1% increase in the fraction of households connected resulting in a Rs.77 increase in the per-capita consumption of the household. The impact of pre-treatment reliability is also considerably more statistically significant (5% level as opposed to 10% level for logs), with every additional hour of pre-treatment power availability resulting in a Rs.318 reduction in per-capita consumption. The negative impact of reliability on both the logs and levels of per-capita consumption may be due to the negative impact of reliability on agricultural labor wages as explored in Chapter 3.

Our results show that there are significant spillover effects of electrification on households without access, even in a predominantly agrarian rural economy like India's. The presence of spillovers implies that future studies need to incorporate strategies to eliminate biases in estimates

$N = 2416$	Δ per-capita household expenditure (2012 Rs.)			
	Logs		Levels	
	Simple DID	p-Weighted DID	Simple DID	p-Weighted DID
	(1)	(2)	(3)	(4)
Intercept	0.2581***	0.3893***	4648.1488***	9138.9564***
	(0.0638)	(0.0818)	(1632.9369)	(2225.2012)
Δ Fraction of Electrified Households in the Village	0.0012*	0.0018**	26.1406	77.1862***
	(0.0007)	(0.0009)	(21.1712)	(26.9947)
Pre-treatment Fraction of Electrified Households in the Village	0.0014*	0.0021**	28.9375	55.3793**
	(0.0008)	(0.0010)	(18.0469)	(23.9016)
Δ Reliability (hours)	-0.0026	-0.0056	50.8144	-125.1991
	(0.0035)	(0.0047)	(77.2223)	(126.7775)
Pre-treatment Reliability (hours)	-0.0067**	-0.0080*	-148.9749*	-318.2548**
	(0.0033)	(0.0046)	(79.5147)	(127.6134)
Years since Connection	-0.0010	-0.0030	-1.3663	-23.5111
	(0.0017)	(0.0021)	(39.0247)	(56.7065)

Table 3.3: Spillover benefits of rural electrification on households without access. Robust standard errors clustered at the village level. The control variables are the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

induced by spillovers. The presence of spillover benefits also implies that randomized control trials are posed with an additional challenge - the potential invalidity of SUTVA. In experimental impact evaluation studies, the sample is usually restricted to a village or a small group of villages. If connected households are measured against unconnected households in a village connected to the grid, the presence of a positive spillover may lead to an underestimate of the impact of access to electricity at the household level. The only alternative is to find pure controls - households in other villages, so the village is yet to be connected to the grid. But then village-level fixed effects would become difficult to account for.

Our results are in line with those of Khandker et al. (2009) in Bangladesh. On the contrary, despite finding positive spillover effects, our results do not agree with those of Van de Walle et al. (2017). While they find that there is a positive spillover effect of the number of years that the village has been connected, we find no such impact. This may be because Van de Walle et al. consider only one channel for spillover effects - the number of years since the village was first connected, while we consider multiple factors such as the fraction of households with access in the village, the quality of electricity available to the village and the years since the village was first connected. The number of years that a village has had access may be correlated with the fraction of households electrified, and the impact of the fraction of households connected may confound the impact of the number of years.

Interestingly, the pre-treatment reliability which represents the lagged reliability has a negative impact on per-capita household consumption, the opposite effect of the fraction of households electrified. This means that although more households with access lead to unelectrified households becoming more affluent, a better quality of electricity reduces consumption. While the positive impact of the fraction of households electrified can be explained through the economy-wide effects of benefits experienced by more households, the negative impact of reliability is more difficult to explain. This prompts us to study the impact of the quality of electricity or reliability on the wages of agricultural laborers, which make up a large fraction of the poor in rural India.

3.3.2 The Effect of Reliability on Agricultural Labor Wages

Table 3.4 presents the results of the difference in differences regressions for the impact of a change in the reliability on the change in casual agricultural labor wage rates. We find that a change in reliability has a negative impact on both the levels of casual agricultural labor wages and the logs. The effect on both the logs and levels is statistically significant at the 1% level, and the incorporation of the levels of reliability and the fraction of households electrified in the initial round leads to a strengthening of the effect on wages.

The results show that for every additional hour of power available to the village, the labor wage rate reduces by Rs. 1.78. The level of electrical reliability in the pre-treatment period also has a negative impact on wage rates, statistically significant at the 1% level, which shows that the overall level of reliability also has a negative impact on the wage rate. On the contrary, a change in the fraction of households connected has a positive impact on wage rates, including the fraction of households connected in the pre-treatment period. This implies that while more houses being connected to the grid has a positive impact on wage rates, improvement in the quality of electricity received by the village has a negative impact.

To study this further, we study the impact on men's and women's agricultural labor wage rates separately. Table 3.5 presents the results of the regressions on men's and women's agricultural labor wage rates. While the results for both men's and women's wages agree broadly with the results on wages, the effects of reliability on men's and women's wage rates are not homogeneous. The impact of reliability, although negative for both men and women, is more strongly negative on

	Coefficients (standard errors)			
	Levels		Logs	
	(1)	(2)	(3)	(4)
Intercept	39.9585***	30.3267***	0.3824***	0.3677***
	(4.5753)	(6.4618)	(0.0360)	(0.0514)
Δ Reliability of Electricity (hours per day)	-0.9785***	-1.7795***	-0.0050***	-0.0090***
	(0.2884)	(0.3747)	(0.0020)	(0.0025)
Pre-treatment Reliability of Electricity (hours per day)		-1.3252***		-0.0069***
		(0.3814)		(0.0024)
Δ Fraction of households connected (%)	-0.0047	0.2846***	0.0002	0.0012**
	(0.0616)	(0.0701)	(0.0004)	(0.0005)
Pre-treatment Fraction of households connected (%)		0.4305***		0.0015***
		(0.0772)		(0.0006)

Table 3.4: The impact of reliability on wage rates. Robust standard errors clustered at the district level. Other control variables are years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

men's wages. While there is approximately a one-rupee reduction in women's wage rates, there's over a two-rupee reduction in men's wage rates. This is seen in the log regressions as well. Where an increase of an hour's availability reduces women's wages by 0.59% and men's wages reduce by 0.93%. Moreover, while the coefficients are significant at the 1% level for both men and women when the outcome variable is the change in levels of wage rates, the impact on women's log wages is only significant at the 5% level while the impact on men's log wages is statistically significant at the 1% level. The fraction of households with access has a positive impact on both men's and women's wages.

The more negative impact on men's wages, when compared to women's wages, may imply that there are certain gendered factors, which make women's wages more resilient to a change in reliability. Comparing the raw changes in wage rates that we can see in Table 1.1, we can observe that the wage gap between the treatment and control groups is lower for women than for men. An improvement in reliability could relieve women of the time burden of collecting fuel, and allow them to participate in the labor market. If the new women in the workforce are productive laborers, the wage rates may improve, or decrease less.

Outcome	Independent Variable	Coefficients (standard errors)			
		Levels		Logs	
		(1)	(2)	(3)	(4)
△ Agricultural Labor Wage Rate for Women (2012 Rs..)	Intercept	32.3924***	19.5856***	-0.2474***	-0.3309***
		(4.2503)	(6.0856)	(0.0402)	(0.0570)
	△ Reliability of Electricity (hours per day)	-0.5253*	-1.0663***	-0.0034	-0.0059**
		(0.2884)	(0.3573)	(0.0021)	(0.0027)
	Pre-treatment Reliability of Electricity (hours per day)		-0.8971**		-0.0040
			(0.3703)		(0.0027)
	△ Fraction of households connected (%)	-0.0276	0.2403***	-0.0000	0.0015**
		(0.0611)	(0.0680)	(0.0005)	(0.0006)
	Pre-treatment Fraction of households connected (%)		0.4085***		0.0023***
			(0.0784)		(0.0006)
△ Agricultural Labor Wage Rate for Men (2012 Rs..)	Intercept	43.7805***	36.3392***	-0.1900***	-0.1688***
		(5.2145)	(7.1990)	(0.0343)	(0.0497)
	△ Reliability of Electricity (hours per day)	-1.1410***	-2.0084***	-0.0045**	-0.0093***
		(0.3388)	(0.4313)	(0.0022)	(0.0028)
	Pre-treatment Reliability of Electricity (hours per day)		-1.4632***		-0.0083***
			(0.4204)		(0.0026)
	△ Fraction of households connected (%)	0.0326	0.3114***	0.0005	0.0012**
		(0.0689)	(0.0757)	(0.0004)	(0.0005)
	Pre-treatment Fraction of households connected (%)		0.4124***		0.0011**
			(0.0801)		(0.0005)

Table 3.5: Heterogeneous Effects: The impact of reliability on wage rates of men and women. Robust standard errors clustered at the district level. Other control variables are years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

Overall, the result that reliability has a consistent, statistically significant, negative effect on agricultural labor wage rates is perplexing, since agriculture is a practice that is largely unrelated to electricity in most Indian farms. To study the mechanisms through which reliability affects wage rates, we study reliability both at the level of villages and households, to explore the effects of reliability on labor demand and supply. In the subsequent section, we explore said mechanisms in detail. We list down hypotheses for how reliability could affect wage rates, and test them out one by one, to determine the actual channels through which reliability affects wage rates.

3.4 Mechanisms

If electricity access brings well-being and prosperity to the village, wealthier households may be more willing to pay for labor, and may even be willing to pay higher wage rates. However, this would result in a positive impact on wages, akin to the impact of the fraction of households connected, which we do not find while studying quality. Instead, we find that better quality electricity in the village has a detrimental impact on agricultural labor wages.

There are several ways in which the reliability of electricity could reduce wages for agricultural laborers. One of the modes may be through the supply of agricultural labor in villages. Since villages in India are predominantly agrarian, agricultural unskilled labor is an extremely common employment option, and changes in the supply of labor in the villages are likely to affect agricultural labor as well. Better quality electricity could increase the supply of labor in several ways. Electricity, in its simplest form, enhances the quality of life in villages. As a result, individuals may be more likely to stay in the village, and less likely to migrate out to other centers in search of seasonal employment. Literature from developing countries (De Brauw and Harigaya, 2007) finds that individuals use seasonal migration as a strategy to improve living conditions. Thus, improved standards of living, and employment opportunity due to improved electricity availability could bring migrants and their families to villages with better reliability, and could potentially saturate the agricultural labor market. Similarly, individuals may be less likely to migrate out of a village that receives better-quality electricity. In addition, the literature on migration finds that individuals may also be influenced by others around them (Epstein and Gang, 2006), and a cascade effect may alter the migration dynamics in villages. On account of such changes, the dynamics of seasonal migration may be affected adequately to significantly increase labor supply in villages, and thus reduce the wage rate of agricultural labor.

There are other ways in which the availability of better quality electricity could affect labor supply. In general, the impact evaluation literature finds a strong relationship between electricity access, and increased participation in rural labor markets and employment, particularly for women (Dinkelman, 2011; Salmon and Tanguy, 2016; Samad and Zhang, 2016). Increased participation of men and women in labor markets could increase labor supply and drive wages down. This could be driven by greater opportunities for automation, and the transition away from traditional fuels which may relieve individuals of the significant time burden of fuel collection (Njenga et al., 2021).

Wages may also be affected despite an insufficient change in labor supply, through changes in labor demand. This is a more direct channel for changes in reliability to manifest in changes in agricultural labor wages. Electricity access could affect the manner in which households practice agriculture (Singh, 2012; Badiani-Magnusson and Jessoe, 2018), or persuade agricultural households to transition away from agriculture, thus affecting the demand for unskilled labor in agriculture. For instance, if decision-makers in agricultural households actively take up other sources of income facilitated by electricity, they could require more labor, and this could drive up labor demand. Better quality electricity may also enable farmers to employ electrical farm equipment such as irrigation pumps or electrical threshers. Agricultural mechanization may supplant human labor used in farms (Baur and Iles, 2023), and the replacement could reduce labor demand. However, the presence of the machines could lead to an increased demand for workers to operate the machines, which could also affect the skill mix of labor.

Electricity also affects the time burden of several household chores. For instance, better reliability could enable households to displace the use of traditional fuels such as biomass or kerosene for heating and lighting. Replacing these fuels with electricity may save considerable time for members of the household. Similarly, other instances of electricity-facilitated automation through electrical appliances may further relieve the time burden of domestic work for individuals in the household. If the time saved due to better quality electricity is reallocated into providing labor in the fields, the household's requirement of labor may diminish considerably. This is particularly relevant for women's participation in family farms, as a reduction in the time demand for domestic work is most likely to free up time for women members of the household who could then supply labor to their household farms.

3.4.1 Labor Supply Channel

Other Unskilled Labor Wages

In order to understand why an improvement in the quality of electricity reduces wage rates, we test several hypotheses. We begin by studying the supply of labor. Since our causal agent is the quality of electricity, several facets of rural life and organization may be affected, which may affect agricultural labor wage rates. If an improvement in reliability results in an increase in labor supply, there could be a reduction in wage rates. There are multiple ways in which better electricity could increase the labor supply in the villages. Better electricity could reduce the time burden of collecting fuel for women, and increase their ability to participate in labor markets. Alternatively, an improvement in the standards of living of villages could make them a more attractive place to live and work in and migrate to. An increase in immigration or a reduction in emigration could increase the casual agricultural labor supply, which would have a negative effect on wages.

While we can argue ways in which supply can increase, it is a difficult quantity to actually test. For instance, while we can study the effect of reliability on the time spent collecting fuel, we cannot, in fact, check if that time saved is being reallocated to the labor market. Similarly, while it is possible to check if there is an impact of reliability on migration, it is still not possible to test whether there is actually an increase in supply. Since we cannot measure labor supply from the data available to us, we use an alternate approach. We postulate that since agricultural labor is casual unskilled labor, an increase in the supply of agricultural labor would mean an increase in the supply of other casual unskilled labor as well - such as domestic labor or casual non-agricultural unskilled labor.

The rationale behind the postulate lies in assuming the mobility of labor between various casual unskilled labor in the village. Thus if there is an increase in the supply of casual agricultural labor, and the wages of casual agricultural unskilled laborers fall, laborers may choose to find work instead as casual non-agricultural unskilled labor. This would thereby distribute the surplus labor among different unskilled labor jobs and lead to a proportional reduction in wage rates. There are cases wherein laborers may opt for a better-paying occupation in their village even if it is underpaid in comparison to the same job in other villages. In most villages, the wage rates are higher for casual non-agricultural unskilled labor, followed by casual agricultural labor, and domestic labor rates were the least. Therefore, even if the surplus supply in agriculture resists mobility to domestic labor due to lower wages, one would still expect a movement to casual non-agricultural work. Therefore, if there were an electricity-induced increase in labor supply, the wage rates of all unskilled labor should suffer.

Therefore, we extend our analysis of agricultural unskilled labor wages to domestic labor wages and non-agricultural unskilled labor wages. Table 3.6 presents the results of our difference-in-difference analysis. Unlike agricultural labor wages, we find no statistically significant impact on the wages of domestic labor and non-agricultural unskilled labor. We find a weak positive effect of the change in reliability on the levels and logs of casual non-agricultural unskilled labor wage rates, although this is not statistically significant at all. With domestic labor, there is a weak negative effect on the level of wages even though this is statistically significant only at the 10% level. However, there is no statistically significant impact whatsoever on the logarithm of domestic labor wage rates.

The absence of effects on other casual unskilled labor wage rates implies that the fall in wages with a rise in reliability is unique to agricultural labor wages. This implies that the reduction

Outcome	Independent Variable	Coefficients (standard errors)			
		Levels		Logs	
		(1)	(2)	(3)	(4)
$N = 1345$					
	Intercept	-30.1812	10.4254	0.3177***	0.3698***
		(40.0535)	(45.7014)	(0.0762)	(0.0947)
	Δ Reliability of Electricity (hours per day)	3.3011	1.6408	0.0037	0.0007
		(2.0901)	(2.3437)	(0.0042)	(0.0047)
Δ Non-Agricultural Unskilled Labor Wage Rate (2012 Rs..)	Pre-treatment Reliability of Electricity (hours per day)		-3.2040		-0.0056
			(2.3324)		(0.0047)
	Δ Fraction of households connected (%)	-0.0503	-0.1281	-0.0006	-0.0005
		(0.3869)	(0.5337)	(0.0008)	(0.0010)
	Pre-treatment Fraction of households connected (%)		-0.1613		0.0001
			(0.6136)		(0.0012)
$N = 629$					
	Intercept	50.0642***	39.9492***	0.5354***	0.5331***
		(9.5847)	(12.4340)	(0.0747)	(0.1062)
	Δ Reliability of Electricity (hours per day)	-0.4066	-1.0493*	-0.0008	-0.0004
		(0.4607)	(0.6176)	(0.0038)	(0.0051)
Δ Domestic Labor Wage Rate (2012 Rs.)	Pre-treatment Reliability of Electricity (hours per day)		-0.9947		0.0007
			(0.8148)		(0.0064)
	Δ Fraction of households connected (%)	-0.0612	0.1826	-0.0005	-0.0006
		(0.1395)	(0.1993)	(0.0011)	(0.0014)
	Pre-treatment Fraction of households connected (%)		0.3653**		-0.0001
			(0.1754)		(0.0013)

Table 3.6: Labor Supply Channel: The impact of reliability on wage rates for domestic labor and Casual non-agricultural unskilled labor. Robust standard errors clustered at the district level. Other control variables are years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

in casual agricultural labor wages is not simply due to a rise in labor supply due to improved electricity quality, for that would result in a similar reduction in the wages of non-agricultural work, which we do not observe from the data. Instead, we find that the effect of reliability on wages is restricted to agricultural labor wages alone. Since non-agricultural labor wages are also higher on average, there is an added incentive for workers to move to non-agricultural work. However, we do not find such an effect. Therefore, the problem may not entirely be related to supply, if at all, and it is possible that the quality of electricity is relevant to how agriculture is practiced in villages, and improved reliability may be reducing demand for labor in agriculture specifically, rather than increasing supply and affecting wages across different unskilled labor jobs.

Women's time allocation to domestic chores

However, we cannot entirely dismiss an increase in supply as a factor for falling wages, as our reasoning was based on the assumption that all forms of casual labor are fungible, i.e., laborers are

$N = 4157$	Coefficients (standard errors)
Intercept	-40.4270
	(61.2623)
Δ Reliability of Electricity (hours per day)	-9.0792***
	(2.1023)
Pre-treatment Reliability of Electricity (hours per day)	-1.1681
	(2.1246)

Table 3.7: Labor Supply Channel: The impact of reliability on the time spent by women in fuel collection. Robust standard errors clustered at the village level. Other control variables are the fraction of households connected in the village, the partial sum of the fraction of households connected in the village pre-treatment, years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

equally skilled in, and willing to provide all forms of labor, causing the surplus labor in one sector to spillover to other labor markets. This may not necessarily be true. For instance, since the wage rates of domestic labor are lower than those of agricultural labor, individuals may continue to work in agriculture, despite the relatively lower wages, when compared to other sectors. There may also be a gendered component to which jobs an individual is willing to do. Women may be less likely to move to some kinds of non-agricultural labor such as construction work. Similarly, men may be less likely to do domestic work, and thus an increase in the supply of agricultural labor may not necessarily lead to an increase in the supply of other forms of casual labor.

One channel through which reliability could reduce the time burden of women in domestic chores may be if reliability reduces the time spent by women in collecting traditional fuels. We test to see if reliability has a statistically significant negative impact on the time spent by women collecting fuel. We only consider households that have access to electricity in both rounds of the survey, so as to be able to isolate the effect of reliability. We then introduce the time spent by women in fuel collection as a control variable and check if this renders the coefficient of reliability statistically insignificant.

In their comprehensive study of the role of reliability in household well-being in India, Samad and Zhang (2016) investigate the impact of reliability on the time spent in fuel collection. However, they find a positive impact of reliability on the time spent in fuel collection. They find that for every one-hour increase in the availability of electricity, there is a 0.17-hour increase in the time spent in fuel collection per month, significant at the 10% level. Therefore, the authors find that better quality electricity leads to higher consumption of traditional fuels and more time spent on fuel collection, which is somewhat counterintuitive. However, the authors find an overall negative aggregate effect on the time spent in fuel collection when they consider households getting connected, along with reliability. But due to the reasons already discussed, the results are difficult to interpret, and we run tests independently considering only households that have access in both time periods, so as to filter out the effect of receiving access.

Table 3.7 presents the results of the DID regression studying the impact of reliability on the time spent by women in fuel collection. We observe an extremely statistically significant (at the 1% level) negative impact of reliability on the time spent in fuel collection. We find that for every additional hour of increased electricity availability, women save over nine minutes per week that they would have otherwise spent collecting fuel. To contrast this with Samad and Zhang's results, we find that every additional hour of electricity availability frees up over 36 minutes or 0.6 hours per month, as opposed to them finding a 0.17-hour increase in the time burden. We attribute this difference in results to our design where we only study households that show no change in their status of being connected, as opposed to including households that may have received new connections or lost existing connections in the period.

Since we find a strong negative effect of reliability on the time spent in fuel collection, we can conclude that better quality electricity relieves women of the time burden of fuel collection and provides them with more time on their hands that they can now spend either in leisure, other household work or by participating in the labor market.

3.4.2 Labor Demand Channel

Another way in which labor wages could be detrimentally affected would be if there is a fall in demand, regardless of whether supply increases or not. A simultaneous decrease in demand and increase in supply would, of course, lead to lower wages. However, even if supply is held constant, a fall in demand could lead to a reduction in wages. A parallel fall in demand and supply together may lead to more unpredictable behavior in terms of labor wages

Unlike casual agricultural labor supply, where the willingness of an individual to work was difficult to quantify, the demand for labor can be studied at the level of the households which hire

Outcome $N = 5249$	Independent Variable	Coefficients (standard errors) (1)	(2)
Δ Agricultural Expenditure on Hired Farm Labor(2012 Rs. per year)	Intercept	-340.4123	-798.4741
		(1546.2548)	(1536.1447)
	Δ Reliability of Electricity (hours per day)	-166.4984**	-134.9467*
		(72.2182)	(69.8937)
	Pre-treatment Reliability of Electricity (hours per day)	-90.9638	-59.0632
		(66.7283)	(67.4040)
	Δ Agricultural Labor Wage Rate (2012 Rs.)		16.4864**
			(7.0029)

Table 3.8: Labor Demand Channel: The impact of reliability on the household expenditure on hired agricultural labor. Robust standard errors clustered at the village level. Other control variables are the fraction of households connected in the village, the partial sum of the fraction of households connected in the village pre-treatment, years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

labor. The IHDS surveys ask agricultural households the amount spent on hired farm labor the previous year, which can be considered a proxy measure for the agricultural household's willingness to pay for hired farm labor. Therefore, we can study if electrical reliability has an effect on the expenditure of an agricultural household on hired labor. If the coefficient of the change in reliability is positive, it implies that better reliability leads to households spending more on labor. To study the willingness to pay, we use a difference-in-differences approach. However, instead of using village-level variables alone, we bring the analysis to the level of the household. Since the IHDS surveys have data for reliability at the household level, and since there is considerable parity between reliability reported at the level of households and at the level of villages, we can use the change in reliability measured at the level of the household as a treatment variable. Since our treatment and outcome variables are at the level of the household, we also include several household-level characteristics such as the number of adult men and women in the household, the presence of flush toilets, the presence of separate kitchens, etc. We also include village-level development characteristics such as the presence of metalled roads, NGOs/development organizations, proximity to markets, banks, and credit cooperatives, and the number of government and private primary, middle, secondary, and higher secondary schools, and primary medical centers. To account for the spillover benefits of electricity, we also include the fraction of connected households in the village and the number of years since the village was connected.

To implement a proper DID analysis, we restrict ourselves to households that already had access to electricity in the first wave of the survey, as an observable change in reliability from 0 hours in a household that did not have electricity, to begin with, need not be attributed entirely to a change in reliability alone, and there is an access factor as well. As we had done with the village-level analysis, we assume that the changes in reliability are random at the level of the household, among households that were already connected to the grid. Since we consider the total expenditure by the landowning household as a metric for demand, it is also important to include the changes in the village-level wage rate as a control variable. This is because a simple reduction in the wage rate due to an unrelated reason could also bring about lesser labor spending. A reduction in household expenditure on farm labor only through reduced wage rates, without a fall in demand directly induced through reliability, could imply that reliability and wage rates are correlated with each other, without reliability actually being correlated with labor demand. Thus, it is important to control for the wage rate and study if there is a reduction in demand that can be attributed directly to changes in reliability, rather than an effect of reliability acting only through reduced wage rates.

Table 3.8 presents the results for the impact of a change in reliability on household expenditure on hired agricultural labor. There is a negative impact on the expenditure on hired agricultural labor. For every one-hour increase in reliability, a household spent nearly Rs.135 less on hired labor, although it is statistically significant only at the 10% level, which indicates that the effect may be somewhat "fuzzy". A negative effect, however, implies that there is a reduction in household expenditure on hired farm labor when the reliability improves, and therefore a better quality of electricity may be associated with falling demand for labor in the agricultural sector. As expected, there's a significantly positive relationship between the change in wage rates and the total expenditure, but it is noteworthy, that despite controlling for changes in wages, the effect of reliability does not vanish.

In the subsequent subsections, we explore the channels through which an improvement in reliability could result in the observed fall in household expenditure on agricultural labor. For any channel to successfully explain the possible fall in demand, two conditions need to be met. Firstly, an increase in reliability should cause a change in a potential explanatory variable (explanatory with respect to the fall in demand). Secondly, the fall in demand should be explained by the variable, i.e., reliability should affect the fall in expenditure through this variable. To test this statistically, we first check if the change in reliability leads to statistically significant changes in the channel variable that could potentially resolve our problem. Along with it, the variable must satisfy another criterion - the introduction of said variable as a control variable should render statistically insignificant, the coefficient of reliability on household expenditure on hired labor. Only

$N = 4993$	Coefficients (standard errors)
Intercept	-0.1975
	(0.1387)
Δ Reliability of Electricity (hours per day)	0.0038
	(0.0042)
Pre-treatment Reliability of Electricity (hours per day)	0.0073*
	(0.0043)

Table 3.9: Labor Demand Channel: The impact of reliability on the share of agriculture in income. Robust standard errors clustered at the village level. Other control variables are the fraction of households connected in the village, the partial sum of the fraction of households connected in the village pre-treatment, years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

if a variable satisfies these two statistical tests do we accept the variable and the associated channel as the explanation for the observed negative impact of reliability on labor demand.

Share of Agriculture in Income

The quality of electricity could potentially affect the demand for agricultural labor by facilitating the transition of rural society away from agriculture. Good quality electricity may allow households to enter businesses that they had hitherto been excluded from. Better quality power could allow electricity-dependent businesses to flourish and not merely supplement a household's agricultural income, but perhaps motivate them to move out of agriculture entirely. This could affect capital such as land holding which could be traded for different forms of capital, based on the demands of the businesses, and reduced investment in agriculture could lead to falling demand for labor, observed in our analysis thus far.

The literature, however, on reliability and household well-being does not lend sufficient support to the hypothesis that improvements in reliability could reduce a household's share of agriculture in income. Samad and Zhang (2016) study the effect of electricity access and reliability on the various components of a household's income on a set of rural households using the IHDS data. They find a negative impact on a household's agricultural and non-agricultural income, although the magnitude of the effect is larger for non-agricultural income, with every additional hour of electricity availability leading to a 2% reduction in non-farm income, as opposed to a meager 0.4% reduction in agriculture. Moreover, the effect on agricultural income is not statistically significant, while the effect on non-agricultural income is significant at the 5% level.

However, Samad and Zhang's results are difficult to interpret. The study does not merely study the effect of reliability but rather studies the effect of access as well, and the reliability of electricity is defined as an indicator variable for whether a household has access to electricity multiplied by the number of hours of electricity available. This definition of reliability makes it difficult to isolate the effects of reliability from the effects of access overall. Instead, in our analysis, we simplify the problem further by considering only households that already have electricity. Therefore, the only treatment is the change in the level of reliability alone, as whether a household has access or not is fixed.

We construct a variable for the share of agricultural in income, which is the ratio of agricultural income to total income in a household. If reliability is found to have a positive or neutral effect on the share of agriculture in income, we cannot attribute the possible fall in demand for laborers to the share of agriculture. On the contrary, if there is a negative effect of reliability on the share of agricultural income in a household, it may lead to falling demand. Then if the introduction of the share of agricultural income as a control variable in the regression studying the impact on labor demand, eradicates the effect of reliability measured, we can successfully attribute the effect of reliability on agricultural labor demand to the falling share of agriculture in income.

Table 3.9 presents the results for the effect of reliability on the share of agriculture in household income. We find a weak positive impact on reliability, and according to the results. A one-hour improvement in reliability leads to a 0.038 increase in the fraction of a household's income from agriculture. But this effect has a large standard error and is not statistically significant. In the absence of a statistically significant effect, nothing concrete can be said about the impact of reliability on the share of farm income, and one cannot ascertain whether a household's share of agricultural income would increase or decrease if the quality of electricity improves. Therefore, the channel fails our first test, and we deduce that it is unlikely that reliability affected labor wage rates through a falling share of agriculture in household income.

$N = 6903$	Coefficients (standard errors)
Intercept	0.1144**
	(0.0486)
Δ Reliability of Electricity (hours per day)	-0.0031
	(0.0020)
Pre-treatment Reliability of Electricity (hours per day)	0.0007
	(0.0021)

Table 3.10: Labor Demand Channel: The impact of reliability on electric pump ownership. Robust standard errors clustered at the village level. Other control variables are the fraction of households connected in the village, the partial sum of the fraction of households connected in the village pre-treatment, years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

Increased Usage of Electrical Machinery

In the absence of a reduced share of agriculture in income attributed to the reliability, there may be other factors that could influence the demand for labor. One of the few ways in which agriculture could be affected in the typical Indian farm could be through the introduction of electrical farm equipment such as irrigation pumps. The introduction of electrical appliances made usable through improvements in reliability, could potentially replace casual agricultural laborers in certain agricultural work such as irrigation. The opposite is possible too, however, where machinery, instead of replacing labor through automation, modern electrical farm equipment could potentially increase the demand for labor, as land-owning farmers may need to hire labor to operate the electricity-powered machinery.

In our study, we would have ideally liked to have a variable that measures the prevalence of electrical farm machines. However, we only have data for the ownership of such appliances. Since it is possible that households may own machinery such as irrigation pumps, electricity poor in quality may render the machines unusable, and the seemingly stochastic improvement in reliability may merely facilitate the usage of the machines, rather than actually drive up ownership of the machines. Thus a regression that simply studies ownership of irrigation pumps may not necessarily paint an accurate picture of the role of reliability in increasing the usage of electrical machinery in farms.

$N = 5214$	Coefficients (standard errors) (1)	(2)
Intercept	-798.4741	-1364.6407
	(1536.1447)	(1555.8046)
Δ Reliability of Electricity (hours per day)	-134.9467*	-128.2940*
	(69.8937)	(71.7529)
Pre-treatment Reliability of Electricity (hours per day)	-59.0632	-71.6028
	(67.4040)	(67.5162)
Δ Number of Irrigation Pumps Owned by the Household		4090.9574***
		(1140.5852)

Table 3.11: Labor Demand Channel: The impact of reliability on the total expenditure on hired farm labor, controlled for irrigation pump ownership. Robust standard errors clustered at the village level. Other control variables are the fraction of households connected in the village, the partial sum of the fraction of households connected in the village pre-treatment, years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

From Table 3.10, we can observe that reliability has no significant impact on pump ownership. This may be because a farmer need not need to be able to operate pumps throughout the day, and a couple of hours of uninterrupted power may be sufficient for the farmer's needs. However, we note that the ownership of pumps need not necessarily imply that pumps can be used to their full potential. It is possible that farmers own pumps but may not be able to reap the full benefits due to poor reliability. Therefore, it may be useful to also study whether pump ownership in itself has a negative impact on expenditure on hired farm labor. Thus, we move on to our second statistical test where we incorporate the number of pumps owned as a control variable for the impact of reliability on a household's expenditure on hired agricultural labor. While we do not expect the ownership of pumps to drive down the effect of reliability, as there is little correlation between reliability and changes in pump ownership (Table 3.10), this regression also helps us study whether pumps have a positive or negative impact on labor demand.

Table 3.11 presents the results for the regression that studies the effect of a change in reliability on the total expenditure on hired agricultural labor, controlled for the change in the number of irrigation pumps owned. For pump ownership and usage to have adequately explained the fall in labor demand, we would have required a negative coefficient on pump ownership, and we would've simultaneously required that the effect of reliability disappear. One can observe from the table that neither of these two criteria is met. The ownership of pumps has a large positive effect on expenditure on labor. A household spends an additional over Rs. 4000 per year on hired labor for each additional irrigation pump it owns. Not only is there a largely positive relationship between pump ownership and expenditure on hired labor, but the effect is also extremely statistically significant (at the 1% level). Therefore, we can reject the hypothesis that increased pump usage replaces labor. The alternate hypothesis that more labor is required to operate the pumps is more likely. Furthermore, the coefficients of the change in reliability and the pre-treatment reliability are similar with or without the incorporation of pump ownership as a control variable, with the standard errors being similar too. As for the question of why the better usability of pumps with better reliability does not adequately create an increased demand for labor, there could be multiple reasons. Irrigation pumps are owned by a small fraction of households (less than 10%), and changes in ownership are smaller still, and thus may not be sufficient to drive up demand for labor.

Reallocation of Time Saved in Fuel Collection

The final hypothesis that we test to explain the fall in demand for agricultural labor is that an increase in reliability reduces the time burden of fuel collection of women, due to which women of the household have more time available which they may allocate to working in the farmland owned by the household. In India and other developing countries, agriculture takes the form of a household enterprise where multiple members of the agricultural household put in labor on their farms. Family farming is one of the most common forms of practiced agriculture and rural India and could potentially explain the fall in labor demand. Women working in their household farms would mimic volunteer labor for a firm, and drive down the demand for additional hired labor.

Since most agricultural households in rural India practice family farming, the women of the households with more time on their hands could also provide labor in the fields and reduce the household's expenditure on hired casual labor. To test this, we perform a regression with the change in the total household expenditure on hired labor as the outcome variable with a change in reliability as the treatment. We use the time spent by women collecting fuel as a control variable in the regression. We know from table 3.8 that in the absence of such a control variable, reliability has a strong negative impact on demand. If the inclusion of time spent in fuel collection as a control variable removes the effect of reliability or renders it statistically insignificant, we can attribute the reliability-induced fall in demand to have acted through the time saved by women in fuel

$N = 1427$	Coefficients (standard errors) (1)	(2)
Intercept	563.5701	3312.6453*
	(1503.5822)	(1888.2954)
Δ Reliability of Electricity (hours per day)	-139.1821**	-79.5313
	(64.1025)	(69.3277)
Pre-treatment Reliability of Electricity (hours per day)	-138.1341**	-30.4178
	(57.0345)	(68.2461)
Δ Time Spent by Women in Fuel Collection (minutes per week)		0.8810
		(0.6616)

Table 3.12: Labor Demand Channel: The impact of reliability on the total expenditure on hired farm labor, controlled for the time spent by women in fuel collection. Robust standard errors clustered at the village level. Other control variables are the fraction of households connected in the village, the partial sum of the fraction of households connected in the village pre-treatment, years since the village was first connected, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, changes in whether there are metalled roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers.
*Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

collection. From Table 3.7, we know that better reliability leads to less time spent by women on fuel collection. Therefore, it remains to be seen whether controlling for the time spent by women, causes the effect of reliability to cease to exist.

Table 3.12 presents the results of the regressions studying the impact of reliability on the total expenditure on hired agricultural labor, with and without the time spent by women in fuel collection as a control variable. We can observe that while there is a statistically significant negative impact of reliability on expenditure when the time spent by women is not accounted for, the effect is both cut down in size and significance with the inclusion of the time spent by women in fuel collection as a control variable. The size of the effect for change in reliability is less than half after controlling for time spent in fuel collection, with a small increase in the standard error.

The results of Table 3.7 and Table 3.12 show that the time spent in fuel collection passed both statistical tests required to explain the negative effect of reliability on labor demand. We, therefore, speculate that village-level reliability has a negative effect on casual agricultural labor wages, an effect that is likely driven by better quality electricity reducing the time burden of fuel collection on women, who spend the surplus hours by providing labor on their household farms, which reduces the demand for hired labor, and thereby has a detrimental effect on wage rates.

Chapter 4

Conclusion

This thesis explores the causal impacts of rural electrification - the good and the bad, using household and village-level data from India. Using data from three panels spread across 18 years, I estimate the causal effects of electricity access on household well-being by studying various development indicators such as consumption, agricultural income, agricultural land ownership, years of schooling, kerosene consumption, and time spent collecting fuel, contrasting the short-term impacts and the long-term impacts of electrification. The time-associated benefits of electricity access are relatively less studied in the impact evaluation literature. Our findings indicate that for most of the outcomes studied, the impacts (if any) of electrification were larger for households that had been connected to the grid longer. The impact on per-capita consumption was found to be two and a half times as large in households that had had access to electricity for over seven years when compared to households that had been connected for fewer than seven years. Statistically significant impacts were also found on the years of schooling of the most educated adult in the household, and on the time spent by women in fuel collection - the long-term effect being the statistically significant one for all three variables.

On investigating the heterogeneity in the results, we found that electricity access may affect different groups differently. For instance, we found that there were typically a greater number of large and statistically significant effects on poor households, than on rich households. While the coefficients measured for the consumption of poor households were larger and more significant on the estimates made for the entire population, there was no statistically significant impact of electricity access on consumption in rich households. Similarly, the effect on agricultural land-holding was also statistically significant only for poor households that had been connected to the grid for over seven years. On the contrary, the effect of electricity access on the time spent in fuel collection is larger, and more statistically significant for richer households. We found similar differences in the effects estimated in the less-developed states of Central India and the development leaders of South India. We found that although long-term effects were found to be consistently greater than short-term effects, the outcomes affected in different regions were different. While consumption showed a statistically significant response only in the less developed Central Indian States, the level of schooling and kerosene consumption showed statistically significant responses only in South India.

Being one of the first studies to contrast the impacts of short-term access and long-term access in a developing country, the short-term versus long-term analysis adds to what we already know about the impacts of electricity access by introducing a new dimension - time. We find that the long-term returns could be significantly higher and researchers and policymakers should look to longer studies to gauge the true potential benefits of rural electrification. We also speculate whether the reason experimental impact evaluation studies are found to estimate lesser positive impacts of electrification than observational studies is the typically shorter duration of experimental studies, which may lead to weaker estimates if the impacts of electrification grow over time. We observe that even observational studies observe positive impacts less frequently if they are constrained by the duration of the study. Most observational studies that find strong effects usually follow households that have been connected to the grid for years, a setting that would be extremely rare in randomized controlled trials due to logistical constraints. We even employ a regression to study how various designs affect the frequency of positive impact measurements and find that the most significant impact is that of time.

In Chapter 3, we study the spillover effects of electricity access. We find that households not connected to the grid benefit from a greater rate of access among other households in the village. We find statistically significant effects of the fraction of households connected in a village on the consumption of households without access. Another relatively unexplored facet of electricity is the quality. The reliability of electricity i.e. the frequency or duration of power outages determines whether villages and households are merely nominally connected to the grid, or whether electricity is available for long enough hours to sufficiently have positive effects on people's well-being. In this thesis, we also investigate the relationship between the hours of electricity available to villages and the agricultural labor wages in the village. We find that the reliability of electricity in a village has a statistically significant negative impact on agricultural labor wages, as high as a

0.9% reduction in casual agricultural labor wage rates for every additional hour's increase in the reliability of electricity. On comparing with other wage rates for unskilled labor in the villages, we find that the negative impact of reliability is exclusive to agricultural labor wage rates and other unskilled labor such as domestic labor and casual non-agricultural unskilled work is unaffected by reliability. Based on the observation, we find the supply of labor to be an unlikely causal agent.

Instead, we focus our attention on labor demand, where we find the total household expenditure on hired-agricultural labor is driven down by improvements in reliability. We test several hypotheses to investigate why reliability drives down labor demand in agriculture. We reject the hypothesis that the lower demand is caused due to electricity enabling households to transition away from agriculture by providing avenues for new businesses. The hypothesis is ruled out as the reliability of electricity is not found to have any consequences on the share of a household's income that comes from agriculture. Our second hypothesis is based on the possibility of agricultural mechanization leads to human labor being supplanted by electrical agricultural machinery such as irrigation pumps. We find that reliability has no impact on the ownership of pumps. Furthermore, the effect of reliability on expenditure on hired labor does not go away even after controlling for pump ownership. Instead, we find that the ownership of electrical pumps drives up demand for agricultural labor, as households that own more pumps spent more on hired labor, after controlling for other confounding factors. This finding is noteworthy to labor economics, where the question of technology replacing labor is an old one, although relatively unexplored in the context of reliability and agriculture. The result most likely signifies that the ownership of pumps creates greater demand for workers to operate the pumps, rather than replacing them.

The final hypothesis we study is that improved reliability results in reduced usage of traditional fuels, relieving the household's women of the time burden of fuel collection, enabling them to supply labor on their family farms, and diminishing the need for additional labor. We confirm statistically, that a change in reliability has a negative impact on the time spent by women in fuel collection, indicating that women have more time on their hands in villages with better reliability. Speculating on whether it is the participation of these women that drives down demand, we control for the time spent by women on fuel collection in the estimation, and find that doing so renders the effect of reliability insignificant, implying that reallocation of time saved in domestic chores could be a potential channel. This has several implications. Firstly, it demonstrates that changes in the quality of electricity can potentially influence lifestyles and alter the time commitments in domestic life and work. It additionally demonstrates that better reliability can lead to greater involvement of women in family farms, leading to a strengthening of the institution, with family members replacing hired farm labor. The result also implies that a reduced time burden of domestic work does not necessarily lead to greater labor market participation of women. A woman working on a farm owned by her family may be beneficial for the household in cutting labor costs, but it keeps the women of the household away from financial independence.

The results of this study from India in the early 21st century are relevant for several other countries in the developing world, currently investing in grid roll-out programs. This is because the state of infrastructure development and energy access and poverty in several countries in the Global South, predominantly in Africa, are akin to their state in India around the time of the IHDS surveys. The results of the thesis are also relevant to several countries in the world plagued by power outages. Countries that have brought access to most of their villages and households would be looking towards improving the quality of those connections over the next few decades, and a comprehensive knowledge of the effects of changes in reliability may help policymakers in the countries to plan the roll-out more smoothly.

Bibliography

Anna Margret Aevarsdottir, Nicholas Barton, and Tessa Bold. The impacts of rural electrification on labor supply, income and health: experimental evidence with solar lamps in tanzania. *Unpublished manuscript, June*, 2017.

Julio Aguirre. The impact of rural electrification on education: A case study from peru. *The Lahore Journal of Economics*, 22(1):91, 2017.

Hunt Allcott, Allan Collard-Wexler, and Stephen D O'Connell. How do electricity shortages affect industry? evidence from india. *American Economic Review*, 106(3):587–624, 2016.

Steven G Allen. Technology and the wage structure. *Journal of Labor Economics*, 19(2):440–483, 2001.

Reena Badiani-Magnusson and Katrina Jessoe. Electricity prices, groundwater, and agriculture: the environmental and agricultural impacts of electricity subsidies in india. In *Agricultural productivity and producer behavior*, pages 157–183. University of Chicago Press, 2018.

Abhijit Vinayak Banerjee, Esther Duflo, and Michael Kremer. The influence of randomized controlled trials on development economics research and on development policy. *The state of Economics, the state of the world*, pages 482–488, 2016.

Douglas F Barnes and Hans P Binswanger. Impact of rural electrification and infrastructure on agricultural changes, 1966-1980. *Economic and Political Weekly*, pages 26–34, 1986.

Manuel Barron and Maximo Torero. Household electrification and indoor air pollution. *Journal of Environmental Economics and Management*, 86:81–92, 2017.

Patrick Baur and Alastair Iles. Replacing humans with machines: A historical look at technology politics in california agriculture. *Agriculture and Human Values*, 40(1):113–140, 2023.

Patrick Bayer, Ryan Kennedy, Joonseok Yang, and Johannes Urpelainen. The need for impact evaluation in electricity access research. *Energy Policy*, 137:111099, 2020.

Gunther Bensch, Jochen Kluve, and Jörg Peters. Impacts of rural electrification in rwanda. *Journal of Development Effectiveness*, 3(4):567–588, 2011.

Clair Brown and Benjamin A Campbell. The impact of technological change on work and wages. *Industrial Relations: A Journal of Economy and Society*, 41(1):1–33, 2002.

Stephan B Bruns, Christian Gross, and David I Stern. Is there really granger causality between energy use and output? *The Energy Journal*, 35(4), 2014.

Fiona Burlig and Louis Preonas. Out of the darkness and into the light? development effects of rural electrification. *Energy Institute at Haas WP*, 268:26, 2016.

Ujjayant Chakravorty, Martino Pelli, and Beyza Ural Marchand. Does the quality of electricity matter? evidence from rural india. *Journal of Economic Behavior & Organization*, 107:228–247, 2014.

Ujjayant Chakravorty, Kyle Emerick, and Majah-Leah Ravago. Lighting up the last mile: The benefits and costs of extending electricity to the rural poor. *Resources for the Future Discussion Paper*, pages 16–22, 2016.

Stephen J Connellan. Lung diseases associated with hydrocarbon exposure. *Respiratory Medicine*, 126:46–51, 2017.

Cynthia C Cook. *Assessing the impact of transport and energy infrastructure on poverty reduction*. Asian Development Bank, 2005.

Duc Anh Dang and Hai Anh La. Does electricity reliability matter? evidence from rural viet nam. *Energy Policy*, 131:399–409, 2019.

Alan De Brauw and Tomoko Harigaya. Seasonal migration and improving living standards in vietnam. *American Journal of Agricultural Economics*, 89(2):430–447, 2007.

Sonalde Desai, Reeve Vanneman, and National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS). *Inter-university Consortium for Political and Social Research [distributor], 2018-08-08.*, 2005.

Sonalde Desai, Reeve Vanneman, and National Council of Applied Economic Research, New Delhi. India Human Development Survey-II (IHDS-II). *Inter-university Consortium for Political and Social Research [distributor], 2018-08-08.*, 2011-2012.

Taryn Dinkelman. The effects of rural electrification on employment: New evidence from south africa. *American Economic Review*, 101(7):3078–3108, 2011.

Gil S Epstein and Ira N Gang. The influence of others on migration plans. *Review of Development Economics*, 10(4):652–665, 2006.

Roger Fouquet. *Heat, power and light: revolutions in energy services*. Edward Elgar Publishing, 2008.

Madhusudan Ghosh. Economic growth and human development in indian states. *Economic and Political Weekly*, pages 3321–3329, 2006.

John Gibson and Susan Olivia. The effect of infrastructure access and quality on non-farm enterprises in rural indonesia. *World Development*, 38(5):717–726, 2010.

David Hamburger, Joel Jaeger, Patrick Bayer, Ryan Kennedy, Joonseok Yang, and Johannes Urpelainen. Shades of darkness or light? a systematic review of geographic bias in impact evaluations of electricity access. *Energy Research & Social Science*, 58:101236, 2019.

Putri Wahyu Handayani, Rus' an Nasrudin, and Jahen Fachrul Rezki. Reliable electricity access, micro-small enterprises, and poverty reduction in indonesia. *Bulletin of Indonesian Economic Studies*, (just-accepted):1–48, 2023.

Guido W Imbens. The role of the propensity score in estimating dose-response functions. *Biometrika*, 87(3):706–710, 2000.

Astrid Kander, Paolo Malanima, and Paul Warde. *Power to the People: Energy in Europe over the Last Five Centuries.* Princeton University Press, 2014.

Shahidur R Khandker, Douglas F Barnes, Hussain A Samad, and Nguyen Huu Minh. Welfare impacts of rural electrification: evidence from vietnam. *World Bank Policy Research Working Paper*, (5057), 2009.

Shahidur R Khandker, Douglas F Barnes, and Hussain A Samad. The welfare impacts of rural electrification in bangladesh. *The Energy Journal*, 33(1), 2012.

Shahidur R Khandker, Douglas F Barnes, and Hussain A Samad. Welfare impacts of rural electrification: A panel data analysis from vietnam. *Economic Development and Cultural Change*, 61 (3):659–692, 2013.

Shahidur R Khandker, Hussain A Samad, Rubaba Ali, and Douglas F Barnes. Who benefits most from rural electrification? evidence in india. *The Energy Journal*, 35(2), 2014.

Kenneth Lee, Edward Miguel, and Catherine Wolfram. Does household electrification supercharge economic development? *Journal of Economic Perspectives*, 34(1):122–44, 2020a.

Kenneth Lee, Edward Miguel, and Catherine Wolfram. Experimental evidence on the economics of rural electrification. *Journal of Political Economy*, 128(4):1523–1565, 2020b.

Joshua Lewis and Edson Severnini. Short-and long-run impacts of rural electrification: evidence from the historical rollout of the us power grid. *Journal of Development Economics*, 143:102412, 2020.

NCAER. Human Development Profile of India (HDPI). *National Council on Applied Economic Research. New Delhi.*, 1994.

M Njenga, J Ka Gitau, and Rc Mendum. Women's work is never done: Lifting the gendered burden of firewood collection and household energy use in kenya. *Energy Research & Social Science*, 77:102071, 2021.

Sheoli Pargal and Sudeshna Ghosh Banerjee. *More power to India: The challenge of electricity distribution*. Washington, DC: World Bank, 2014.

Vanessa Mae Pepino, Majah-Leah V Ravago, and Karl Jandoc. Does quality of electricity matter? household-level evidence from the philippines. *Journal of the Asia Pacific Economy*, 26(4): 720–747, 2021.

Jörg Peters and Maximiliane Sievert. Impacts of rural electrification revisited–the african context. *Journal of Development Effectiveness*, 8(3):327–345, 2016.

Amod K Pokhrel, Michael N Bates, Sharat C Verma, Hari S Joshi, Chandrashekhar T Sreeramareddy, and Kirk R Smith. Tuberculosis and indoor biomass and kerosene use in nepal: a case–control study. *Environmental health perspectives*, 118(4):558–564, 2010.

Rajendra Prasad, Singh Abhijeet, Rajiv Garg, and Giridhar B Hosmane. Biomass fuel exposure and respiratory diseases in india. *Bioscience trends*, 6(5):219–228, 2012.

Lant Pritchett and Justin Sandefur. Learning from experiments when context matters. *American Economic Review*, 105(5):471–75, 2015.

Narasimha D Rao. Does (better) electricity supply increase household enterprise income in india? *Energy policy*, 57:532–541, 2013.

Martin Ravallion. On multidimensional indices of poverty. *The Journal of Economic Inequality*, 9:235–248, 2011.

Juan Pablo Rud. Electricity provision and industrial development: Evidence from india. *Journal of development Economics*, 97(2):352–367, 2012.

Jeffrey Sachs, Christian Kroll, Guillame Lafortune, Grayson Fuller, and Finn Woelm. *Sustainable development report 2022*. Cambridge University Press, 2022.

Claire Salmon and Jeremy Tanguy. Rural electrification and household labor supply: Evidence from nigeria. *World development*, 82:48–68, 2016.

Hussain A Samad and Fan Zhang. Benefits of electrification and the role of reliability: evidence from india. *World Bank policy research working paper*, (7889), 2016.

Amartya Sen. Social choice and justice: A review article, 1985.

Amartya Sen. *Inequality reexamined*. Harvard University Press, 1995.

Amartya Sen. *Choice, welfare and measurement*. Harvard University Press, 1997.

Karam Singh. Electricity subsidy in punjab agriculture: Extent and impact. *Indian Journal of Agricultural Economics*, 67(902-2016-66726), 2012.

Andrew Sumner. Economic well-being and non-economic well-being. *A Review of the Meaning and Measurement of Poverty*, 30(49):33–50, 2004.

Daniel Robert Thomas, SP Harish, Ryan Kennedy, and Johannes Urpelainen. The effects of rural electrification in india: An instrumental variable approach at the household level. *Journal of Development Economics*, 146:102520, 2020.

Dominique Van de Walle, Martin Ravallion, Vibhuti Mendiratta, and Gayatri Koolwal. Long-term gains from electrification in rural india. *The World Bank Economic Review*, 31(2):385–411, 2017.

Appendix A

Propensity-Score Weighted Regressions

A.1 First Stage Results

For the elimination of selection bias, we use a two-stage propensity score weighted regression. Since we have three categories of the treatment, including the control group, we use a generalized propensity score. In the first stage, the generalized propensity score is estimated using multinomial logistic regression. The outcome variable is the categories of treatment that a household receives - households with long-term connections, households with short-term connections, and households that are yet to be connected. The multinomial logistic regression model estimates the probability that the i^{th} household is assigned the k^{th} of K (here, $K = 3$) treatments. Therefore, if there are K categories, we can run $K - 1$ parallel regressions while using one category as a reference or a pivot.

$$Pr(i,k) = \frac{exp\Bigg(\alpha + \beta_k^1 x_i^1 + \beta_k^2 x_i^2 + ... + \beta_k^m x_i^m\Bigg)}{1 + \sum_{k=1}^{K-1} exp\Bigg(\alpha + \beta_k^1 x_i^1 + \beta_k^2 x_i^2 + ... + \beta_k^m x_i^m\Bigg)} \tag{A.1}$$

where $x_i^1, x_i^2, \ldots, x_i^m$ are the independent variables which determine the propensity for household i being assigned treatment k. The household and village characteristics that we use as determinants in the estimation of propensity scores are the pre-treatment levels of the log of the total household consumption, the number of adult men and women, the presence of water sources, flush toilets, and separate kitchens in the house, whether the village has a population of less than or equal to 1000, whether the village has a population of over 5000, presence of paved roads and development groups/NGOs, distance to the nearest bank branch office/credit cooperative, and the number of primary, middle, secondary, and higher secondary schools, both government and private, and the number of primary healthcare centers, the fraction of brahmins, non-brahmin forward castes, other backward castes, scheduled castes, and scheduled tribes in the village.

Table A.1 presents the stage 1 results for the multinomial logistic regression used in the calculation of propensity scores. The table presents the results estimated on the sample of households studied in the estimation of the impact of electricity access on the logarithm of the per-capita consumption of the households. The results show that several of the variables have a significant effect in determining which households receive the treatments - short-term and long-term connections, while the control is the reference level. One may expect particularly significant coefficients for the long-term access category as these households may have already benefited from short-term gains of electrification. However, we observe that there are also statistically significant coefficients for the assignment to short-term electricity access confirming that the assignment was non-random.

We find that the logarithm of total household consumption is an important determinant of the assignment to treatment. Both the coefficients of the pre-treatment log of total household consumption, for short-term and long-term assignments, are significant at the 0.1% level, with the coefficient greater for long-term connections, perhaps due to a compounded effect of more affluent households being connected earlier, and gaining additional benefits from electricity access. Other variables that were statistically significant determinants of both short-term access and long-term access were proximity to banks and credit cooperatives, proximity to markets and shops, the fraction of households in a village that were already connected to the grid by 2005, the fraction of the population of the village that belonged to the Brahmin caste - all significant at least at the 1% level for both short-term and long-term access categories. Caste seemed to be a particularly important determinant for short-term access where the fraction of brahmins, other forward castes, scheduled tribes, and other backward castes all being important determinants of short-term connections. However, the reference caste is other castes in the village, and almost all castes seem to be a negative determinant of short-term access, and it is unclear why that may be the caste.

Note that the results presented in Table A.1 are for estimates on the sample of households for which we have data on the log of per-capita household consumption. Since other outcome variables have slightly different sample sizes, which may even be influenced by factors such as influence, such as with respect to outcome variables relevant to agricultural households, there could be slight variations in the coefficients and significance measured for each independent variable

Independent variable	Estimate (Standard Error)	
	Short-term electricity access	Long-term electricity access
Intercept	-0.439391	-10.098906***
	(1.077072)	(1.114974)
Pre-treatment level of log-total household consumption (2012 Rs.)	0.314083***	0.896803***
	(0.088296)	(0.088070)
Pre-treatment level of the number of adult men in the household	0.012534	0.080608
	(0.063999)	(0.063160)
Pre-treatment level of the number of adult women in the household	0.032900	0.096071
	(0.076181)	(0.075069)
Pre-treatment level of water source presence inside the house	0.024369	-0.223268
	(0.113408)	(0.116178)
Pre-treatment level of flush toilet presence in the house	0.013538	0.514605***
	(0.097538)	(0.087343)
Pre-treatment level of separate kitchen in the house	-0.086487	-0.146988*
	(0.062684)	(0.063177)
Pre-treatment level of metalled road presence in the village	-0.155257*	-0.034794
	(0.078904)	(0.080816)
Pre-treatment level of the number of government primary schools	-0.026735	-0.046190
	(0.030395)	(0.031557)
Pre-treatment level of the number of private primary schools	-0.080189	-0.188505***
	(0.048460)	(0.046083)
Pre-treatment level of the number of government middle schools	0.160627	-0.001686
	(0.101061)	(0.101566)
Pre-treatment level of the number of private middle schools	0.266060*	0.209316
	(0.114802)	(0.111300)
Pre-treatment level of the number of government secondary schools	-0.433936**	-0.161050
	(0.135860)	(0.132667)
Pre-treatment level of the number of private secondary schools	0.366870*	0.471076**
	(0.181839)	(0.174670)
Pre-treatment level of the number of government higher secondary schools	-0.067652	-0.176047
	(0.218990)	(0.213822)
Pre-treatment level of the number of private higher secondary schools	0.013337	0.016806
	(0.066154)	(0.062192)
Pre-treatment level of the distance to the closest bank/credit cooperative	-0.023798**	-0.032600***
	(0.008686)	(0.009479)
Pre-treatment level of the distance to the closest shop/market	0.067750***	0.063217***
	(0.014200)	(0.014312)
Pre-treatment level of the presence of NGOs in the village	0.105620	0.064212
	(0.145701)	(0.146989)
Pre-treatment level of the presence of primary healthcare centers in the village	-0.012015	0.067202
	(0.166885)	(0.163213)
Pre-treatment level of the fraction of households in the village with electricity access (%)	0.009734***	0.030321***
	(0.001642)	(0.001672)
Years since Village was connected	0.006869	0.012190**
	(0.004194)	(0.004173)
Whether the Village is small	0.158081	0.116552
	(0.126890)	(0.128018)
Whether the Village is big	-0.155035	-0.406225**
	(0.147292)	(0.149172)
Fraction of population Brahmin (%)	-0.038360***	-0.020691**
	(0.007434)	(0.007607)
Fraction of population Forward (%)	-0.024989***	-0.015925*
	(0.005644)	(0.006283)
Fraction of population OBC (%)	-0.026034***	-0.009922
	(0.005600)	(0.006210)
Fraction of population SC (%)	-0.025955	-0.011492
	(0.006091)	(0.006687)
Fraction of population ST (%)	-0.023071***	-0.003998
	(0.005784)	(0.006399)

Table A.1: Results of the first-stage regression for the estimation of coefficients of independent variables, in the households being assigned short-term and long-term treatment. The reference level of the estimates is the control group which still lacks access. The sample of households used in the estimation includes households used to estimate the impact of electricity access on consumption.

*Significant at the 5% level, **Significant at the 1% level, ***Significant at the 0.1% level across samples.

It is noteworthy that several of the determinants with extremely significant coefficients such as the pre-treatment level of consumption, the fraction of households already electrified, and the caste compositions of the villages, are not a part of the set of observable determinants chosen in Samad and Zhang (2016) and therefore, our results may vary significantly with their short-term results.

A.2 Estimating weights for the Second Stage

From the coefficients estimated, we can estimate the propensity score $p(i,k) = \hat{Pr}(i,k)$ that household i gets assigned a particular treatment. From the propensity score, we can then calculate the weights according to the formula

$$w_i = \frac{1}{\hat{Pr}(i,k)} \tag{A.2}$$

By weighting each household by the inverse of its propensity score, we scale the contribution of each household to the estimation in the second stage by the probability of being assigned to a particular treatment. In doing so, we can control the selection bias of households being assigned a particular treatment based on observable characteristics.