Electrification and Cooking Fuel Choice in Rural India

Martino Pelli* Ridhima Gupta†

September 2018

Abstract

This study investigates the causal link between electrification and the adoption of modern (and cleaner) cooking fuels, more specifically Liquefied Petroleum Gas (LPG). We exploit an instrumental variables approach that allows us to capture the part of the variation in electrification which is not related to factors that are also likely to affect a household's choice of cooking fuel. Our instrument interacts state-level supply shifts in hydro-electric power availability with the initial level of electrification of each district (11 years before the beginning of our sample). The results seem to contradict the predictions of the traditional energy ladder hypothesis. We find that electrification leads to an increase in the probability of adoption of biomass fuels and a decrease in the probability of adoption of modern cooking fuels. These results are statistically significant only for the poorest 50% of households in our sample, while they become statistically insignificant when we move to the richest 50%. The same is true for the share of expenditure in a specific fuel over total fuel expenditures. These results seem to indicate that electrification by creating an additional strain on households' finances pushes them back on the energy ladder.

Keywords: rural electrification, cooking fuel, energy ladder, fuel stacking.
JEL Classification Codes: O12, O13, Q56

*Department of Economics, University of Sherbrooke, 2500 Blvd de l'Université, Sherbrooke, Q.C., Canada and CIREQ; Martino.Pelli@USherbrooke.ca; +1 (819) 821-8000 (ext 61358).

†Department of Economic Sciences, Indian Institute of Science Education and Research, Bhopal, India; ridhima@iiserb.ac.in; + 91 755 669 2392

1 Introduction

A large literature on households' energy transition has focused on the energy ladder hypothesis. This hypothesis explains the substitution of traditional fuels with modern ones through increases in income and in the socio-economic status of the household (Bruce et al., 2000; Hosier & Dowd, 1987; Leach, 1992; van der Kroon et al., 2013). The traditional view of the energy ladder sees energy transitions as a series of disjointed steps. Yet, more recently, a growing body of literature (for instance Heltberg, 2004; Masera et al., 2000; Ruiz-Mercado, 2015), has shown that instead of switching from one fuel to the other, households simultaneously use multiple fuels. This is known as fuel stacking.

The fuel choice of households for everyday activities, such as cooking and heating, impacts on several factors influencing general wellbeing, from health to time use and exposure to financial risks (see for instance Bruce et al., 2000; Dherani et al., 2008; Khandker et al., 2013; Kishore & Spears, 2014; Peters & Sievert, 2016; Po et al., 2011; van de Walle et al., 2017). In India, electricity is rarely used for cooking. The benefits of electricity consist primarily in improved lighting and in providing power for consumer appliances. Yet, only 19% of the rural households uses Liquified Petroleum Gas (LPG) as a primary fuel for cooking, while the majority of the rural population still cooks with biofuels such as fuelwood and crop residues.

Therefore, it is not surprising that emissions from domestic fires are the largest contributor to emissions of black carbon in South Asia (Bond et al., 2013). Black carbon is the second most important greenhouse agent after carbon dioxide. About 400,000-550,000 premature deaths occur annually in India from indoor air pollution exposures to children under five and adult women (Smith, 2000). The average particle level in these households ranges from 1000-2000 $\mu g/m^3$ of PM_{10} (particulate mass of particles smaller than 10 micron diameter, Smith, 2000). This is 10-20 times higher than the national ambient air quality standard of PM_{10} set by the government of India. It is therefore crucial to understand the determinants of fuel choice and stacking and the role played by the electrification status of a household in spurring adoption of modern cooking fuels.

We investigate the existence of a causal link between electrification and the adoption of modern (and cleaner) cooking fuels, like LPG. Following the traditional energy ladder theory, and results from Chakravorty et al. (2014), Lipscomb et al. (2013), and Dinkelman (2011), our hypothesis is that the increase in income generated by electrification should push households to increase their use of LPG, a normal good, and decrease their use of fuelwood, dung and crop residues, inferior goods. In order to achieve our goal, we use a nationally representative panel of rural households constructed from three waves of the National Sample Survey (NSS), 2004, 2009 and 2011. This exercise may allow us to identify another important effect of electrification. This additional channel may play a role in the design of new policies aimed at increasing the uptake of modern cooking fuels such as LPG.

Identifying the impact of electrification on modern fuel use presents a number of empirical challenges. These challenges are common in the literature that identifies casual effects of big infrastructure projects (see for instance Allcott et al., 2016; Duflo & Pande, 2007; Röller & Waverman, 2001; Aschauer, 1989; Garcia-Mila & McGuire, 1992; Holtz-Eakin, 1993). Electrification, like other infrastructure investments, is not randomly assigned. Governments, for instance, may aim infrastructure investment to areas which are already growing faster and, therefore, electrification may always depend on a range of unobservable or omitted variables. We tackle this problem of endogeneity by using an instrumental variables approach.

The use of an instrumental variable allows us to capture the part of the variation in electrification which is not related to factors that are also likely to affect the household's choice of cooking fuel. Our instrument is constructed starting from the work of Allcott et al. (2016). In their paper, they instrument electricity shortages with state-level supply shifts in hydro-electric power availability. Following Bartik (1994) and Chakravorty et al. (2014), we then multiply state level shifts in hydro-electric power availability by the initial level of electrification of each district. In other words, we weight the supply shifts by the initial level of electrification of each district. The initial level of electrification is going to be defined as the district's electrification level 11 years before the beginning of our sample. We measure a district's initial electrification rate using the mean light intensity emanating from it, measured using satellite data. Besides controlling for a variety of household characteristics, such as size, land owned, characteristics of its head, religion and caste, we also control for fuel prices (LPG and fuelwood) and for an array of unobservable characteristics that may affect fuel choice by including district and year fixed effects. We analyse both the intensive and the extensive margin of decision regarding the choice of cooking fuel. The extensive margin focuses on the adoption decision to use a given cooking fuel. Hence our outcome of interest is a binary indicator variable indicating whether a household uses a particular cooking fuel or not. The extensive margin denotes the intensity of use of a given fuel. We measure the intensity of use by calculating the share of expenditure on a given fuel type over total monthly expenditure by a household. We separately examine the choice of Liquified Petroleum Gas (LPG) and fuelwood. Finally, we jointly investigate the extensive and the intensive margin, using a tobit specification.

The results from the main specification are rather surprising. Electrification leads to an increase in the probability of adoption of fuelwood and a decrease in the probability of adoption of LPG. This seems to go against our working hypothesis and the predictions of the energy ladder models. For this reason, we push the exercise further and discover that these effects are statistically significant only for the poorest 50% of households in our sample, while they become statistically insignificant when we move to the richest 50%. The same is true for the intensive margin. Several hypotheses may explain these results. First, it seems that for poor households, receiving an electricity connection and, therefore, facing a new monthly expenditure creates an additional strain on their budget. As a consequence, in order to enjoy electric light they reduce expenses on the more costly modern cooking fuels and revert to the use of biomass. Another hypothesis, somehow related to the first, is related to the quality of the power supply. Having a poor quality power supply, as found in Chakravorty et al. (2014), has a much smaller impact on income then a high quality power supply, therefore, not allowing households' income to increase sufficiently to push them to the next step of the energy ladder.

Precise estimations are crucially important for policymakers when taking decisions related to rural electrification. These electrification initiatives are very expensive and, therefore, have historically occurred mainly in places near the main axes of communication or the existing power grid. The cost to extend the power grid to a village which is more than 15km away from the existing infrastructure is estimated at around 150,000 dollars (Greenstone et al., 2014). That is why, in order to justify these important expenses, it is important to provide accurate figures on the benefits of a new connection. As noted above, this paper focuses on a potential additional benefit of a connection to the grid that has been previously neglected.

The remainder of the paper is organized as follows. Section 2 presents the data used. Section 3 focuses on the empirical strategy adopted and Section 4 discusses the results. Section 5 describes a few robustness test and, finally, section 6 draws some conclusions.

2 Data

Data on households expenditures are taken from surveys conducted by the National Sample Survey Organisation (NSSO). These data are representative of the rural and urban population of India and are collected every 5 years. For our analysis, we use data for the rural sample for the years 2004-2005 (61st Round), 2009-2010 (66 st Round) and 2011-2012 (68th Round). These three datasets give us a sample containing 197,231 households.[1]

Table 1 reports summary statistics for the variables used in the paper. Panel A reports statistics for the dependent variables, Panel B for the variable of interest and the instruments used and, finally, in Panel C we present all the control variables.

Only 19% of households use LPG for cooking in rural India, and 70% of rural households still rely on fuelwood.[2].

NSS asks households to recall expenditures incurred and quantities purchased of almost all items of domestic consumption over the last month. Total expenditure on any item includes money spent on purchase and value of consumption out of home production. The latter is valued at the average retail prices prevailing in the village of residence of the household.

We use this data to calculate expenditure shares of all fuel items for each household. Panel

[1] After dropping outliers.

[2] The remaining 11% is divided in the following way: 7.68% in biomass; 2.29% in other category; 0.89% use kerosene; 0.55% has no cooking arrangement; and 0.07% uses electricity

A of Table 1 shows that while households using LPG spend roughly 1% of their monthly expenditures on it, fuelwood constitute 5.1% of the expenditures of fuelwood consuming households. This big difference is probably explained by the fact that households using LPG are on average richer than households using fuelwood as primary cooking fuel.

The survey also collects information on several household-level demographic and economic characteristics such as household size, religion, caste and occupation, which we use as household level controls. We also obtain prices paid by households for the various items purchased by dividing expenditures by the number of unit bought. Since many households do not consume some of the fuel items in which we are interested – such as Liquified Petroleum Gas (LPG) – we use the average price of the item in the state of residence of the household as a proxy for the price. To account for differences in quality, we control for fuel prices prevailing in the village of residence of the household. In order to make prices and expenditure comparable across the three rounds of the NSS we use state level poverty lines as reported by the Planning Commission of India from the three time periods to deflate all 2009 and 2011 values to 2004 equivalent values.

Panel C of Table 1 shows the large difference in price between fuelwood and LPG. The rest of the control variables concern the socio-economic status of each household, measured through its land holdings, its religion, the age and the education level of its head and whether the household belong to a schedule caste/tribe or another backward caste.

72% of the households in our sample have a connection to the power grid, yet we have no information about the quality of the power supply they receive. Panel B also presents statistics for the instrument used. As exposed above, the instrument we use in this paper is composed by two separate variables that we interact. The first variable comes from Allcott et al. (2016) and is predicted hydro generation as a share of predicted electricity demand at the state level. These shocks, related to hydro generation, represent on average 13% of predicted electricity demand, yet there is a lot of variation across states. The shocks are then interacted with the initial electrification status of each district. As mentioned above, the initial electrification status of a district is measured by the level of light intensity emanating from the district 11 years before the beginning of our sample, i.e. in 1993.

The database on night light for India has been constructed by the University of Michigan in collaboration with the World Bank, using images taken by the Defense Meteorological Satellite Program (DMSP), run by the U.S. Department of Defense. This satellite program took pictures of the earth at night for 20 years, 1993 to 2013. These images have a resolution of 30 arc-seconds (i.e. roughly 1 square km at the equator). Each pixel is assigned a value between 0 and 63, where 0 indicates no light output and 63 is the highest level of light output. Based on these images, the India Lights API data is freely available for each district and month from 1993 to 2013.[3] Figure 1 shows district averages for the reference year (1993), used in the construction of our instrument. Darker shades represent a higher light output. It is easy to identify the biggest cities as the poles of highest light intensity, Delhi, Mumbai, Kolkata, Bangalore and Chennai to name a few. The interaction between this variable and the predicted shift variable takes into account the fact that districts that already had more electricity connections benefit more from the positive hydro shocks than districts with lower degrees of electrification. No district reaches the maximum light output of 63. The maximum observed over the sample is of 55.44. The average light intensity is low, at 2.56.

3 Empirical Strategy

The question of interest here is whether an electricity connection affects the probability to switch to a more modern cooking fuel, such as LPG. In order to study this question, we use the following specification

$$y_{hdst} = \alpha + \beta E_{hdst} + \gamma_1 X_{hdst} + \gamma_2 V_{dst} + \delta_d + \delta_t + \varepsilon_{hdst} \tag{1}$$

where y represents different outcome variables. First, we will investigate the impact of an electric connection on the extensive margin of adoption of LPG. In this case, the outcome

[3] http://india.nightlights.io/#/nation/2006/12

variable is binary and takes value 0 if a household does not use LPG and 1 otherwise. Second, we will investigate the intensive margin impact, i.e. does the fact of having an electricity connection increase the amount a household spends on LPG. In this case, we use the relative share of expenditures on LPG out of expenditures on cooking fuels (such as fuelwood, dung or crop residue). Subscripts h, d, s and t denote household, district, state and time, respectively. E is the electrification status of a household and X includes time-varying household specific controls, i.e. household size, religion and total land owned, information on the household head, age, sex and education and finally whether the household belongs to a scheduled caste or tribe or another backward caste. V is a matrix of village level controls, and it contains the price of fuelwood and the price of LPG. δ_d and δ_t denote district and year fixed effects, respectively. Finally, ε is the error term.

Fuel choice and the presence of an electricity connection could depend on a variety of unobservable factors. If this were the case, equation (1) cannot be interpreted casually. As said before, establishing causality in the case of important infrastructure investments presents important econometric challenges. These challenges are mainly due to the endogeneity in the placement of infrastructure. In order to tackle this endogeneity and to establish causality going from electrification to the adoption of modern cooking fuels, we use a standard instrumental variable approach.

A series of factors may hide behind the decision to electrify some areas rather than others. Governments generally reserve new investment in infrastructure to areas already experiencing more growth. Other unobservable economic trends may also influence investment decisions. For example, a richer village may have a higher probability of being connected than a poor village. The probability of connection may also depend on the proximity to a major city or the density of the population in a certain area. For all these reasons, it is difficult to isolate the effects of infrastructure investments on development measures. This issue has been widely discussed in the literature, for example, Allcott et al. (2016), Duflo & Pande (2007), Röller & Waverman (2001), Aschauer (1989), Garcia-Mila & McGuire (1992) and Holtz-Eakin (1993), just to mention a few.

Our instrument is constructed starting from the work done by Allcott et al. (2016). In their paper, they instrument electricity shortages with state-level supply shifts in hydro-electric power availability. One of the advantages of this instrument is that it also captures the quality of the supply of electricity. By capturing variations in electricity supply, this instrument actually captures shortages. If shortages are too frequent and, therefore, the quality of the power supply is low, households may assume not to have electricity even though they do have a connection. Since they cannot use for instance a fridge or other appliances that need a regular supply of electricity.

Allcott et al. (2016)'s instrument is constructed at the state level. Following Bartik (1994) and Chakravorty et al. (2014), we multiply state level shifts in hydro-electric power availability by the initial level of electrification of each district in order to obtain a district-level instrument. This is equivalent to weighting the shifts by the initial level of electrification of each district. The importance that the shifts measured by Allcott et al. (2016) will play in each district is likely contingent on its initial level of electrification. A higher initial level of electrification allows us to use more effectively the positive shifts, while no initial electrification would render them useless. Along the same narrative, negative shifts would have a bigger impact on districts with a lower initial level of electrification and a smaller one on districts with a higher initial level of electrification. The initial level of electrification is going to be defined as the district's electrification level 11 years before the beginning of our sample. We measure a district's initial electrification rate using the mean light intensity emanating from it. The mean light intensity is measured as the average of the light intensity of each pixel composing the district. Light intensity varies between 0 and 63.

This instrument constructed by Allcott et al. (2016) consists of state level supply shifts from hydroelectric generation.[4] Given the low marginal costs of hydroelectric power plants, their annual output depends mainly on water availability. The instrument constructed in

[4] Allcott et al. (2016) is interested in quantifying the impact of power outages on industrial production in Indian states. Yet, power outages and industrial production may be correlated. For instance, if a state is growing faster and, consequently, its energy demand is higher, it may experience more power outages. In order to eliminate this endogeneity problem, Allcott et al. (2016) looks for an instrumental variable (the use of an instrumental variable also solves for the possible measurement error coming from the way in which shortages are measured). For this reason, the authors need an instrument affecting shortages but not affecting industrial production, or better, that affects industrial production only through its impact on power shortages.

Allcott et al. (2016) is predicted hydro generation as a share of predicted electricity demand. Allcott et al. (2016) divides predicted hydro generation by predicted demand in order to obtain the relative share of hydro generation across the different states.

Actual demand could be affected by shortages and, for this reason, the instrument is not based on actual demand but on predicted demand. Demand is predicted using the average share demanded by a given state between 1992 and 2010 multiplied by electricity demand in all other states.[5] Computing predicted electricity demand in this way minimizes fluctuations due to important shortages.

Allcott et al. (2016) also needs to use predicted hydro generation, because – since water can be kept in reservoirs – in years of low industrial demand less hydro energy will be produced and, therefore, industrial production and hydro generation may be correlated. For this reason, they use a prediction instead of actual generation. Predicted generation capacity is computed in a similar way as predicted demand, and is based on reservoirs inflow and run-of-river plants. The first element needed is the "state predicted annual hydrogeneration capacity" (C), which is computed in the same way as predicted demand. This term is then multiplied by a "state average capacity factor". The capacity factor depends on two main elements. First, on the "share of output of a given reservoir which is contractually allocated to the state" and, second, on the "demeaned inflow predicted capacity factor for a given reservoir in a given year". This last element is computed individually for each reservoir-year couple. Allcott et al. (2016) first regresses generation on inflow (for each reservoir) in order to produce inflow-predicted generation, which is then divided by the generation capacity of the reservoir (capacity x 8,760 hours in a year).

[5] Denoting by s and t the state and year at hand, respectively and r other states and y years, predicted electricity demand $\tilde{D}_{st}$ is defined as

$$\tilde{D}_{st} = \sum_{r \neq s} D_{rt} \cdot \sum_{y=1992}^{2010} \frac{D_{sy}}{\sum_{r \neq s} D_{ry}} \tag{2}$$

The second multiplicative term computes the electricity demand/consumption in a given state s as a share of total electricity demand/consumption in all other states and takes the average of this share over the period 1992 to 2010. Once obtained this average share, it is multiplied by total consumption in all other states in the year of interest, obtaining in this way a prediction of consumption for that particular year.

The first stage of our specification takes the following form

$$E_{hdst} = \alpha + \beta_1(H_{st} * L_{sd0}) + \beta_2 H_{st} + \gamma X_{hdst} + \delta_d + \delta_t + u_{hdst} \tag{3}$$

where L represents a district's light intensity and H is Allcott et al. (2016)'s state-level instrument. δ_d and δ_t denote district and year fixed effects, respectively. $H_{st} * L_{sd0}$ is our instrument, the interaction between the Allcott et al. (2016) instrument and a district's initial light intensity.[6] Finally X is a matrix of time-varying household specific controls and u is the error term.

Table 2 reports results for first stage estimations. The table is organized as follows, columns (1) and (2) contain first stage estimations for the whole sample, first with only the instrument and the relevant fixed effects and then with the full set of household controls. Column (3) shows the first stage estimation for individuals whose household's per adult equivalent monthly expenditure is below the median level (of roughly 600 Rs), while column (4) shows it for households whose income is above the median level. Standard errors are clustered at the district level. Let us focus on the effect of the interaction, i.e. our instrument. Its impact is positive and statistically significant at the 1% level across all specifications. This means that a hydro shock combined with a higher level of initial electrification in the district increases the probability of electrification by roughly 11 percentage points. The direction of this coefficient is the one expected. The negative value on *Lag Hydro* cannot be interpreted on its own, it has to be interpreted together with the coefficient on the interaction term. In the rare cases where the initial level of electrification is zero, we may observe a negative impact of hydro on the probability of receiving an electricity connection. Yet, as the initial level of electrification increases this negative effect decreases and becomes positive. When running the first stage

[6] The condition which needs to be satisfied in order for our identification to be consistent is that $E[L_{ds0} * H_{st} * \varepsilon_{idst}] = 0$. We can easily verify this by taking the limit over districts and states $\lim_{D,S\to\infty} \frac{1}{D*S} \sum_d \sum_s (L_{ds0} * H_{st} * \varepsilon_{idst}) = 0$. Since H_{st} does not depend on the district, we can extract it from the sum and re-write it as $\lim_{D,S\to\infty} \frac{1}{D*S} \sum_s H_{st} \sum_d (L_{ds0} * \varepsilon_{idst}) = 0$. In order for this to be verified we only need to argue that $\lim_{D\to\infty} \frac{1}{D} \sum_d (L_{ds0} * \varepsilon_{idst}) = 0$, or equivalently that $E[L_{ds0} * \varepsilon_{idst}] = 0$. This means that shocks in cooking fuels decisions in 2004, 2009 and 2011 need not to be correlated with the district light intensity in 1992. While the trend in cooking fuel choices may be related to the initial level of electrification of a district, shocks to the path are highly unlikely to be related to initial electrification. For this reason, we are confident in claiming that this condition is satisfied.

on the whole sample, we obtain an F statistic of 13.2, above the critical level of 10. When we split the sample the value of the F statistic decreases to 9.83 and 3.89, respectively. This means that our instrument is weak when applied to the two subsamples. In order to insure the robustness of our inference, we compute the Anderson-Rubin F statistics for weak instruments, which in both cases rejects the hypothesis that the coefficients of interest may be equal to zero. Therefore, although the instrument becomes weaker, the inference stays robust.

4 Results

First, we present an extensive margin analysis, i.e. we focus on the adoption of LPG as a cooking fuel versus fuelwood. Second, we look at the intensive margin. The intensive margin analysis measures the impact of electrification on spending in LPG versus fuelwood. Eventually, we look at a tobit estimation, in order to account for the intensive and extensive margin at the same time. This allows us to correct for the large amount of zeros in the dataset.

Extensive margin

Table 3 reports results for the extensive margin estimations, i.e. adoption. For each estimation we report OLS and IV coefficients of a linear probability model. Columns (1) and (2) show results for LPG, while (3) and (4) for fuelwood. Standard errors are clustered at the district level in all specifications.

Let us focus on the IV results for this linear probability model. The results on the variable of interest, *Electricity*, are all statistically significant at least at the 5% level. These coefficients are negative for LPG and positive for fuelwood. This implies that when a household gets connected to the power grid, the probability that it will use LPG for cooking decreases by 63.6 percentage points, while the probability that it will use fuelwood increases by 70.7 percentage points. These results go against conventional theories. One would expect that electricity constitutes a push towards modernization and, therefore, towards the adoption of more modern and less polluting cooking fuels. The signs of the coefficients on the control variables, instead, are all as expected. The price of fuelwood has a positive impact on the adoption of modern fuels and a negative one on the adoption of biomass, while the price of LPG has the opposite effect. Bigger households tend to be poorer and, therefore, use more biomass and less modern (and usually more expensive) fuels. A male, more educated and older head of household is usually associated with a higher economic status, and this is exactly what we find in our data. This typology of household will use modern fuels more often than biomass. Belonging to a lower caste also means a higher probability of cooking using biomass.

Why are these results counterintuitive? Why does a connection to the grid decrease the probability of adoption of modern cooking fuels? The story lying behind these results is probably linked to a financial constraint. Cooking with LPG is more expensive than with biomass. Biomass, consisting mainly of fuelwood, dung and crop residue, can easily be collected, while LPG has to be bought. Households connected to the grid have to pay for the electricity they consume, and this extra expenditure leaves a smaller disposable budget for cooking fuels. In order to verify whether this is what is happening, we split our sample in two. The split is based on the median monthly per capita expenditures. In the median household, the monthly per capita expenditure is of roughly 600 Rs (the exact value is 603.5 Rs). Therefore, in the first sub-sample we keep households characterized by a monthly per capita expenditure lower than 600 Rs, while in the second sub-sample we keep richer households, with expenditures per capita higher than 600 Rs.

Tables 4 presents results for this test, which seems to confirm our hypothesis. Columns (1) through (4) present results for LPG and fuelwood adoption for households whose monthly per capita expenditures are below 600 RS., while columns (5) through (8) present them for households situated above the 600 Rs. threshold. The coefficients obtained are slightly bigger in magnitude and negative for the poorer households, while they are not statistically significant for the richer households. It seems that a power connection allows poor households to have better lighting, yet puts an additional constraint on their tight energy budget, pushing them to revert to fuelwood for cooking. Instead richer households do not suffer from the additional financial strain. For them, obtaining a power connection does not have an impact on the choice of cooking fuel.

Intensive margin

Table 5 presents the intensive margin results, i.e. how do spending in LPG and fuelwood evolve as a consequence of electrification. Columns (1) and (2) show results for LPG, while columns (3) and (4) report fuelwood results. The picture presented here is quite similar to the one on the extensive margin.

A new electrical connection not only decreases the probability of adoption of LPG as a cooking fuel, but seems also to decrease the share of spending households dedicate to it. When focusing on fuelwood we observe the opposite pattern, exactly as for the extensive margin, a new electrical connection increases spending in fuelwood. The fact that the results follow the same pattern for the extensive and the intensive margin of the analysis reinforces our believe in an explanation related to the budget constraint. In order to investigate this claim, we proceed as before and split the sample according to the median monthly per capita expenditure.

Results for this can be found in table 6. Columns (1) through (4) present results for households with monthly per capita spending above the 600 Rs. threshold, while columns (5) through (8) present results for households with spending below it. As in the extensive margin analysis, results are statistically significant only for the poorer 50% of households. These households, when faced with the expenditures related to electrical power have to cut back on expenditures in other expensive fuels and, therefore, we observe a decrease in the share dedicated to LPG and an increase in the share dedicated to fuelwood, the cheaper alternative.

Tobit

Table 7 shows results for a tobit estimation over the whole sample.

The interest of a tobit model lies in accounting for data censoring, in our case for left censoring since our database contains many zero values as the majority of the households do not spend on LPG. In our case, the tobit model weakens the results obtained for the intensive margin. The IV result for LPG is not statistically significant anymore, while the one for fuelwood becomes statistically significant only at the 10% level. As before, when moving on and splitting the sample according to income – in Table 8 – we observe that this effect on fuelwood becomes stronger, i.e. larger in magnitude and statistically significant at the 5 % level. Yet, the effect on LPG remains statistically insignificant. As before, richer households do not seem to react to electrification.

5 Robustness

In order to further verify our claim, we perform several robustness checks. First, we split the sample into quartiles instead of just focusing on the median income. Second, we measure the economic status of a household according to state-specific and country-specific poverty lines. Finally, we verify our claim that electricity puts further stress on the poorest households' budget by looking at the impact of electrification on their energy budget.

First, we analyze the evolution of the coefficient on an electricity connection across quartiles of monthly per capita spending. This evolution is shown in Figure 2 and Figure 3. When splitting the sample four-ways, the coefficients become less precise. Yet, we observe the same patterns found before. Let us start with Figure 2. This figure compares the coefficients obtained when regressing adoption of LPG as a cooking fuel on access to electricity by quartile of monthly per capita expenditures. These are all instrumental variable regressions with errors clustered at the district level. The figure shows the same result we obtained earlier, in response to a connection to the power network poorer households abandon modern cooking fuels in favour of solid ones, or simply have a lower probability of adoption, while richer households do not seem to be affected by it. Figure 3 instead focuses on the adoption of fuelwood as a cooking fuel. Again, we observe the same story as before. Results for the bottom of the income distribution are statistically stronger, and point towards a higher probability of adoption of fuelwood following a connection to the grid.

Second, instead of splitting the sample across the median or quartiles, we split the sample according to state-specific poverty lines calculated by Rangarajan et al. (2014).[7] The poverty lines are defined as the per capita monthly household expenditure that needs to be incurred to achieve a minimum standard of living. Any household that fails to meet this level of consumption expenditure can be treated as a poor household. Using this definition, we classify each household in the sample as poor and non-poor. The results are shown in Table 9. The first thing we notice when splitting the sample along the poverty line is that we end up with two unbalanced samples. Over 150,000 households are above the poverty line, while only 44,000 are below. This fact is going to have an impact on our results. While we find a statistically significant, positive impact on the probability of adoption of fuelwood as a cooking fuel only for households below the poverty line, the impact on the adoption of LPG is negative and statistically significant for both groups. These results seem at odds with those presented above, yet once we take into account how stringent is the definition of the poverty line, we should not be surprised by them.

Finally, if a new connection to the power grid poses an extra constraint on people's budget, we should be able to pick this up. In Table 10, we regress the share of fuel expenditures out of a household's total expenditures on access to a connection to the power grid. Columns (1) and (2) show the OLS and IV results for the whole sample and seem to indicate, as expected, an increase in the share of income that is dedicated to energy. When we split the sample between households below and above the median monthly per capita expenditure we find an additional argument strengthening our claim. Poorer households already spend a larger share of their income on energy (on average 17% of their expenditures are dedicated to energy), and the new power connection increases this share in a significant way, by 16 percentage points on average, and this result is statistically significant at the 5% level. If instead we move to the richer half of the sample, that on average dedicates 14% of its expenditures to energy, the coefficient becomes statistically insignificant. Therefore, it seems that the new connection

[7] If instead of using a state-wise poverty we use the national poverty line the results are unchanged. These results are available upon request.

does not put any additional financial stress on richer households.

6 Conclusions

In this paper, we investigate the impact of the fact that a household in rural India is connected to the power grid on its choice of cooking fuel. The ideas, coming from the energy ladder literature, is that electrification, via an increase in income, should push households to the next step on the energy ladder. According to the fuel stacking literature, this may not be a clear jump, but we should observe the appearance of modern fuels together with the persistence of biomass.

Our results are surprising and, for this reason, important from a policy perspective. We observe that, while electrification has basically no impact on the choice of a cooking fuel for richer households, its impact is the opposite of what one would want for poorer households. Poorer households, when faced with the possibility of having access to electric light, they will take it. Yet, probably because of their tight budget constraint, in order to be able to afford electricity, they cut back on other energy related expenses, reverting from modern cooking fuels to biomass.

References

Allcott, H., Collard-Wexler, A., & O'Connell, S. D. (2016). How do electricity shortages affect industry? evidence from india. *The American Economic Review*, 106(3), 587–624.

Aschauer, D. A. (1989). Is public expenditure productive? *Journal of monetary economics*, 23(2), 177–200.

Bartik, T. J. (1994). The effects of metropolitan job growth on the size distribution of family income. *Journal of Regional Science*, 34(4), 483–501.

Bond, T. C., Doherty, S. J., Fahey, D., Forster, P., Berntsen, T., DeAngelo, B., Flanner, M., Ghan, S., Kärcher, B., Koch, D., et al. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11), 5380–5552.

Bruce, N., Perez-Padilla, R., & Albalak, R. (2000). Indoor Air Pollution in Developing Countries: A Major Environmental and Public Health Challenge. *Bulletin of the World Health Organization*, 78, 1078–1092.

Chakravorty, U., Pelli, M., & Marchand, B. U. (2014). Does the quality of electricity matter? evidence from rural india. *Journal of Economic Behavior & Organization*, 107, 228–247.

Dherani, M., Pope, D., Mascarenhas, M., Smith, K., Weber, M., & Bruce, N. (2008). Indoor Air Pollution from Unprocessed Solid Fuel Use and Pneumonia Risk in Children Aged under Five Years: A Systematic Review and Meta-Analysis. *Bulletin of the World Health Organization*, 86, 390–398C.

Dinkelman, T. (2011). The Effects of Rural Electrification on Employment: New Evidence from South Africa. *American Economic Review*, 101(7), 3078–3108.

Duflo, E. & Pande, R. (2007). Dams. *The Quarterly Journal of Economics*, 122(2), 601–646.

Garcia-Mila, T. & McGuire, T. J. (1992). The contribution of publicly provided inputs to states' economies. *Regional science and urban economics*, 22(2), 229–241.

Greenstone, M. et al. (2014). Energy, growth and development. *International Growth Center Evidence Paper*.

Heltberg, R. (2004). Fuel switching: evidence from eight developing countries. *Energy economics*, 26(5), 869–887.

Holtz-Eakin, D. (1993). State-specific estimates of state and local government capital. *Regional science and urban economics*, 23(2), 185–209.

Hosier, R. & Dowd, J. (1987). Household Fuel Choice in Zimbabwe. *Resources and Energy*, 9, 347–361.

Khandker, S., Barnes, D., & Samad, H. (2013). Welfare Impacts of Rural Electrification: A Panel Data Analysis from Vietnam. *Economic Development and Cultural Change*, 61(3), 659–692.

Kishore, A. & Spears, D. (2014). Having a Son Promotes Clean Cooking Fuel Use in Urban India: Women's Status and Son Preference. *Economic Development and Cultural Change*, 62(4), 673–699.

Leach, G. (1992). The Energy Transition. *Energy Policy*, 20(2), 116–123.

Lipscomb, M., Mobarak, A., & Barham, T. (2013). Development Effects of Electrification: Evidence from the Geologic Placement of Hydropower Plants in Brazil. *American Economic Journal: Applied Economics*, 5(2), 200–231.

Masera, O., Saatkamp, B., & Kammen, D. (2000). From Linear Fuel Switching to Multiple Cooking Strategies: A Critique and Alternative to the Energy Ladder Model. *World Development*, 28(12), 2083 – 2103.

Peters, J. & Sievert, M. (2016). Impacts of rural electrification revisited – the African context. *Journal of Development Effectiveness*, 8(3), 327–345.

Po, J., Fitzgerald, J., & Carlsten, C. (2011). Respiratory Disease Associated with Solid

Biomass Fuel Exposure in Rural Women and Children: Systematic Review and Meta-Analysis. *Thorax*, 66, 232–239.

Rangarajan, C., Mahendra Dev, S., Sundaram, K., Vyas, M., & Datta, K. (2014). Report of the expert group to review the methodology for measurement of poverty. *New Delhi*.

Röller, L.-H. & Waverman, L. (2001). Telecommunications infrastructure and economic development: A simultaneous approach. *American economic review*, (pp. 909–923).

Ruiz-Mercado, I., M. O. (2015). Patterns of Stove Use in the Context of Fuel-Device Stacking: Rationale and Implications. *EcoHealth*, 12, 45–56.

Smith, K. R. (2000). National burden of disease in india from indoor air pollution. *Proceedings of the National Academy of Sciences*, 97(24), 13286–13293.

van de Walle, D., Ravaillon, M., Mendiratta, V., & Koolwal, G. (2017). Long Term Gains from Electrification in Rural India. *The World Bank Economic Review*, 31(2), 385–411.

van der Kroon, B., Brouwer, R., & van Beukering, P. (2013). The Energy Ladder: Theoretical Myth or Empirical Truth? Results from a Meta-Analysis. *Renewable and Sustainable Energy Reviews*, 20, 504–513.

7 Tables

Table 1: Summary Statistics of Variables used in the QAIDS Model

Variable	Mean	SE	Min	Max
Panel A				
Share of LPG	0.19	0.39	0	1
Share of fuelwood	0.70	0.46	0	1
Percent share of LPG expenditure	1.02	2.16	0	27.82
Percent share of Fuelwood expenditure	5.09	4.62	0	65.75
Panel B				
Share of electricity	0.72	0.45	0	1
Nightlights 1993	2.56	3.24	0	55.44
Lag Hydro Instrument	0.13	0.19	0	0.93
Interaction Hydro Instrument	0.30	0.59	0	4.91
Panel C				
Price Fuelwood (Rs./Kg)	1.23	2.74	0.01	625
Price LPG (Rs./ Kg)	15.65	11.99	5	654.48
HH Size	4.93	2.42	1	43
Total land owned (Hectares)	1.10	30.39	0	9120.27
Education head of HH (Years)	4.77	3.28	1	13
Sex head of HH	0.89	0.31	0	1
Age head of HH	46.54	13.33	0	108
Hindu	0.77	0.42	0	1
Scheduled tribe	0.16	0.37	0	1
Scheduled caste	0.18	0.38	0	1
Other backward caste	0.39	0.49	0	1

Notes: Percent shares of expenditures are computed over the last 30 days.

Table 2: First stage results

	Dep. variable: electricity			
	Full sample		**Below**	**Above**
	(1)	(2)	(3)	(4)
Lag Hydro	-0.453^{***}	-0.417^{***}	-0.480^{***}	-0.150^{*}
	(0.083)	(0.085)	(0.111)	(0.078)
*Lag Hydro*Initial elec*	0.132^{***}	0.113^{***}	0.120^{***}	0.060^{***}
	(0.032)	(0.031)	(0.038)	(0.022)
Price Fuelwood		0.001	0.0005	0.00003
		(0.001)	(0.002)	(0.001)
Price LPG		-0.0004	0.0003	-0.001^{**}
		(0.001)	(0.001)	(0.001)
HH size		0.010^{***}	0.017^{***}	0.013^{***}
		(0.001)	(0.001)	(0.001)
Total land owned		0.0002	0.00002	0.0005
		(0.000)	(0.000)	(0.001)
Education head of HH		0.024^{***}	0.024^{***}	0.015^{***}
		(0.001)	(0.001)	(0.001)
Sex of head of HH		-0.017^{***}	-0.009	-0.020^{***}
		(0.004)	(0.006)	(0.004)
Age head of HH		0.002^{***}	0.002^{***}	0.001^{***}
		(0.000)	(0.000)	(0.000)
Hindu		0.008	0.006	0.002
		(0.006)	(0.008)	(0.005)
Scheduled tribe		-0.124^{***}	-0.122^{***}	-0.084^{***}
		(0.011)	(0.012)	(0.011)
Scheduled caste		-0.072^{***}	-0.066^{***}	-0.049^{***}
		(0.005)	(0.007)	(0.005)
Other backward caste		-0.028^{***}	-0.030^{***}	-0.019^{***}
		(0.004)	(0.006)	(0.004)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Observations	196,181	195,272	96,741	98,531
F-stat	16.68	13.20	9.83	3.89
AR F-stat	21.74	34.34	36.92	13.16

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Baseline results – extensive margin

	Dep. variable:			
	LPG		**Fuelwood**	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Electricity	0.082^{***}	-0.636^{***}	-0.074^{***}	0.707^{**}
	(0.005)	(0.154)	(0.007)	(0.296)
Price Fuelwood	0.003^{**}	0.004^{**}	-0.005^{***}	-0.006^{***}
	(0.001)	(0.001)	(0.002)	(0.002)
Price LPG	-0.001	-0.001^{**}	0.001	0.002^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
HH size	-0.003^{***}	0.005^{***}	0.005^{***}	-0.003
	(0.001)	(0.002)	(0.001)	(0.003)
Total land owned	0.0001	0.0002	-0.0000	-0.0002
	(0.000)	(0.000)	(0.000)	(0.000)
Education head of HH	0.038^{***}	0.055^{***}	-0.035^{***}	-0.054^{***}
	(0.001)	(0.004)	(0.001)	(0.007)
Sex of head of HH	-0.048^{***}	-0.060^{***}	0.046^{***}	0.059^{***}
	(0.004)	(0.006)	(0.004)	(0.008)
Age head of HH	0.003^{***}	0.004^{***}	-0.002^{***}	-0.003^{***}
	(0.000)	(0.000)	(0.000)	(0.001)
Hindu	0.003	0.009	0.001	-0.006
	(0.006)	(0.009)	(0.007)	(0.010)
Scheduled tribe	-0.086^{***}	-0.175^{***}	0.105^{***}	0.202^{***}
	(0.008)	(0.022)	(0.009)	(0.038)
Scheduled caste	-0.084^{***}	-0.136^{***}	0.080^{***}	0.136^{***}
	(0.005)	(0.012)	(0.006)	(0.022)
Other backward caste	-0.055^{***}	-0.076^{***}	0.045^{***}	0.067^{***}
	(0.004)	(0.007)	(0.005)	(0.010)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Observations	195,272	195,272	195,272	195,272
F-stat		13.20		13.20
AR F-stat		34.34		34.34

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Baseline results – extensive margin below/above 600

	Dep. variable:							
	Below 600				Above 600			
	LPG		Fuelwood		LPG		Fuelwood	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Electricity	0.035***	−0.521***	−0.030***	1.070***	0.142***	−0.513	−0.126***	0.167
	(0.003)	(0.141)	(0.006)	(0.336)	(0.007)	(0.372)	(0.009)	(0.574)
Price Fuelwood	0.002***	0.002	−0.008***	−0.008**	0.002	0.002	−0.002	−0.002
	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)	(0.001)	(0.002)	(0.002)
Price LPG	−0.0004	−0.0003	0.001	0.0004	−0.002**	−0.003***	0.003**	0.003**
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
HH size	0.003***	0.013***	−0.003***	−0.022***	0.003***	0.011**	0.002*	−0.002
	(0.000)	(0.003)	(0.001)	(0.006)	(0.001)	(0.005)	(0.001)	(0.007)
Total land owned	−0.000	0.000	0.0001**	0.000	−0.002***	−0.002**	0.002**	0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Education head of HH	0.012***	0.026***	−0.012***	−0.038***	0.046***	0.056***	−0.042***	−0.046***
	(0.001)	(0.004)	(0.001)	(0.008)	(0.001)	(0.006)	(0.001)	(0.009)
Sex of head of HH	−0.019***	−0.024***	0.027***	0.037***	−0.072***	−0.085***	0.060***	0.066***
	(0.003)	(0.005)	(0.004)	(0.008)	(0.005)	(0.010)	(0.005)	(0.013)
Age head of HH	0.001***	0.002***	−0.0003***	−0.002***	0.003***	0.004***	−0.002***	−0.003***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Hindu	0.007**	0.010*	−0.002	−0.009	−0.007	−0.006	0.008	0.007
	(0.004)	(0.006)	(0.006)	(0.011)	(0.008)	(0.009)	(0.008)	(0.009)
Scheduled tribe	−0.036***	−0.104***	0.053***	0.188***	−0.118***	−0.173***	0.133***	0.158***
	(0.005)	(0.019)	(0.006)	(0.044)	(0.012)	(0.034)	(0.013)	(0.050)
Scheduled caste	−0.033***	−0.070***	0.030***	0.103***	−0.095***	−0.128***	0.093***	0.108***
	(0.004)	(0.011)	(0.006)	(0.024)	(0.007)	(0.019)	(0.007)	(0.029)
Other backward caste	−0.020***	−0.037***	0.010*	0.044***	−0.060***	−0.073***	0.054***	0.060***
	(0.004)	(0.007)	(0.006)	(0.013)	(0.005)	(0.009)	(0.006)	(0.013)
Year F.E.	yes	yes	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes	yes	yes
Observations	96,741	96,741	96,741	96,741	98,531	98,531	98,531	98,531
F-stat		9.83		9.83		3.89		3.89
AR F-stat		36.92		36.92		13.16		13.16

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 5: Baseline results – intensive margin

	Dep. variable:			
	LPG		**Fuelwood**	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Electricity	0.611***	−1.699**	−1.487***	5.569*
	(0.029)	(0.792)	(0.067)	(3.284)
Price Fuelwood	0.017**	0.019**	0.126***	0.118**
	(0.007)	(0.008)	(0.047)	(0.048)
Price LPG	−0.008**	−0.009***	−0.015*	−0.010
	(0.003)	(0.003)	(0.008)	(0.009)
HH size	−0.041***	−0.017**	−0.188***	−0.262***
	(0.003)	(0.009)	(0.008)	(0.037)
Total land owned	−0.001	−0.0003	−0.005	−0.007
	(0.000)	(0.000)	(0.004)	(0.006)
Education head of HH	0.168***	0.224***	−0.326***	−0.497***
	(0.004)	(0.020)	(0.009)	(0.083)
Sex of head of HH	−0.282***	−0.322***	−0.196***	−0.074
	(0.023)	(0.031)	(0.052)	(0.087)
Age head of HH	0.013***	0.017***	−0.012***	−0.026***
	(0.001)	(0.002)	(0.001)	(0.007)
Hindu	0.095***	0.113***	−0.031	−0.086
	(0.031)	(0.039)	(0.061)	(0.088)
Scheduled tribe	−0.473***	−0.757***	1.202***	2.072***
	(0.045)	(0.107)	(0.109)	(0.442)
Scheduled caste	−0.446***	−0.612***	0.939***	1.445***
	(0.025)	(0.059)	(0.059)	(0.241)
Other backward caste	−0.271***	−0.334***	0.374***	0.569***
	(0.022)	(0.031)	(0.042)	(0.101)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Observations	196,344	196,344	196,283	196,283
F-stat		13.41		13.37

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Baseline results – intensive margin below/above 600

	Dep. variable:							
	Below 600				Above 600			
	LPG		Fuelwood		LPG		Fuelwood	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Electricity	0.324***	−2.923***	−1.005***	10.272**	0.906***	−0.171	−1.287***	2.712
	(0.026)	(0.843)	(0.066)	(4.537)	(0.038)	(2.061)	(0.082)	(4.810)
Price Fuelwood	0.017***	0.018*	0.140***	0.137**	0.009	0.009	0.135**	0.134**
	(0.006)	(0.010)	(0.046)	(0.057)	(0.007)	(0.006)	(0.054)	(0.055)
Price LPG	−0.005*	−0.004	−0.021**	−0.024*	−0.014**	−0.015**	0.007	0.013
	(0.003)	(0.003)	(0.010)	(0.014)	(0.005)	(0.006)	(0.008)	(0.010)
HH size	0.008***	0.064***	−0.360***	−0.555***	−0.042***	−0.029	−0.217***	−0.267***
	(0.003)	(0.015)	(0.012)	(0.082)	(0.004)	(0.026)	(0.009)	(0.062)
Total land owned	−0.0001	−0.0001	−0.002	−0.002	−0.026***	−0.025***	−0.002	−0.003
	(0.000)	(0.000)	(0.001)	(0.003)	(0.004)	(0.004)	(0.008)	(0.009)
Education head of HH	0.091***	0.170***	−0.152***	−0.424***	0.185***	0.201***	−0.281***	−0.342***
	(0.004)	(0.022)	(0.008)	(0.111)	(0.004)	(0.032)	(0.009)	(0.076)
Sex of head of HH	−0.170***	−0.201***	−0.489***	−0.388***	−0.369***	−0.391***	0.002	0.084
	(0.023)	(0.032)	(0.073)	(0.108)	(0.029)	(0.053)	(0.049)	(0.116)
Age head of HH	0.006***	0.011***	0.005***	−0.012	0.015***	0.016***	−0.007***	−0.011**
	(0.000)	(0.002)	(0.002)	(0.008)	(0.001)	(0.002)	(0.001)	(0.005)
Hindu	0.080***	0.098**	−0.079	−0.145	0.079*	0.081*	0.017	0.009
	(0.027)	(0.040)	(0.064)	(0.117)	(0.041)	(0.043)	(0.061)	(0.069)
Scheduled tribe	−0.292***	−0.690***	0.785***	2.172***	−0.583***	−0.672***	1.005***	1.337***
	(0.037)	(0.114)	(0.110)	(0.618)	(0.067)	(0.182)	(0.112)	(0.408)
Scheduled caste	−0.279***	−0.492***	0.591***	1.337***	−0.468***	−0.520***	0.626***	0.823***
	(0.028)	(0.064)	(0.075)	(0.321)	(0.031)	(0.103)	(0.050)	(0.237)
Other backward caste	−0.161***	−0.259***	0.178***	0.523***	−0.276***	−0.296***	0.331***	0.406***
	(0.027)	(0.041)	(0.062)	(0.168)	(0.025)	(0.046)	(0.040)	(0.095)
Year F.E.	yes	yes	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes	yes	yes
Observations	96,901	96,901	96,772	96,772	96,443	96,443	99,511	99,511
F-stat		9.83		9.83		3.89		3.89
AR F-stat		36.92		36.92		13.16		13.16

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 7: Baseline results – tobit

	Dep. variable:			
	LPG		**Fuelwood**	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Electricity	1.159***	0.544	−1.041***	4.926*
	(0.033)	(1.238)	(0.048)	(2.608)
Price Fuelwood	0.016***	0.016***	0.075***	0.061**
	(0.006)	(0.006)	(0.029)	(0.027)
Price LPG	−0.005*	−0.006**	−0.008	−0.003
	(0.003)	(0.003)	(0.006)	(0.007)
HH size	−0.005**	0.001	−0.103***	−0.155***
	(0.002)	(0.013)	(0.005)	(0.029)
Total land owned	0.001*	0.001	−0.003	−0.004
	(0.001)	(0.001)	(0.002)	(0.004)
Education head of HH	0.167***	0.181***	−0.258***	−0.377***
	(0.005)	(0.028)	(0.007)	(0.066)
Sex of head of HH	−0.307***	−0.316***	−0.070*	0.038
	(0.022)	(0.027)	(0.038)	(0.065)
Age head of HH	0.015***	0.016***	−0.009***	−0.020***
	(0.001)	(0.002)	(0.001)	(0.005)
Hindu	0.038	0.042	−0.042	−0.084
	(0.028)	(0.030)	(0.048)	(0.067)
Scheduled tribe	−0.615***	−0.685***	0.921***	1.562***
	(0.051)	(0.151)	(0.077)	(0.349)
Scheduled caste	−0.504***	−0.544***	0.739***	1.094***
	(0.025)	(0.081)	(0.044)	(0.193)
Other backward caste	−0.220***	−0.234***	0.328***	0.461***
	(0.018)	(0.034)	(0.033)	(0.080)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Censored Observations	144,936	144,936	28732	28732
Uncensored Observations	49940	49940	166,078	166,078
P-value of Wald Test of Exogeneity		0.63		0.02

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Baseline results – tobit below/above 600

	Dep. variable:							
	Below 600				**Above 600**			
	LPG		**Fuelwood**		**LPG**		**Fuelwood**	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Electricity	1.009^{***}	-5.693	-0.800^{***}	7.349^{**}	1.242^{***}	1.794	-0.848^{***}	3.596
	(0.044)	(3.988)	(0.053)	(3.147)	(0.042)	(2.936)	(0.051)	(5.734)
Price Fuelwood	0.023^{***}	0.037^{*}	0.096^{***}	0.078^{**}	0.007	0.007	0.073^{**}	0.069^{**}
	(0.008)	(0.021)	(0.033)	(0.036)	(0.006)	(0.006)	(0.029)	(0.030)
Price LPG	-0.005	-0.006	-0.014^{*}	-0.014	-0.009^{**}	-0.009	0.006	0.012
	(0.004)	(0.007)	(0.008)	(0.009)	(0.004)	(0.006)	(0.006)	(0.010)
HH size	0.056^{***}	0.214^{***}	-0.270^{***}	-0.363^{***}	0.013^{***}	0.006	-0.103^{***}	-0.154^{**}
	(0.004)	(0.079)	(0.009)	(0.056)	(0.003)	(0.035)	(0.006)	(0.075)
Total land owned	0.000	0.001	-0.001	-0.001	-0.006^{**}	-0.006^{*}	0.011^{**}	0.008
	(0.000)	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.005)	(0.007)
Education head of HH	0.155^{***}	0.422^{***}	-0.128^{***}	-0.299^{***}	0.169^{***}	0.162^{***}	-0.212^{***}	-0.271^{***}
	(0.005)	(0.123)	(0.007)	(0.077)	(0.005)	(0.035)	(0.006)	(0.095)
Sex of head of HH	-0.341^{***}	-0.610^{***}	-0.338^{***}	-0.209^{***}	-0.338^{***}	-0.330^{***}	0.072^{**}	0.159
	(0.037)	(0.117)	(0.057)	(0.074)	(0.024)	(0.045)	(0.034)	(0.129)
Age head of HH	0.012^{***}	0.031^{***}	0.004^{***}	-0.009^{*}	0.014^{***}	0.014^{***}	-0.006^{***}	-0.010
	(0.001)	(0.008)	(0.001)	(0.005)	(0.001)	(0.002)	(0.001)	(0.006)
Hindu	0.064	0.142	-0.061	-0.098	0.033	0.032	-0.012	-0.020
	(0.041)	(0.102)	(0.054)	(0.082)	(0.033)	(0.033)	(0.045)	(0.054)
Scheduled tribe	-0.673^{***}	-1.959^{***}	0.647^{***}	1.522^{***}	-0.601^{***}	-0.561^{**}	0.754^{***}	1.086^{**}
	(0.063)	(0.615)	(0.086)	(0.427)	(0.063)	(0.219)	(0.075)	(0.500)
Scheduled caste	-0.516^{***}	-1.298^{***}	0.502^{***}	0.945^{***}	-0.465^{***}	-0.442^{***}	0.494^{***}	0.690^{**}
	(0.039)	(0.346)	(0.062)	(0.222)	(0.028)	(0.121)	(0.035)	(0.294)
Other backward caste	-0.214^{***}	-0.558^{***}	0.159^{***}	0.377^{***}	-0.231^{***}	-0.223^{***}	0.276^{***}	0.346^{***}
	(0.033)	(0.165)	(0.051)	(0.117)	(0.020)	(0.047)	(0.029)	(0.116)
Year F.E.	yes	yes	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes	yes	yes
Censored Observations	89,618	89,618	8177	8177	55318	55318	20555	20555
Uncensored Observations	6987	6987	88,296	88,296	42953	42953	77782	77782
P-value of Wald Test of Exogeneity		0.00		0.00		0.84		0.45

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 9: Baseline results – extensive margin below/above state poverty line

	Dep. variable:							
	Below state poverty line				**Above state poverty line**			
	LPG		**Fuelwood**		**LPG**		**Fuelwood**	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Electricity	0.019***	−0.293***	−0.019***	0.984**	0.103***	−0.696**	−0.089***	0.607
	(0.003)	(0.114)	(0.005)	(0.431)	(0.006)	(0.283)	(0.008)	(0.464)
Price Fuelwood	0.002***	0.002	−0.012***	−0.010	0.002*	0.003**	−0.004**	−0.004**
	(0.001)	(0.002)	(0.003)	(0.007)	(0.001)	(0.001)	(0.001)	(0.002)
Price LPG	−0.001**	−0.000	0.001*	−0.001	−0.002*	−0.003***	0.002*	0.003**
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
HH size	0.002***	0.009***	0.000	−0.021**	−0.002**	0.009**	0.004***	−0.006
	(0.000)	(0.002)	(0.001)	(0.009)	(0.001)	(0.004)	(0.001)	(0.006)
Total land owned	0.000	0.000	0.000**	0.000	0.0002	0.002*	0.0001	−0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Education head of HH	0.006***	0.012***	−0.006***	−0.027***	0.041***	0.056***	−0.037***	−0.051***
	(0.001)	(0.003)	(0.001)	(0.009)	(0.001)	(0.006)	(0.001)	(0.009)
Sex of head of HH	−0.007**	−0.008**	0.015***	0.019**	−0.059***	−0.075***	0.054***	0.068***
	(0.003)	(0.004)	(0.005)	(0.009)	(0.005)	(0.009)	(0.005)	(0.012)
Age head of HH	0.000***	0.001***	−0.000	−0.001**	0.003***	0.004***	−0.002***	−0.003***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Hindu	0.002	0.002	0.009	0.008	−0.000	0.004	0.001	−0.002
	(0.004)	(0.005)	(0.008)	(0.013)	(0.007)	(0.009)	(0.007)	(0.009)
Scheduled tribe	−0.014***	−0.047***	0.027***	0.136***	−0.100***	−0.177***	0.121***	0.187***
	(0.005)	(0.015)	(0.007)	(0.052)	(0.009)	(0.029)	(0.010)	(0.046)
Scheduled caste	−0.013***	−0.027***	0.012	0.060**	−0.087***	−0.133***	0.084***	0.125***
	(0.003)	(0.007)	(0.007)	(0.023)	(0.006)	(0.017)	(0.007)	(0.027)
Other backward caste	−0.009***	−0.015***	−0.002	0.020	−0.055***	−0.075***	0.047***	0.065***
	(0.003)	(0.005)	(0.007)	(0.014)	(0.005)	(0.009)	(0.006)	(0.013)
Year F.E.	yes	yes	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes	yes	yes
Observations	44,790	44,788	44,790	44,788	150,482	150,482	150,482	150,482
F-stat		6.51		6.51		5.23		5.23
AR F-stat		16.71		8.63		27.82		7.18

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 10: Robustness: share of fuel expenditures

	Dep. variable: Share fuel expenditures					
	Full sample		**Below median**		**Above median**	
	(1)	(2)	(3)	(4)	(5)	(6)
Electricity	1.35443***	11.092**	2.208***	16.034**	1.752***	−13.350
	(0.089)	(4.845)	(0.097)	(6.264)	(0.121)	(9.463)
Price Fuelwood	0.174**	0.163**	0.237***	0.233***	0.169**	0.170**
	(0.071)	(0.071)	(0.078)	(0.089)	(0.075)	(0.077)
Price LPG	−0.044***	−0.038***	−0.035***	−0.039**	−0.025***	−0.047**
	(0.011)	(0.014)	(0.011)	(0.018)	(0.009)	(0.019)
HH size	−0.764***	−0.865***	−1.031***	−1.271***	−0.889***	−0.670***
	(0.013)	(0.055)	(0.018)	(0.114)	(0.016)	(0.119)
Total land owned	−0.013	−0.015	−0.003	−0.003	−0.041***	−0.034**
	(0.010)	(0.013)	(0.002)	(0.004)	(0.013)	(0.014)
Education head of HH	−0.407***	−0.644***	−0.137***	−0.472***	−0.251***	−0.019
	(0.011)	(0.122)	(0.012)	(0.155)	(0.012)	(0.144)
Sex of head of HH	−1.277***	−1.107***	−1.728***	−1.594***	−1.040***	−1.345***
	(0.084)	(0.132)	(0.105)	(0.149)	(0.093)	(0.220)
Age head of HH	−0.0004	−0.019**	0.026***	0.004	0.020***	0.035***
	(0.002)	(0.010)	(0.003)	(0.010)	(0.002)	(0.010)
Hindu	0.100	0.022	0.076	−0.006	0.191**	0.220*
	(0.083)	(0.118)	(0.096)	(0.156)	(0.090)	(0.122)
Schedued tribe	1.014***	2.222***	0.307*	2.006**	0.521***	−0.741
	(0.152)	(0.675)	(0.169)	(0.865)	(0.154)	(0.798)
Scheduled caste	1.294***	1.994***	0.631***	1.542***	0.514***	−0.237
	(0.092)	(0.368)	(0.104)	(0.446)	(0.096)	(0.475)
Other backward caste	0.490***	0.762***	0.171*	0.592**	0.254***	−0.033
	(0.067)	(0.157)	(0.090)	(0.232)	(0.069)	(0.197)
Year F.E.	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes
Observations	195,099	195,099	96,645	96,645	98,454	98,454
F-stat		13.15		9.71		3.96

Notes: All regressions contain a constant. Standard errors in parentheses are robust and clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.

8 Figures

Figure 1: District wise night light intensity, 1993

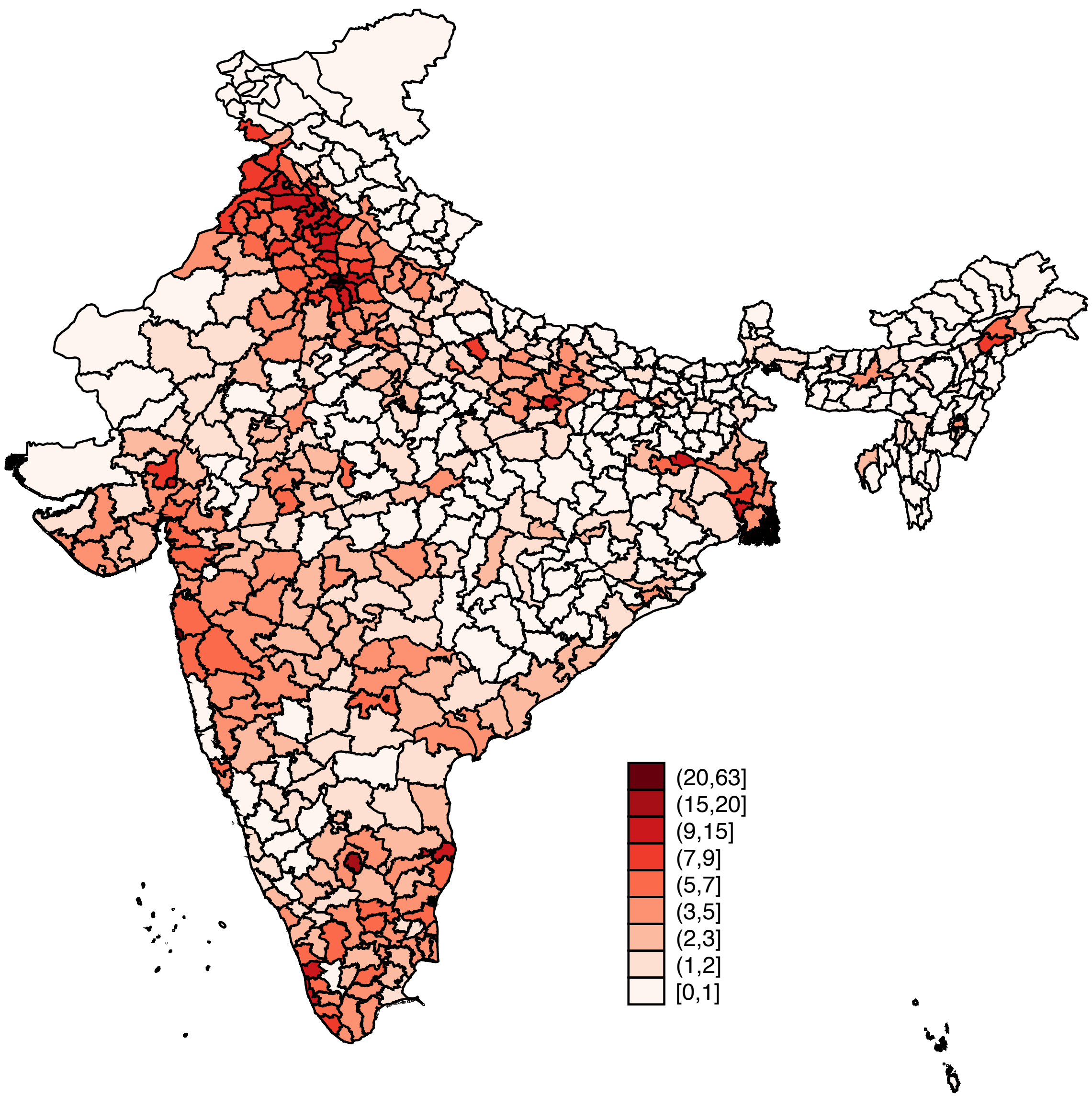

Figure 2: Coefficients on LPG adoption by quartile

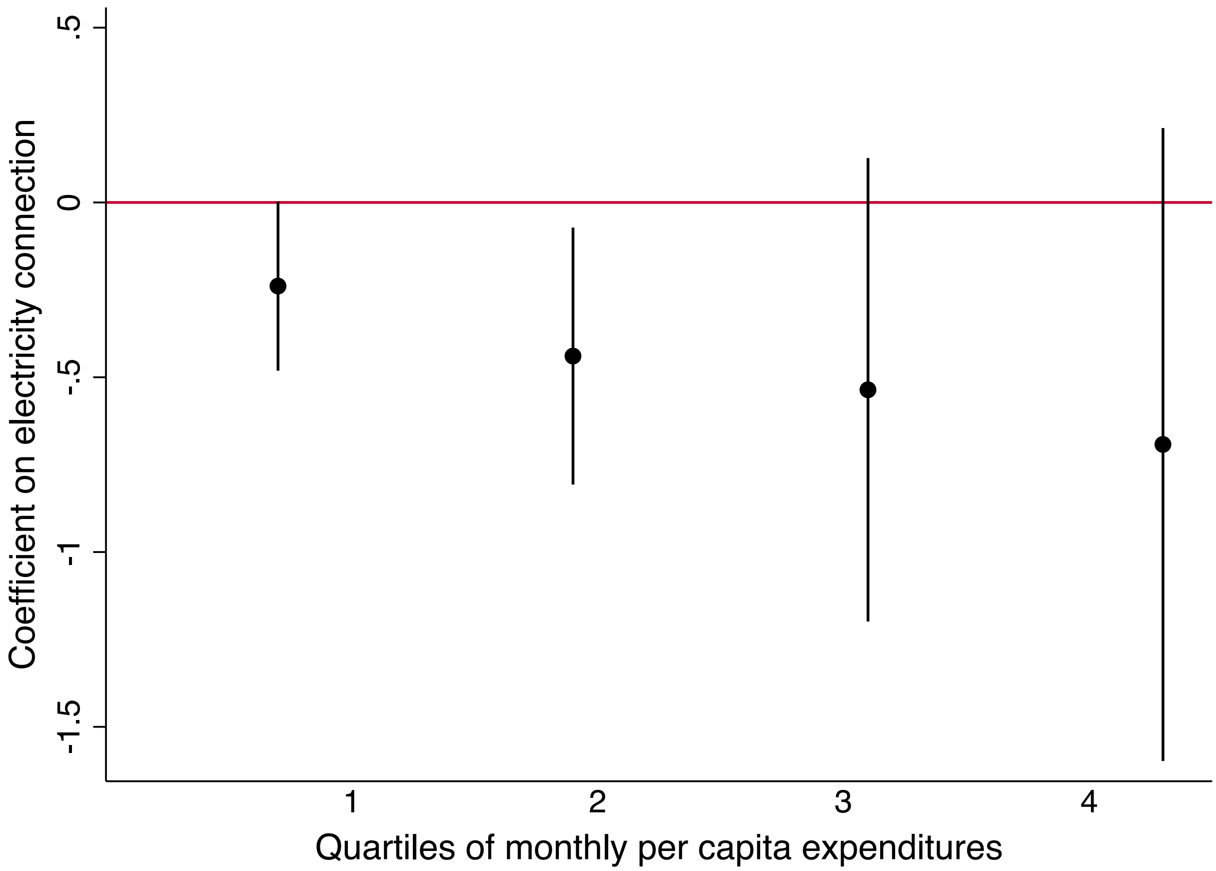

Figure 3: Coefficients on fuelwood adoption by quartile