CIRANO
Allier savoir et décision

Electrification and Cooking Fuel Choice in Rural India

RIDHIMA **GUPTA**
MARTINO **PELLI**

2021S-19
CAHIER SCIENTIFIQUE
CS

CIRANO

Knowledge into action

Center for Interuniversity Research and Analysis on Organizations

ISSN 2292-0838 (online version)

Electrification and Cooking Fuel Choice in Rural India *

Ridhima Gupta †, Martino Pelli ‡

Abstract/Résumé

This study investigates the causal link between electrification and the adoption of modern (and cleaner) cooking fuels, more specifically Liquefied Petroleum Gas (LPG). In order to correct for the potential endogeneity in the placement of electrical infrastructure, we exploit an instrumental variable approach. Our instrument interacts state-level supply shifts in hydroelectric power availability with the initial level of electrification of each district. The results are consistent with a choice set expansion under a fixed budget constraint. We find that electrification leads to an increase in the probability of adoption of (free) biomass fuels and a decrease in the probability of adoption of (costly) modern cooking fuels. These results are statistically significant only for the poorest households in our sample, while they become statistically insignificant when we move to richer households. The same is true for the share of expenditure in a specific fuel. These results seem to indicate that electrification, by creating an additional strain on households' finances, pushes them back on the energy ladder.

Keywords/Mots-clés: Rural Electrification, Cooking Fuel, Energy Ladder, Fuel Stacking

JEL Codes/Codes JEL: O12, O13, Q56

* We thank, without implicating them, Eswaran Somanathan, Daniel Kaffine, Sarah Jacobson, Jörg Peters, Soham Sahoo, Subhrendu Pattanayak, and participants at the Canadian Economic Association Meeting (2019), the Canadian Resource and Environmental Economics Association Meeting (2019), the 13th Annual Meeting of Environment For Development (2019), the 4th Meeting of the Sustainable Energy Transition Initiative (2019), and the Annual Economic Growth and Development Conference at ISI Delhi (2018), as well as seminar participants at Kansas University and Lulea University of Technology for helpful comments and suggestions. Rakesh Sinha provided outstanding research assistance. Funding: This work was supported by the Sustainable Energy Transition Initiative (SETI) [grant number 16018]; the *Fonds de Recherche Société et Culture Québec* [grant number 208237]; and the JanWallander and Tom Hedelius foundation [grant number P17-0059].

† Corresponding author, Faculty of Economics, South Asian University, India; Email: ridhima@sau.int.

‡ Department of Economics, Université de Sherbrooke, Sherbrooke, Quebec, Canada CIREQ, CIRANO, and GREDI. Email: martino.pelli@usherbrooke.ca.

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 (see for instance Hosier and Dowd, 1987; Leach, 1992; Bruce et al., 2000; 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 (e.g. Masera et al., 2000; Heltberg, 2004; Ruiz-Mercado and Masera, 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; Po et al., 2011; Khandker et al., 2012; Kishore and Spears, 2014; Peters and Sievert, 2016; Van de Walle et al., 2017). In India, electricity is rarely used for cooking. The benefits of electricity consist primarily of improved lighting and providing power for consumer appliances or agriculture. Only 19% of the rural households use Liquified Petroleum Gas (LPG) as a primary fuel for cooking, while the majority of the rural population still cooks with biomass fuels such as fuelwood or crop residues.[1]

For these reasons, it is not surprising that 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 exposure of children under five and adult women (Smith, 2000). The daily average particle level in these households ranges from between 1000 and 2000 $\mu g/m^3$ of PM_{10} (particles of mass smaller than 10 micron diameter,

[1] The primary cooking fuel is the fuel used in the majority of cases by the household.

Smith, 2000). This is 10-20 times higher than the national ambient air quality standard of PM_{10} set by the government of India (Pant et al., 2019). Exposure to ambient and household $PM_{2.5}$ (particles of mass smaller than 2.5 micron diameter) is among the largest single causes of premature mortality in India (Cohen et al., 2017). 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 the 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. Other papers, such as Barron and Torero (2017), focus on the impact of electrification on indoor pollution. We focus more specifically on the channel through which indoor air pollution may be affected.

The arrival of an electric connection may impact households in two different ways. Electricity is a fundamental input for the utilisation of a large variety of appliances and, therefore, gaining access to a connection opens up a lot of new consumption possibilities. The size of the choice set of a recently electrified household increases significantly. With a larger choice set, households face a new trade-off. Households may decide to: *i)* invest in new appliances that are going to increase their productivity level (such as an irrigation pump or a small mill) or help them optimize their labor supply (for instance a fridge that allows less frequent visits to the market), or *ii)* invest in entertainment technology (such as a televisions set). Clearly, the way in which a household decides to use the recently acquired electricity is going to determine the impact that electricity has on its budget and, through this bias, its impact on other consumption choices, like the ones related to energy.

If households decide to use electricity for productive activities, following the traditional energy ladder theory and results from a large literature (see for instance International Evaluation Group, 2008; Khandker et al., 2012; Lipscomb et al., 2013; Rao, 2013; Chakravorty et al., 2014, 2016; Bridge et al., 2016; Van de Walle et al., 2017), we would conclude 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. Instead, if households are going to favour entertainment goods, or any other use of electricity which is not productivity enhancing, we would not observe an increase in the household income, but we would observe an increase in their expenditures.

In order to disentangle these two effects, we use a nationally representative repeated cross-section of rural households constructed from two waves of the National Sample Survey (NSS). The waves used are 2004 and 2009. They cover one of the periods in which the government was most active with respect to rural electrification. These years correspond to the first phase of the Rajiv Gandhi Grameen Vidyutikaraan Yojana (RGGVY) or Rural Electricity Infrastructure and Household Electrification Scheme. Understanding this new and important effect of electrification 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 the use of modern fuels presents a number of empirical challenges. These challenges are common in the literature that identifies causal effects of big infrastructure projects (see for instance Aschauer, 1989; Garcia-Mila and McGuire, 1992; Holtz-Eakin, 1993; Röller and Waverman, 2001; Duflo and Pande, 2007; Dinkelman, 2011; Allcott et al., 2016). Electrification, like other infrastructure investments, is not randomly assigned. Governments may aim infrastructure investments to areas that are already growing faster and, therefore, electrification may always depend on a range of unobservables. In the specific case of this paper, some omitted variable could contemporaneously impact a higher rate of electrification and a higher rate of adoption of modern cooking fuels. We tackle this endogeneity problem by using an instrumental variable 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 a 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 multiply these state level shifts in hydro-electric power availability by the initial level of electrification of each district.[2] In other words, we weight the supply shifts by the initial level of electrification of each district. Because of the high cost of installing new electric infrastructure (transmission and distribution networks and sub-stations), it is more feasible and affordable for states to add additional connections in districts that are already characterized by higher electrification rates. 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.[3] The selection of new sites for hydropower plants depends primarily on topographic considerations, and is therefore plausibly exogenous to economic activity.[4] Hydro is the second largest source of power in India after coal, and it represents roughly 14% of installed capacity.

Given the expected heterogeneity of our treatment effect, throughout the analysis we estimate Local Average Treatment Effects (LATE), i.e. the effect of electrification on the subpopulation which received a connection because of a change in the availability of hydro supply.[5] Despite the fact that the reaction of households to a new connection may be similar irrespectively of how the power is generated, one should keep in mind that the sample of

[2]A district is an administrative division of an Indian state or territory. In some cases districts are further subdivided into sub-divisions, and in others directly into tehsils or talukas. According to the 2011 decadal Census, India is composed of 640 districts (but the the Reserve Bank of India (RBI), counts 644 districts and the state-level Planning Commission bodies, that perform the crucial task of collecting development indictors, stop at 556 districts). The database used in this paper, NSS, uses the definition used by the Census.

[3]The 11 years gap is dictated by the availability of the night lights data.

[4]This paper focuses on the period between 2004 and 2009, it is likely that most of the more promising hydro-sites were already exploited by 2004, leaving only second choice sites available.

[5]Overall, the generation of thermal power increased more than that of hydro power over the period of interest. Thermal power accounts for 69% of the additional power, while hydro power account for only 26% of the additions (`http://www.cea.nic.in/reports/others/planning/pdm/index.pdf`). These numbers vary greatly across states. Yet, as will be clear later in the paper, the weak monotonicity assumption holds, i.e. our instrument affects all households in the same direction.

population analyzed in this paper, the compliers of our analysis, is characterized by an improvement in its electricity connection (either a new connection or one providing more power) due to a shift in hydro supply weighted by the historical electrification of the region.

We analyze both the intensive and the extensive margin of the decision regarding the choice of the primary 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 intensive margin, instead, 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.

The results from the main specification point in the direction of the second mechanism highlighted above: households receiving a connection invest in new goods which are not going to enhance their productivity. For this reason, we find that electrification leads to an increase in the probability of adoption of fuelwood and a decrease in the probability of adoption of LPG. One way to test whether the results found are related to a change in the choice set is to look at the heterogeneity of the results across households' wealth levels. We find that these effects are statistically significant for the poorer part of the sample – households whose expenditures per month are below average – while they are statistically insignificant for the richest part.[6] The same is true for the intensive margin. We also find that these effects are more accentuated in male-led households compared to female-led ones. This result is consistent with Duflo (2003) and Qian (2008), which show that, on average, women take less impulsive decisions, aimed at improving the welfare of the household. We then run a series of robustness tests using alternative definitions of a household's wealth level

[6]The instrument becomes weak when we split the sample across several dimensions. In these cases we construct the Conditional Likelihood Ratio (CLR) confidence set, which is robust to weak instruments.

and a placebo test for the exclusion restriction which gives a certain degree of trust to our results.

Electricity allows people to include goods such as television sets, or electric fans in their consumption basket. Goods which are in principle not accessible without an electricity connection. Television sets do not have any close substitute, while LPG does have several substitutes, which are completely free, such as fuelwood, dung, or crop residues. This paper does not claim that the only adjustment that households make after receiving an electrical connection is through energy consumption, yet, energy is going to be one of the dimensions along which households adjust their spending.

Precise estimates of these effects are 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). In order to justify these important expenditures, 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. Yet, in light of our results, it seems that, for the poor, access to electricity makes the switch to cleaner cooking fuels even harder. As a consequence, one of the main take away of this paper for policymakers is that the introduction of electricity should not be done without a careful consideration of its impact on the poorest households and the provision of subsidies (for instance on LPG cylinders) in order to help them avoid the negative consequences of an additional regular expenditure on an already tight budget.

Our results are in line with recent findings by Lee et al. (2020a). They also conclude that only certain subgroups of the population – those willing to pay for it – benefit from electrification. As a consequence, electrification should never come alone, but always be combined with complementary programs. In a recent paper, Burlig and Preonas (2016) find that electrification has much smaller effects that what has been found in the literature so far. Differently from Burlig and Preonas (2016), our paper focuses on household level outcomes and not on village level outcomes. Unfortunately, the perfect identification strategy to capture the effects of electrification has not yet been developed and, therefore, the impact of electrification is always dependent on the identification strategies adopted and the sample of population analyzed.[7]

The remainder of the paper is organized as follows. Section 3 presents the data used. Section 4 focuses on the empirical strategy and Section 5 discusses the results. Section 6 describes robustness tests and, finally, Section 7 concludes.

2 Background

In this section, we provide some background information on rural electrification, hydropower generation, and LPG utilization in India. This information will help the reader to frame our results in the right context and to better understand our identification strategy.

2.1 Rural electrification in India

Electricity prices are highly subsidized in India, especially for the agricultural sector (Ryan and Sudarshan, 2020), partly for this reason supply lags demand, especially during peak periods, leading to frequent outages and voltage fluctuations (Modi, 2005). Between 2004 and 2009 Indian average per capita electricity consumption moved from 451.6 to 598.5 kilowatt

[7] As the literature shows, results very greatly across countries, methodologies and outcome variables. Dinkelman (2011) finds a positive effect on female employment in South Africa and, similarly, Grogan and Sadanand (2013) find an increase in the probability of labor force participation for females. Bensch et al. (2011) and Lenz et al. (2017) find a positive effect on lighting hours but no other effect in Rwanda, also Lee et al. (2020b) finds no effect of electrification in Kenya. Khandker et al. (2013) find a positive impact on households' income and expenditures in Vietnam; while Lipscomb et al. (2013) find similar, yet smaller, effects from Brazil and Chakravorty et al. (2016) finds larger results for the Philippines.

hour (kWh), a significant increase (32.5%), which left Indian consumption just above Sub-Saharan Africa. Cost recovery is low and has actually declined over time (69% in 2001-2002). These low rates are mainly due to losses during transmission and distribution, which rose from 25% in 1997-1998 (Modi, 2005) to about 39% in 2000-2001 (Oda and Tsujita, 2011). Estimates suggest that only 55% of the power supplied is billed and only 41% is paid for (Modi, 2005). Infrastructure theft has led to further declines in coverage (Balachandra, 2011). However, there has been significant improvement in average consumption and production of electricity. The generation capacity grew from 1362 MW in 1947 to nearly 74,699 MW in 1991. Over the same period, per capita consumption increased from 15.55 kWh to 252.7 kWh (Modi, 2005). Currently, total capacity in India exceeds 371 GW.

Over the last decades, India has gone through a series of national electrification programs in order to bring the number of electrified villages from 1,500 out of 597,464 in 1947 (year of its independence) to 576,554 in 2014 (Burlig and Preonas, 2016). These programs started as early as 1950 and culminated in 2005 with the launch of the Rajiv Gandhi Grameen Vidyutikaraan Yojana (RGGVY) or Rural Electricity Infrastructure and Household Electrification Scheme, which merged all the existing programs.[8] The goal of this program was to fulfill the National Common Minimum Programme (NCMP) goal of universal electrification within five years through a program financed 90% by the Central government and 10% by the Rural Electrification Corporation (REC). The scheme covered all hamlets with a population larger than 100.[9]

Under the scheme, state power utilities had to prepare detailed project reports including a

[8]RGGVY covered all states except Andaman & Nicobar Islands, Chandigarh, Dadra & Nagar Haveli, Daman & Diu, Delhi, Goa, Lakshadweep, and Puducherry

[9]According to the program, in 2005, a village was declared electrified if three conditions were fulfilled: *i*) basic infrastructure such as a distribution transformer and distribution lines are provided in the inhabited locality as well as the dalit basti hamlet where it exists (i.e. hamlet of backward castes); *ii*) Electricity is provided to public places like schools, panchayat office, health centers, dispensaries, community centers, etc.; *iii*) The number of households electrified should be at least 10% of the total number of households in the village.

list of all the un-electrified villages and hamlets (with population above 100). Utilities could also include already electrified villages in need of intensive electrification. All below poverty line (BPL) households were to be provided with 100% capital subsidy for the connection cost of Rs. 2,200 (approximately USD 32). While above poverty line (APL) households would receive 10% of the cost as a soft loan from the REC.

Between 2004 and 2008, 47,826 villages and 2,771,610 households have been electrified through RGGVY of which 2,293,770 were BPL. Moreover, in 40,838 already electrified villages electrification has been intensified. Our sample is well suited for the research question at hand since it covers this very active period. According to Section 43 of the Electricity Act (2003) the distribution licensee has to provide electricity within one month of application for a connection. Section 6 of the same act states that state governments and the central government have to act jointly in order to provide universal access to electricity. It is the responsibility of state governments to supply at least 6 to 8 hours of power per day to villages electrified under RGGVY.

Despite the fact that many poor rural households received a connection to the grid for free, monthly electricity costs still represent an important expenditure for poor rural households. As pointed out in Agrawal et al. (2019), households face several problems related to electricity expenditures. First, a large share of the connections in rural India are unmetered and, therefore, incur a fixed cost. For instance, an unmetered connection in Uttar Pradesh is billed 400 Rs. a month even for a minimum use of power. Second, most metered connections are not billed regularly, in some cases just once a year or once every several years, increasing the burden of arrears for late payment.

2.2 Hydro power generation in India

India has been among the first countries in Asia to adopt hydropower, with the hydropower plants at Darjeeling and Shivanasamudram, established in 1898 and 1902, respectively. Over the last century India remained a dominant player in global hydroelectric power development, and is currently ranked fifth in the world in terms of hydropower generation installed capacity, with 50.07 GW (International Hydropower Association, 2020). In 2004 India had 183 hydro power plants with a total capacity of 31,218.25 MW, corresponding to the blue circles in the map on the left of Figure 1 (Global Energy Observatory, 2019). The map shows that the location of these plants is not correlated with economic activity, but to the topology of the country. Between 2005 and 2009, 18 additional hydro power plants were added to the system, for a total of an additional 6,806 MW of capacity. These additional 18 plants are shown in the right panel of Figure 1.

[Figure 1 here]

While the absolute number of plants grew only by 9.8%, installed capacity grew by 21.8%, corresponding to an average increase of 4% per year. This large expansion of the hydro sector corresponds with the implementation of the RGGVY scheme. In the following 9 years (2010-2018), 31 new hydro power plants were commissioned totalling 7,427.02 MW, so capacity increased by an extra 19.5% corresponding to a measly 1.9% yearly growth rate of installed capacity.

The main advantage of using supply shifts coming from hydropower and not for instance from thermal power plants, is that hydropower cannot be built everywhere. In order to be able to build an additional hydropower plant the topology of the land has to be appropriate. As shown in the map on the left of Figure 1 we observe many hydropower plants along big mountain ranges, like the Himalayas in Jammu & Kashmir, Himachal Pradesh, and Uttarakhand, the Western Ghats in Maharashtra, Goa, Karnataka, and Kerala, or the Eastern Ghats in Odisha, Telangana, Andhra Pradesh, and Tamil Nadu (Global Energy Observatory, 2019).

2.3 LPG use in Rural India

The use of traditional solid fuels for cooking (e.g. coal, fuelwood or dung) has negative externalities at many levels: health (Lim et al., 2012), environment (Masera et al., 2006; Ghilardi et al., 2009; Boskovic et al., 2018), socio-economic status (Duflo et al., 2008; Kowsari and Zerriffi, 2011), and climate (Bond et al., 2004; Jeuland and Pattanayak, 2012). LPG is the most accessible clean cooking fuel in rural India, like in many other parts of the world and, therefore, more likely to deliver substantial benefits (Simon et al., 2014; Bruce et al., 2017). Yet, according to the 2011 Census only 11% of rural households use it as a primary cooking fuel, this number is even lower than the one observed in our dataset, 8.6% in 2004 and 12.2% in 2009.

While LPG has been present in India since 1950, over the last 7 years the Indian government has introduced several policies aimed at increasing LPG use. Launched in 2013, the Direct Benefit Transfer of LPG (DBTL), or Pahal, puts subsidies for the purchase of LPG cylinders directly into recipients' bank accounts. The goal of this program is to eliminate misuse of funds. In 2015, the government launched the program *Give It Up* to motivate households able to afford LPG at its market price to surrender their subsidies and transfer them to poorer households. Finally, in 2016, the *Pradhan Mantri Ujjwala Yojana* (commonly known as Ujjwala) program was started. This is the best known program, which aimed at providing LPG connections to 80 million poor households by 2019. Ujjwala offered a subsidized LPG connection. All these programs must account for fuel stacking, that is happening everywhere and allows households to quickly switch back to solid fuels (Mukhopadhyay et al., 2012; Hollada et al., 2017; Troncoso and Soares da Silva, 2017).

The cost of LPG (connection and regular fuel cost), and also the "lumpiness" associated with it, are important obstacles to LPG adoption and use for rural households (Lewis and Pattanayak, 2012; Rehfuess et al., 2014; Puzzolo et al., 2016; Gould and Urpelainen, 2018). Although it is clear that LPG offers important benefits, the determinants of its adoption and use are still not completely understood. This paper contributes to this large literature by investigating the role played by electrification. Previous contributions have looked at affordability (Cheng and Urpelainen, 2014; Alkon et al., 2016), characteristics of the head of household and of the household (Lewis and Pattanayak, 2012; Bhojvaid et al., 2014; Sehjpal et al., 2014).

3 Data

Data on households' expenditures are taken from surveys conducted by the National Sample Survey Organisation (NSSO).[10] These data are representative of the rural and urban population of India.[11] For our analysis, we use data for the rural sample for the years 2004-2005 (61^{st} Round), and 2009-2010 (66^{th} Round). These two datasets give us a repeated cross-section sample containing 136,221 households.[12] Table 1 reports summary statistics for the variables used in the paper. Panel A reports statistics for the variables of interest and the dependent variables, Panel B for the instruments used and, finally, in Panel C we present all the control variables used in the different specifications.

[Table 1 here]

68% of the households in our sample have a connection to the power grid.[13] Only 16.2% of the households in the sample use LPG as the primary fuel for cooking, while 73% of rural

[10]The National Sample Survey (NSS) has been set up by the Government of India in 1950 to collect socio-economic data employing scientific sampling methods.

[11]In order to ensure the representativeness of the sample, the survey collects, for each state or union territory an urban and a rural sample in proportion to the population reported by the Census. The survey uses a multi-stage sampling method, where the fist-stage units consist of villages in the rural sector and urban frame survey blocks in the urban sector. Villages are selected by a probability proportional to their size with a replacement method. Urban blocks include roughly 200 households each and are completely updated every 5 years (20% update every year). Second stage sampling is conducted in the field, since no list of households is available. Household lists are drawn up in the selected villages, urban blocks. Household sampling is then undertaken using simple random sampling without replacement.

[12]136,211 is reached after dropping outliers. The sample is composed of 78,017 households interviewed in 2004, and 58,204 in 2009.

[13]For most households in India, having a power connection does not mean receiving regular power, as the

households still rely on fuelwood.[14] NSS asks households to recall expenditures incurred and quantities purchased for almost all items of domestic consumption over the last 30 days. Total expenditures on any item include money spent on purchases and value of consumption out of home production. The latter is valued at the average retail prices prevailing in the household's village of residence. We use this data to calculate expenditure shares for all fuel items for each household. Panel A of Table 1 shows that while LPG represents roughly 1% of a household's monthly expenditures, fuelwood constitute 5.3%. These expenditure shares are computed conditional on the household using the specific fuel. The difference between the two shares is partly explained by the fact that households using LPG are on average richer than households using fuelwood as the primary cooking fuel.

Panel B presents statistics for the instrument used. As discussed in the introduction, the instrument we use in this paper is composed of the interaction of two separate variables. 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.[15] These supply shifts, related to hydro generation, represent on average 13% of predicted electricity demand, yet there is a lot of variation across states.[16] These supply shares 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. 11 years is as far back as we can go due to

quality of supply is fairly low. Unfortunately, the NSS survey has no information about the quality of the power supply. In this paper we define a household as electrified if it declares using electricity as the primary fuel for lighting. Hence, some households that switched status from unconnected to connected may have been connected all along but had not received a sufficient amount of power to use electric lighting as the main source of lighting.

[14] All these statistics are based on what each household declares as the primary cooking fuel and, therefore, it is possible that for a household reporting fuelwood as the primary cooking fuel we still observe expenditures in LPG. The remaining 11% is divided in the following way: 7.34% use biomass; 2.09% are in *other category*; 0.92% use kerosene; 0.39% has no cooking arrangement; and 0.06% use electricity.

[15] More details about the construction of this variable are presented in Section 4.

[16] In 2004/05 hydro generation represented 14.38% of total generation in India, Central Electricity Authority (2006)

data availability.

The database on night lights for India has been constructed by the University of Michigan in collaboration with the World Bank, using images taken by the Defence Meteorological Satellite Program (DMSP), run by the U.S. Department of Defence. This satellite program took pictures of the earth at night for 20 years, from 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. The India Lights API data is based on these images and is freely available for each district and month from 1993 to 2013.[17] Figure 2 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. 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 2.53. The interaction between this variable and the predicted hydro supply shift takes into account the fact that districts that already had more electricity connections (and, therefore, a higher light output in 1993) benefit more from the positive hydro supply shifts than districts with lower degrees of electrification.

[Figure 2 here]

The NSS survey 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 units bought. Since many households do not consume some of the fuel items in which we are interested – such as LPG – we proceed in the following way. 19% of the households in our sample use LPG and, therefore, report

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

a price. We use these values in order to compute a state level average LPG price, which we then attribute to all the households not reporting an LPG price. After having attributed a price to each household in our sample, we compute village level average prices, which, at this point, are composed by a mix of household reported prices and state level averages. We proceed in this way because LPG prices tend to be fixed at the state level.[18] For consistency, we used the same methodology to compute the village-level price of fuelwood. This approach allows us to minimize the risk of measurement error. In order to make expenditure and prices comparable across rounds we use industrial and agricultural consumer price indices reported by the Labour Bureau of India to deflate 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 reflects the socio-economic status of each household, measured through its size, land holdings, religion, the age, gender and education level of its head and caste.

4 Empirical Strategy

In order to study the impact of a connection to the grid on the choice of cooking fuel we split the analysis into two parts. First, we analyze the extensive margin response, i.e. the adoption of a specific fuel as primary cooking fuel and, second, we look at the intensive margin response, i.e. how is the share of expenditure in a given fuel going to be affected. The specification we use in both cases takes the following form:

$$y_{hdst} = \alpha_0 + \alpha_1 E_{hdst} + \alpha_2 X_{hdst} + \alpha_3 V_{dst} + \delta_d + \delta_t + \varepsilon_{hdst} \quad (1)$$

[18] "Here's How LPG Prices Have Changed In Past One Year. Find Out Latest Rates Here", *NDTV*, July 12^{th} 2020; Singh, S.

where y represents different outcome variables. Subscripts h, d, s and t denote household, district, state and year, respectively. First, we investigate the impact of an electric connection on adoption. In this case, the outcome variable is binary and takes the value 0 if a household does not use LPG (fuelwood) and 1 otherwise. Second, we investigate the intensive margin, i.e. does an electric connection increase the amount a household spends on LPG (fuelwood). In this case, we use the relative share of expenditure on LPG (fuelwood) in total monthly expenditure of the household. E is the electrification status of a household and X includes time-varying household specific controls, i.e. household size, religion, total land owned, information on the household's head age, sex, education level, and 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, clustered at the district level across all specifications.

Fuel choice and the presence of an electricity connection could depend on a variety of unobservable factors. If this were the case, equation (1) could not be interpreted causally. As mentioned above, 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. For instance, governments may reserve new investments in infrastructure to areas already experiencing more growth. Other unobservable economic trends can 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 population density in certain areas. For all these reasons, it is difficult to isolate the effect of infrastructure investments on development outcomes. This issue has been widely discussed in the literature, see for example, Aschauer (1989), Garcia-Mila and McGuire (1992) Holtz-Eakin (1993), Röller and Waverman (2001), Duflo and Pande (2007) and Allcott et al. (2016) just to mention a few.

We are going to instrument electrification using electricity supply-shifts due to the coming online of new hydroelectric power plants. Hydroelectric power in India is much more reliable than coal generated power (see Chan et al., 2014, which shows that between 1994 and 2009 Indian coal power plants were offline 28% of the time). An increase in the available supply is going to be correlated with an increase in the probability of receiving a new connection. The additional supply will also allow existing connections to receive more power and, therefore, increase the proportion of households using electricity as a primary source of lighting. Another advantage of hydroelectric generation is that the location of new power plants is closely related to the topology of the terrain and not so much to economic considerations, which would make these new plants endogenous. Our instrument stems from Allcott et al. (2016). In their paper, they instrument electricity shortages with state-level supply shifts in hydro-electric power availability.

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 supply shifts by the initial level of electrification of each district. The importance that the supply shifts measured by Allcott et al. (2016) will play in each district is likely to be contingent on the district's initial level of electrification. A higher initial level of electrification allows us to use more effectively the positive supply shifts, while no initial electrification would render them virtually useless. As pointed out in Greenstone et al. (2014), the cost to extend the power grid to a village which is more than 15km away from the existing infrastructure is very high, and estimated at around 150,000 dollars. For this reason, we believe that it is more feasible and affordable for the state to add additional connections in districts that are already characterized by higher electrification rates. If a district has a low level of electrification it also means that it has a smaller transmission and distribution network, and less sub-transformers and, therefore, the investment needed to electrify additional households is more important. 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 a district.

The instrument constructed by Allcott et al. (2016) consists of predicted state-level supply shifts from hydroelectric generation.[19] More precisely, it 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.[20] Computing predicted electricity demand in this way

[19] 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) look 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.

[20] Denoting by s and t the state and year at hand, and by r and y other states and 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.

minimizes fluctuations due to important shortages.

Allcott et al. (2016) also need 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. 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 hydro generation 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 regress 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).[21]

Our first stage specification takes the following form

$$E_{hdst} = \beta_0 + \beta_1(H_{st-1} * L_{sd0}) + \beta_2 H_{st-1} + \beta_3 X_{hdst} + \delta_d + \delta_t + u_{hdst} \tag{3}$$

where L represents a district's light intensity and H is Allcott et al. (2016)'s state-level instrument. We use the lag of the hydro instrument to allow enough time for the positive supply shifts to take effect and impact the electrification level. 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)'s instrument and a district's initial light intensity.[22] Finally X is a

[21] As shown in Table 1 the average hydro supply shift is 13%, which compares well with average hydro generation in India, which was 14.38% in 2004/05, Central Electricity Authority (2006).

[22] 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 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 true we

matrix of time-varying household specific controls and u is the error term, clustered at the district level.

Table 2 reports first stage estimates. The table is organized as follows, columns (1) and (2) contain first stage estimates 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 households with per adult equivalent monthly expenditure below the mean level (of about 724 Rs, equivalent to roughly 10 USD), while column (4) shows it for households whose income is above the mean level. Let us focus on the effect of the interaction, i.e. our instrument in Column (2), the full specification. Its impact is positive and statistically significant at the 1% level across most specifications. This means that a higher hydro share 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. The best way to look at this result is to compute the marginal effect of a supply shift across the full spectrum of initial levels of electrification. We plot the marginal effect in Figure 3. The gray histogram shows that most of the districts had low levels of initial electrification. We can also notice that for districts with no initial electrification, the impact of a supply shift is statistically insignificant. Yet, as the initial level of electrification increases this insignificant effect becomes positive and statistically significant.

[Table 2 here]

[Figure 3 here]

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 variables left in the error term in the model for cooking fuel decisions in 2004 and 2009 after controlling for district and year fixed effects and a number of time-varying household- and village-level controls should not to be correlated with the district light intensity in 1993. Hence, we are confident in claiming that this condition is satisfied.

When running the first stage on the whole sample, we obtain an F statistic of 12.1, above the critical level of 10. When we split the sample the value of the F statistic decreases to 6.76 and 8.03, respectively. This means that our instrument becomes weaker when applied to the two subsamples. For these cases, we first compute the modified version of the F statistics proposed by Montiel Olea and Pflueger (2013) in order to ensure that we are effectively dealing with a weak instrument problem. In all the cases where this alternative F statistic is below the critical value of 10, we compute the 95% Conditional Likelihood Ratio (CLR) confidence set. These confidence sets are the best suited for weak instrument robust inference in the case of one endogenous variable and multiple instruments, as shown in Mikusheva (2010). These confidence sets are based on a Limited Information Maximum Likelihood (LIML) estimator, and not a simple two stage least square (2SLS) estimator and, therefore, we also report the LIML coefficients. These extra steps allow us to evaluate how much trust we should put in our estimates and whether they are robust to weak instruments.

Before moving on to the results, let us discuss the external validity of the analysis. The response of each household to our treatment may be different, i.e. we are estimating heterogenous treatment effects. In practice, this means that the treatment effect identified in this paper is valid only for the subpopulation of *compliers* (i.e. households which received a connection *because* of a hydro supply shift) and, therefore, we are identifying a LATE. The implication of this identification strategy is that, even if the results are internally valid, some caution should be exercised when we want to generalize them. While it is plausible to think that all households receiving a connection react in a similar manner, the results obtained from a LATE estimation are directly applicable only to the group of compliers, in this case households whose electricity connection improved because of a change in hydro supply and that live in districts that historically have known higher rates of electrification.

The rollout of the power grid and of the LPG network may have followed the same pattern. Even thought the 11 years lag between the initial level of electrification and the current choices with respect to cooking fuel adoption should ensure a low correlation between the two, if the two networks developed following a similar pattern the initial level of electrification may not be completely exogenous to LPG adoption today.

5 Results

First, we present the results for the extensive margin, focusing on the adoption of LPG and fuelwood as primary cooking fuels. Second, we look at the intensive margin, measuring the impact of electrification on expenditures on LPG and fuelwood.

Extensive margin

Table 3 reports results for the extensive margin estimations, i.e. adoption. For each specification we report OLS and IV coefficients for a linear probability model.[23] Columns (1) and (2) show results for LPG, while (3) and (4) for fuelwood.

[Table 3 here]

Let us focus on the IV results for the linear probability model.[24] The results on the variable of interest, *Electricity*, are all statistically significant at least at the 10% level. These coefficients are negative for LPG and positive for fuelwood. This implies that, when a

[23]We also run the model using a probit specification instead of a linear probability model, the results are very similar in terms of magnitude, sign and statistical significance.

[24]From Table 3, and all the following tables, one sees that the direction of the bias in the OLS specification is positive in the case of LPG and negative in the case of fuelwood. This is easily explained by the sign of the covariances: *i)* between the electricity connection variable and the instrument and, *ii)* between the electricity connection variable and the dependent variable. While the covariance between electricity connection and the instrument is positive in both cases (see results of the first stage specification), the covariance between electricity and LPG use is expected to be positive and the one between electricity and fuelwood use is expected to be negative. For instance, we expect areas experiencing a higher rate of economic growth to have a larger number of electricity connections and at the same time more people using LPG and less people using fuelwood. Since $cov(E, y)$ and $cov(E,(H * L))$ move in the same direction in the case of LPG, we expect the bias to be positive. In the case of fuelwood, instead, $cov(E, y)$ and $cov(E,(H * L))$ move in opposite directions and, therefore, we expect the bias to be negative.

household gets connected to the power grid, the probability that it will use LPG for cooking decreases by 56.6 percentage points, while the probability that it will use fuelwood increases by 52 percentage points. Between 2004 and 2009, the electrification rate in India went up from 0.64 to 0.74, implying a change of ten percentage points. These numbers, together with our coefficients, imply that, *ceteris paribus*, fuelwood adoption increased by roughly 5.53 percentage points and LPG adoption decreased by 6.02 percentage points over this 5-year period as a result of the increase in the electrification rate.[25] According to the energy ladder theory, one would expect electricity to constitute a push towards modernization and, therefore, towards the adoption of more modern and less polluting cooking fuels. Yet, as we discussed in the introduction, electrification also expands the choice set of households. A new connection allows households to invest in new appliances for productivity or for entertainment, putting households, especially the poorer ones, in front of difficult choices when deciding how to allocate their limited resources.

The signs of the coefficients on the control variables 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. Richer households, with larger landholdings and a higher level of education, use less biomass and more modern (and usually more expensive) fuels. A more educated and older head of household is also 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. A male head of household is more likely to use fuelwood, while a female has a higher probability to use LPG. Since women are usually solely responsible for cooking, it make sense that when they have the choice they opt in larger numbers for the cleaner fuel.

[25]This numbers are obtained by multiplying the coefficients 0.52 and -0.566 by the improvement in the electrification rate, 0.7432-0.6369=0.1063, so 10.63 percentage points.

As we mentioned above, our results may be explained by an expansion of the choice set available to households. A natural way to test for this is to look at the behaviour of households as a function of the stringency of their budget constraints. Cooking with LPG is more expensive than with biomass. Biomass, consisting mainly of fuelwood, dung and crop residue, can easily be collected (for free), while LPG has to be bought (a five-member household cooking almost exclusively with LPG spends roughly 530 Rs. per month, based on a per capita consumption of about 5.2 kg/month at a price of 20.3 Rs/kg. The price comes from our summary statistics).[26] Households connected to the grid have to pay for the electricity they consume, irrespective of whether they use it for light or entertainment (for appliances such as television sets), and this without accounting for the costs of investments such as a television set, or an electric fan. These extra expenditures leave 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 average monthly per capita expenditure. In the average household, the monthly per capita expenditure is about 724 Rs (the exact value is 724.09 Rs, about 10 USD). Therefore, in the first sub-sample we keep the 92,773 households characterized by a monthly per capita expenditure lower than 724 Rs, while in the second sub-sample we keep the 43,381 richer households, with expenditures per capita equal or higher than 724 Rs.[27]

Panel A of table 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 724 Rs., while columns (5) through (8) present them for households situated above the 724 Rs. threshold. The IV coefficients obtained are statistically significant at the 5% level for the poorest part of the sample, while

[26]The price of 20.3 Rs/kg is an average of the price between 2004 and 2009. In 2020, the price of LPG is of around 56 Rs/kg and, therefore, the same five-member household would spend roughly 1,456 Rs. per month.

[27]If we use the median instead of the mean to divide the sample, the results are similar in terms of sign and magnitude but less statistically significant. These results are available upon request.

they are statistically insignificant for the richer part.[28] For households below the average expenditure level, the coefficient on LPG adoption is -0.526, so slightly smaller in magnitude than the one obtained for the baseline estimation. The coefficient for fuelwood adoption instead increases in magnitude to a value of 0.758. The electrification rate for this group of households went up from 0.55 to 0.66, i.e. an increase of 11 percentage points. This change implies that over this 5-year period, *ceteris paribus*, we would have observed a decrease in the adoption of LPG by 5.8 percentage points and an increase in the adoption of fuelwood by 8.34 percentage points.[29]

[Table 4 here]

These results seem consistent with the expansion of the choice set hypothesis. Electricity allows households to adopt new consumption habits, for instance, 35.2% of the households with expenditures below the mean own a television set. These new choices impose an additional constraint on their tight budget. An easy way to free up some extra budget is to revert to fuelwood for cooking, instead of buying LPG cylinders, household members collect fuelwood for free. Conversely, richer households do not suffer from the additional financial strain, and, for them, obtaining a power connection does not have an impact on the choice of cooking fuel. Table 5 provides more information on appliance uptake following electrification. This table shows instrumental variable results for the impact of electrification on television set, radio and electric fan adoption. As expected, electrification has no impact on the adoption of radios, since most of these can be run on batteries. Yet, we observe a big jump for television sets and electric fans adoption, an increase by 63.3 and 79.3 percentage

[28] In order to ensure that the results for the two subsamples are statistically different we construct a confidence set for their difference. Assuming that the two subsamples are independent, the 95% confidence set around the difference of the two coefficients for LPG is [-0.4405;-0.4289], while for fuelwood is [0.8999;0.9121]. This confirms that the results for richer and poorer households are statistically different.

[29] The exact numbers for the electrification rate are 0.6621 for 2009 and 0.5526 for 2004, giving a difference of 10.95 percentage points, which we then multiply by the two coefficients.

point for television sets and fans, respectively. Taking into account the increase in electrification of 10.63 percentage points, these coefficients imply that electrification increased ownership of television sets by 4.99 percentage points over the five years studied and of fans by 8.4%. Mean ownership across the sample is of 46.98% for television sets and 54.61% for electric fans. So the increases are larger then 10%.

[Table 5 here]

A considerable body of literature (see for instance Duflo, 2003; Qian, 2008) shows that women tend to take decisions that are less impulsive and aimed more at increasing the general welfare of the household, while men tend to take more impulsive decisions. It is therefore interesting to investigate whether we observe a difference in behaviour between male-headed and female-headed households when faced with a change in the available choice set. In order to do this, we interact the electricity dummy with a gender dummy for the head of the household. This leaves us with two endogenous variables. In order to take care of this additional endogenous variable we also interact the instrument with the gender dummy. The results of this specification are presented in Table 6 and are not surprising. In columns (1) to (3), we present results for LPG, first for the whole sample, followed by households below and above the average level of expenditures. Columns (4) to (6) have the same structure and show results for fuelwood. We observe that male-led households have higher positive effect on the probability of reverting to fuelwood and a higher negative effect on the probability of adopting LPG compared to female-led household, both effects are statistically significant. Moreover, once we interact with gender, the effect appears also for households whose expenditures level is above the average, and behaves similarly.

[Table 6 here]

Intensive margin

Table 7 presents intensive margin results, i.e. how does spending on 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 similar to the one presented above for the extensive margin.

[Table 7 here]

A new electrical connection decreases the probability of adoption of LPG as a cooking fuel, but seems to have no impact on the share of spending households dedicate to it. The coefficient obtained is 0.094 and it is statistically insignificant. When focusing on fuelwood we observe the same pattern as for the extensive margin, a new electrical connection increases spending in fuelwood, the coefficient here is 6.835, but it is only statistically significant at the 10 percent level. This coefficient implies an increase in the share of expenditures in fuelwood over the 5 years of 72.7 percentage points. At first, this number may seem large. Yet, the largest share of a household expenditure on fuelwood in the NSS sample is imputed using consumption multiplied by the prevailing regional price. This means that a large part of the increase does not come from an actual increase in spending but from an increase in the collection of fuelwood. Poorer households seem to decide to spend more on consumption goods becoming available because of the new form of energy, while cutting spending wherever possible, for instance by collecting fuelwood instead of buying LPG cylinders.

The pattern of the intensive margin results is similar to the one observed for the extensive margin. This similarity, reinforces our believe in an explanation related to an increase of the consumption choice set linked with a tight budget constraint. As before, in order to investigate this claim we split the sample according to the average monthly per capita expenditure. Results for this can be found in Panel B of Table 4. Columns (1) through (4) present results for households with monthly per capita spending below the 724 Rs. threshold, while columns (5) through (8) present results for households with spending above it. As in the extensive margin analysis, results are statistically significant only for the poorer part of the sample. These households, when faced with new expenditures derived from electric power seem to cut back on expenditures. We observe a decrease on expenditures on expensive fuels and, an increase in the share of income dedicated to fuelwood, the cheaper alternative.

Panel B of Table 6 shows results for the interaction with the gender of the head of the household for the intensive margin. Also in this case, results imply that women tend to spend more on cleaner fuels compared to men.

6 Robustness

This section is divided into two parts. First, we perform several tests aimed at our instrument and at whether it respects the exclusion restriction. Second, in order to further verify the claim that households revert to the use of biomass because electrification opens up a new array of possible consumption choices, we perform several robustness checks. First, instead of using expenditures to measure the wealth of each household, we use caste, education level, and an interaction with the expenditure level. Second, we verify our claim that electricity puts further stress on the budget of poor households by looking at the impact of electrification on their energy budget.

Instrument

We perform two separate tests to verify whether the exclusion restriction holds. First we run a placebo test and, second, we look at whether heavy rainfalls impact contemporaneously the instrument and the variable of interest.

The placebo test consists in running a reduced form specification on districts where the electrification rate did not change between 2004 and 2009, in order to show that if not through changes in the electrification rate, the instrument does not affect the adoption rate of LPG and fuelwood. First, we identified the 173 districts in which the electrification rate stays constant between 2004 and 2009.[30] The placebo test focuses on these districts. We run a reduced form specification, i.e. we regress the variable of interest of the structural equation directly on the instrument. If our instrument has an effect on the variable of interest only through a change in the electrification rate, the coefficients on the instruments in this regression should not be statistically significant. If this were the case the exclusion restriction would be verified. Table (8) reports results for the reduced form for LPG and Fuelwood adoption, estimated only over the districts that did not experience changes in their electrification rate. As expected the coefficients on the instruments are not statistically significant (not even when taken jointly).[31]

[Table 8 here]

The main source of variation of our instrument is the opening of new hydropower plants, yet it could still be weakly correlated with rainfall, as it relaxes the water availability constraint for hydroelectric power plants. Heavy rainfall could, therefore, create a positive supply shift, associated with an increase in the electrification rate and, at the same time, a shortage of LPG by disrupting the supply chain. In order to rule out the correlation between rainfall and the uptake of LPG, we run our main specification eliminating information collected during the monsoon months, characterized by heavy precipitation and, therefore, more likely to create a problem with the exclusion restriction. In India, the monsoon season

[30]We define the electrification rate as constant if the change between 2004 and 2009 stays within a 5 percentage points margin, in absolute value.

[31]The p-value for the coefficient on Hydro interacted with initial light intensity is 0.15 for LPG, and the one for a joint test on the coefficient on Hydro and Hydro interacted with initial light intensity is 0.35. If we look at the companion placebo test, the one for fuelwood, the p-value for the coefficient on Hydro interacted with initial light intensity is 0.57 and the p-value for a joint test on the coefficient on Hydro and Hydro interacted with initial light intensity is 0.75. The p-values on these coefficients taken together allow us to put a certain degree of trust in the placebo test, suggesting that our IV does not directly impact the decision to use either LPG or fuelwood as a cooking fuel.

occurs every year between the months of July and September. Since NSS asks respondent to recall their consumption over the last 30 days, we consider that households interviewed between August and October report their "monsoon" consumption. Table 9 reports results for our main specification (intensive and extensive margin for LPG and fuelwood) excluding observations collected during the monsoon months (roughly 20% of the sample). If the coefficients were biased by the correlation between LPG supply disruptions and heavy rainfalls, they should have become smaller or less statistically significant once the data collected during the monsoon month is discarded. Instead, we notice that the coefficients remain unchanged, passing this robustness test.

[Table 9 here]

In order to deal with the decrease in the strength of the instrument when we split the sample between richer and poorer households we also present the results of a specification using an interaction term instead of a split in the sample. The results for this specification can be found in table 10. The variable *Expenditure* is a dummy taking value 1 if a household's expenditures are below the mean level of expenditures (724 Rs.) and 0 otherwise. In this case the F statistic for both first stages is above the critical level of 10. The findings coincide with our baseline results in terms of sign and magnitude for the extensive margin, but become more statistically significant. We observe the same pattern for the intensive margin, in fact the statistical significance of the results is higher.

[Table 10 here]

Alternative measures of wealth

In order to confirm our findings, we split the sample according to two additional measures of wealth. First, we divide households by caste, and, second, we interact the impact of electrification with the education level of the head of the household.

Caste The caste system used to be linked to individuals' wealth level. It has been shown that a correlation between caste and wealth still exists (see for instance Zacharias and Vakulabharanam, 2011). We split our sample in five different ways. First, we identify the three backward castes: scheduled tribes, scheduled castes and other backward castes, we then take all the backward castes together and, finally, the other (upper) castes. All the backward castes taken together represent roughly 72.3% of our sample. Table 11 reports results for the extensive margin, while Table 12 for the intensive margin.

The extensive margin results, reported in Table 11, confirm what we observed before. When we take all the backward castes together, in column (4), an electric connection decreases the probability of adopting LPG, while it increases the probability of adopting fuelwood. The size of the coefficients obtained is similar to the one obtained in Table 4. Instead, when we look at the other castes, the result become statistically insignificant. We obtain the strongest results for other backward castes, in column (3) and schedule tribes, in column (1), while the results for scheduled castes are statistically insignificant (column 2).

[Table 11 here]

Table 12 shows results for the intensive margin, which are also similar to the one obtained in Table 4. When pulling all the backward castes together we find the same results we had for households below mean expenditures, no impact on LPG expenditures and a statistically significant, albeit only at 10%, result on fuelwood expenditures. Interestingly, when looking at the other castes we find a positive and statistically significant result on LPG expenditures following electrification. This result fits the standard energy ladder hypothesis.

[Table 12 here]

Education Education also serves as a proxy for wealth. On average, richer households have a higher level of education. We focus on primary education and create a dummy variable that takes value 1 if the head of the household has an education level higher than primary school and 0 if the level is at most primary school. In India primary school goes from grade 1 to grade 4. 49.8% of households have a head with at most a primary school education level.

We then proceed in the same way used when dealing with the gender of the head of household and interact the education dummy with the household's electrification status. We also interact it with our instrument in order to obtain a second instrument. Results are shown in Table 13. As expected, a lower education level further increases adoption of fuelwood and decreases the likelihood of picking up LPG. A higher level of education is likely to be linked with a better understanding of the dangers of indoor air pollution, while households with a lower level of education are more likely to be more oblivious to the dangers of it and use electricity for other purposes. The effect becomes noticeable even for households above the mean level of monthly expenditures.

[Table 13 here]

Budget reaction

Finally, if a new connection to the power grid poses an extra constraint on households' budget, we should be able to pick this up. In Table 14, we regress the share of monthly energy expenditures out of a household's total monthly expenditure 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 mean monthly per capita expenditures, 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 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 negative and statistically insignificant. Therefore, it seems that the new connection does not put any additional financial stress on richer households. Actually, if we focus on the CLR confidence set, the negative coefficient seem to be statistically significant at the 5% level.

[Table 14 here]

7 Conclusion

The goal of this paper is to identify a causal relationship between access to electricity and the adoption of modern cooking fuels, such as LPG. Access to a new electricity connection may impact households in two different ways. First, electricity opens up a whole new set of consumption opportunities, both productive – such as irrigation pumps – and non-productive – such as entertainment. Faced with these new consumption opportunities, a household may decide to divert to them some of the money that previously was devoted to energy, given the availability of free choices for energy (fuelwood, dung, crop residues). Second, according to the existing literature (e.g. International Evaluation Group, 2008; Khandker et al., 2012; Lipscomb et al., 2013; Chakravorty et al., 2014, 2016; Van de Walle et al., 2017) a connection to the grid contributes to increase rural households' incomes. This increase in socio-economic status, according to the energy ladder theory should push households towards the adoption of more efficient and less polluting fuels, such as kerosene or LPG. 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. This causal relationship would constitute an additional channel through which rural electrification contributes to an improvement in households' welfare.

Our results show that the first possibility is the one at play. We observe that, while electrification has no impact on the choice of a cooking fuel for richer households, its impact is consistent with what one would expect for poorer households when offering them a wider choice set. Poorer households (below mean expenditure), when faced with the possibility of having access to electric light, will take it, since electricity allows them to have good quality lightning. Yet, these households run on a tight budget constraint. As a consequence, in order to be able to afford electricity, they cut back on other energy related expenses, which leads them to revert from modern cooking fuels, such as LPG, to freely available biomass. As underlined earlier, we should remember that the estimates produced in this paper are LATE and, therefore, can be directly applied only to the population of compliers.

In order to improve our understanding of the results obtained, we should ideally investigate how these coefficients are affected by the quality of power supply. Presently, in rural India, electricity is available on average for only about 12 and a half to 13 hours in a day (Chakravorty et al., 2014; Aklin et al., 2016). Chakravorty et al. (2014) found that while a good quality connection may increase a household income by up to 27%, a low quality connection will not do better then 9%. The bad quality of power supply could have two consequences. First, incomes may not increase sufficiently enough to compensate for the new bill that needs to be paid and enable the financially constrained households to afford LPG. Second, because of the low quality of the power supply, households may not invest in many electrical appliances, i.e. their choice set would actually not be affected. This lack of appliances together with the fact that they only receive power for less than 13 hours per day may result in small expenditures for electricity and, therefore, we may not observe the change in cooking fuel choice for poor households receiving a bad quality of power supply.

Unfortunately, NSS does not contain information about the quality of power supply and, therefore, we are unable to further investigate this channel.

Indoor pollution generated by the burning of biomass is a major health concern in India, between 400,000 and 500,000 people die of household air pollution every year. Pollution from dirty biomass fuels used for cooking contributes to 12% of still birth. Moreover, this is also a gender issue, since women are the primary cooks in many Indian households and household air pollution leads to a range of diseases, every year we observe: over 2.4 million cases of chronic bronchitis, over 300,000 cases of tuberculosis and over 5 million cases of cataract.[32] For these reasons the government of India is trying to push LPG adoption through the Ujjwala scheme.[33] Introduced in 2016, this scheme provides a subsidy to government-owned oil manufacturing companies for each LPG gas connection that they install in poor rural households without one. Yet, this program only focuses on the sunk costs related to the uptake of LPG, what may be inferred from our results is that the running costs, e.g. the costs of refilling the LPG cylinder, may still be a hurdle for the poorer households. Subsidies on LPG cylinders are currently low and decreasing in India.[34]

Our findings align with Lee et al. (2020a) along two dimensions. First, they imply that, for the poor, access to electricity makes the switch to cleaner cooking fuels even harder. Second, one of the main take away of the paper, is that the introduction of electricity should always be accompanied by complementary policies. In this specific case, they could consist of subsidies for the adoption of cleaner cooking fuels, in order to avoid a worsening of indoor pollution for the poorest households.

[32] "How withdrawal of LPG subsidy can raise India's healthcare costs," *Business Standard*, August 3rd 2017.

[33] `http://www.pmujjwalayojana.com`

[34] "Govt orders LPG prices to be hiked by Rs 4 per month", *The Times of India*, July 31st 2017; Singh, S., "How withdrawal of LPG subsidy can raise India's healthcare costs," *Business Standard*, August 3rd 2017.

References

Agrawal, S., Bali, N., and Urpelainen, J. (2019). Rural Electrification in India: Customer Behaviour and Demand. Report, Smart Power India & Initiative for Sustainable Energy Policy.

Aklin, M., Cheng, C.-y., Urpelainen, J., Ganesan, K., and Jain, A. (2016). Factors affecting household satisfaction with electricity supply in rural india. *Nature Energy*, 1(11):16170.

Alkon, M., Harish, S., and Urpelainen, J. (2016). Household Energy Access and Expenditure in Developing Countries: Evidence from India, 1987–2010. *Energy for Sustainable Development*, 35:25 – 34.

Allcott, H., Collard-Wexler, A., and 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.

Balachandra, P. (2011). Dynamics of Rural Energy Access in India: An Assessment. *Energy*, 36(9):5556–5567.

Barron, M. and Torero, M. (2017). Household Electrification and Indoor Air Pollution. *Journal of Environmental Economics and Management*, 86:81–92.

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.

Bensch, G., Kluve, J., and Peters, J. (2011). Impacts of Rural Electrification in Rwanda. *Journal of Development Effectiveness*, 3(4):567–588.

Bhojvaid, V., Jeuland, M., Kar, A., Lewis, J., Pattanayak, S., Ramanathan, N., Ramanathan, V., and Rehman, I. (2014). How do People in Rural India Perceive Improved Stoves and Clean Fuel? Evidence from Uttar Pradesh and Uttarakhand. *International Journal of Environmental Research and Public Health*, 11(2):1341–1358.

Bond, T., Venkataraman, C., and Masera, O. (2004). Global Atmospheric Impacts of Residential Fuels. *Energy for Sustainable Development*, 8(3):20 – 32.

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.

Boskovic, B., Chakravorty, U., Pelli, M., and Risch, A. (2018). The Effect of Forest Access on the Market for Fuelwood in India. Working Paper 05-2018, CIREQ.

Bridge, B., Adhikari, D., and Fontenla, M. (2016). Electricity, Income, and Quality of Life. *The Social Science Journal*, 53(1):33–39.

Bruce, N., Aunan, K., and Rehfuess, E. (2017). Liquified Petroleum Gas as a Clean Cooking Fuel for Developing Countries: Implications for Climate, Forests, and Affordability. Report, KfW Development Bank: Materials on Development Financing.

Bruce, N., Perez-Padilla, R., and 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.

Burlig, F. and Preonas, L. (2016). Out of the Darkness and Into the Light? Development Effects of Rural Electrification. Working Paper 268, Energy Institute at HAAS.

Central Electricity Authority (2006). *Review of Performance of Hydro Power Stations 2005-2006*. Government of India, Ministry of Power.

Chakravorty, U., Emerick, K., and Ravago, M.-L. (2016). 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.

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

Chan, H., Cropper, M., and Malik, K. (2014). Why Are Power Plants in India Less Efficient Than Power Plants in the United States? *American Economic Review*, 104(5):586–590.

Cheng, C. and Urpelainen, J. (2014). Fuel Stacking in India: Changes in the Cooking and Lighting Mix, 1987–2010. *Energy*, 76:306 – 317.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., et al. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. *The Lancet*, 389(10082):1907–1918.

Dherani, M., Pope, D., Mascarenhas, M., Smith, K., Weber, M., and 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. (2003). Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *World Bank Economic Review*, 17(1):1–25.

Duflo, E., Greenstone, M., and Hanna, R. (2008). Indoor Air Pollution, Health and Economic Well-Being. *S.A.P.I.EN.S. [Online]*, 1(1).

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

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

Ghilardi, A., Guerrero, G., and Masera, O. (2009). A GIS-Based Methodology for Highlighting Fuelwood Supply/Demand Imbalances at the Local Level: A Case Study for Central Mexico. *Biomass and Bioenergy*, 33(6):957 – 972.

Global Energy Observatory (2019). Global power plant database v1.2.0. Database, Google, KTH Royal Institute of Technology in Stockholm, Enipedia, World Resources Institute.

Gould, C. F. and Urpelainen, J. (2018). Lpg as a clean cooking fuel: Adoption, use, and impact in rural india. *Energy Policy*, 122:395–408.

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

Grogan, L. and Sadanand, A. (2013). Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua. *World Development*, 43(C):252–265.

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

Hollada, J., Williams, K., Miele, C., Danz, D. aand Harvey, S., and Checkley, W. (2017). Perceptions of Improved Biomass and Liquefied Petroleum Gas Stoves in Puno, Peru: Implications for Promoting Sustained and Exclusive Adoption of Clean Cooking Technologies. *International Journal of Environmental Research and Public Health*, 14(2):182.

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

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

International Evaluation Group (2008). *The welfare impact of rural electrification: A reassessment of the costs and benefits*. The World Bank.

International Hydropower Association (2020). 2020 Hydropower Status Report: Sector trends and insights. Report.

Jeuland, M. and Pattanayak, S. (2012). Benefits and Costs of Improved Cookstoves: Assessing the Implications of Variability in Health, Forest and Climate Impacts. *Plos One*, 7(2):1–15.

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

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

Kishore, A. and 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.

Kowsari, R. and Zerriffi, H. (2011). Three Dimensional Energy Profile: A Conceptual Framework for Assessing Household Energy Use. *Energy Policy*, 39(12):7505 – 7517. Clean Cooking Fuels and Technologies in Developing Economies.

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

Lee, K., Miguel, E., and Wolfram, C. (2020a). Does Household Electrification Supercharge Economic Development? *Journal of Economic Perspectives*, 34(1):122–44.

Lee, K., Miguel, E., and Wolfram, C. (2020b). Experimental Evidence on the Economics of Rural Electrification. *Journal of Political Economy*, 128(4):1523–1565.

Lenz, L., Munyehirwe, A., Peters, J., and Sievert, M. (2017). Does Large-Scale Infrastructure Investment Alleviate Poverty? Impacts of Rwanda's Electricity Access Roll-Out Program. *World Development*, 89:88–110.

Lewis, J. and Pattanayak, S. (2012). Who Adopts Improved Fuels and Cookstoves? A Systematic Review. *Environmental Health Perspectives*, 120(5):637.

Lim, S., Vos, T., Flaxman, A., and al. (2012). A Comparative Risk Assessment of Burden of Disease and Injury Attributable to 67 Risk Factors and Risk Factor Clusters in 21 Regions, 1990–2010: A Systematic Analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859):2224 – 2260.

Lipscomb, M., Mobarak, A., and 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., Ghilardi, A., Drigo, R., and Trossero, M. (2006). WISDOM: A GIS-Based Supply Demand Mapping Tool for Woodfuel Management. *Biomass and Bioenergy*, 30(7):618 – 637.

Masera, O., Saatkamp, B., and 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.

Mikusheva, A. (2010). Robust Confidence Sets in the Presence of Weak Instruments. *Journal of Econometrics*, 157:236–247.

Modi, V. (2005). Improving Electricity Services in Rural India. Working Paper 30, Center on Globalization and Sustainable Development. The Earth Institute at Columbia University.

Montiel Olea, J. and Pflueger, C. (2013). A Robust Test for Weak Instruments. *Journal of Business and Economic Statistics*, 31(3):358–369.

Mukhopadhyay, R., Sambandam, S., Pillarisetti, A., Jack, D., Mukhopadhyay, K., Balakrishnan, K., Vaswani, M., Bates, M., Kinney, P., Arora, N., and Smith, K. (2012). Cooking Practices, Air Quality, and the Acceptability of Advanced Cookstoves in Haryana, India: An Exploratory Study to Inform Large-Scale Interventions. *Global Health Action*, 5(1):19016. PMID: 28140858.

Oda, H. and Tsujita, Y. (2011). The Determinants of Rural Electrification: The Case of Bihar, India. *Energy Policy*, 39(6):3086–3095.

Pant, P., Lal, R. M., Guttikunda, S. K., Russell, A. G., Nagpure, A. S., Ramaswami, A., and Peltier, R. E. (2019). Monitoring particulate matter in india: Recent trends and future outlook. *Air Quality, Atmosphere & Health*, 12(1):45–58.

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

Po, J., Fitzgerald, J., and 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.

Puzzolo, E., Pope, D., Stanistreet, D., Rehfuess, E., and Bruce, N. (2016). Clean Fuels for Resource-Poor Settings: A Systematic Review of Barriers and Enablers to Adoption and Sustained Use. *Environmental Research*, 146:218 – 234.

Qian, N. (2008). Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance. *The Quarterly Journal of Economics*, 123(3):1251–1285.

Rao, N. (2013). Does (Better) Electricity Supply Increase Household Enterprise Income in India? *Energy Policy*, 57:532 – 541.

Rehfuess, E., Puzzolo, E., Stanistreet, D., Pope, D., and Bruce, N. (2014). Enablers and Barriers to Large-Scale Uptake of Improved Solid Fuel Stoves: A Systematic Review. *Environmental Health Perspectives*, 122(2):120.

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

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

Ryan, N. and Sudarshan, A. (2020). Rationing the Commons. Working Paper 27473, NBER.

Sehjpal, R., Ramji, A., Soni, A., and Kumar, A. (2014). Going Beyond Incomes: Dimensions of Cooking Energy Transitions in Rural India. *Energy*, 68:470 – 477.

Simon, G., Bailis, R., Baumgartner, J., Hyman, J., and Laurent, A. (2014). Current Debates and Future Research Needs in the Clean Cookstove Sector. *Energy for Sustainable Development*, 20:49 – 57.

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.

Troncoso, K. and Soares da Silva, A. (2017). LPG Fuel Subsidies in Latin America and the Use of Solid Fuels to Cook. *Energy Policy*, 107:188 – 196.

Van de Walle, D., Ravallion, M., Mendiratta, V., and 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., and 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.

Zacharias, A. and Vakulabharanam, V. (2011). Caste stratification and wealth inequality in india. *World Development*, 39(10):1820–1833.

8 Tables

Table 1: Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
<u>*Panel A:*</u>				
Share of electricity	0.68	0.47	0	1
Share of LPG‡	0.16	0.37	0	1
Share of fuelwood‡	0.73	0.44	0	1
LPG share of expenditures (%)	0.94	2.16	0	27.82
Fuelwood share of expenditures (%)	5.29	4.59	0	64.88
<u>*Panel B:*</u>				
Nightlights 1993	2.53	2.85	0	55.44
Hydro Instrument	0.13	0.19	0	0.93
Interaction Hydro Instrument	0.32	0.60	0	4.91
<u>*Panel C:*</u>				
Price Fuelwood (Rs./Kg)	1.44	1.76	0.05	112.97
Price LPG (Rs./ Kg)	20.27	9.99	8.36	215.83
Household Size	5.00	2.46	1	30.00
Total land owned (Hectares)	1.01	2.42	0	310.31
Education head of household (Years)	4.39	3.10	1	13.00
Age head of household	46.44	13.29	14	100
Gender head of household‡	0.89	0.31	0	1
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‡	0.38	0.49	0	1
Below Primary‡	0.50	0.50	0	1

Notes: Our sample contains 136,221 households, 78,017 interviewed in 2004 and 58,204 interviewed in 2009. *Panel A* reports statistics for the variables of interest and the dependent variables, *Panel B* for the instruments used and, *Panel C* for all the control variables used in the different specifications. The share of expenditures is computed over the last 30 days. *Sex head of household* equals 1 if the head of the household is a male and 0 otherwise. *Hindu*, *Scheduled tribe*, *Scheduled caste*, *Other backward caste*, and *Below Primary* equal 1 if the household belongs to the group and 0 otherwise. ‡ denotes indicator variables.

Table 2: First stage

	Dep. variable: electricity connection			
	All		Below	Above
	(1)	(2)	(3)	(4)
Hydro	−0.301**	−0.274**	−0.323**	0.022
	(0.133)	(0.123)	(0.147)	(0.137)
*Hydro*Initial elec*	0.140***	0.118***	0.123***	0.045**
	(0.026)	(0.025)	(0.034)	(0.023)
Price Fuelwood		0.002	0.002	−0.001
		(0.001)	(0.002)	(0.002)
Price LPG		−0.000	−0.000	−0.001
		(0.001)	(0.001)	(0.001)
HH size		0.010***	0.016***	0.011***
		(0.001)	(0.001)	(0.001)
Total land owned		0.004***	0.004***	0.001
		(0.001)	(0.001)	(0.001)
Education head of HH		0.028***	0.028***	0.015***
		(0.001)	(0.001)	(0.001)
Sex of head of HH		−0.019***	−0.012**	−0.021***
		(0.005)	(0.006)	(0.005)
Age head of HH		0.002***	0.002***	0.001***
		(0.0001)	(0.0001)	(0.0001)
Hindu		0.010	0.010	0.002
		(0.007)	(0.007)	(0.007)
Scheduled tribe		−0.143***	−0.144***	−0.080***
		(0.013)	(0.013)	(0.014)
Scheduled caste		−0.074***	−0.071***	−0.046***
		(0.006)	(0.006)	(0.008)
Other backward caste		−0.027***	−0.028***	−0.020***
		(0.005)	(0.006)	(0.005)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Observations	136,221	136,154	92,773	43,381
F-stat	16.68	12.08	6.76	8.03
Montiel-Pflueger F-stat			4.86	2.77

Notes: *All* includes the whole sample, *Below* is the sub-sample of households below the mean of monthly household consumption expenditure of 724 Rs. (roughly 10 USD) and *Above* is the sub-sample of households below the mean of monthly household consumption expenditure of 724 Rs. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Baseline – extensive margin

	Dep. variable: adoption			
	LPG		Fuelwood	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Electricity	0.081***	−0.566**	−0.074***	0.520*
	(0.005)	(0.231)	(0.007)	(0.308)
Price Fuelwood	0.003**	0.004**	−0.005***	−0.006***
	(0.001)	(0.002)	(0.002)	(0.002)
Price LPG	−0.003***	−0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
HH size	−0.003***	0.004	0.006***	−0.001
	(0.001)	(0.002)	(0.001)	(0.003)
Total land owned	0.001	0.004***	−0.000	−0.003*
	(0.001)	(0.001)	(0.001)	(0.002)
Education head of HH	0.037***	0.056***	−0.036***	−0.053***
	(0.001)	(0.007)	(0.001)	(0.009)
Sex of head of HH	−0.047***	−0.060***	0.044***	0.055***
	(0.004)	(0.007)	(0.004)	(0.008)
Age head of HH	0.002***	0.004***	−0.002***	−0.003***
	(0.000)	(0.001)	(0.000)	(0.001)
Hindu	0.005	0.011	−0.004	−0.010
	(0.006)	(0.008)	(0.008)	(0.009)
Scheduled tribe	−0.075***	−0.168***	0.098***	0.183***
	(0.008)	(0.035)	(0.009)	(0.046)
Scheduled caste	−0.078***	−0.126***	0.075***	0.119***
	(0.005)	(0.012)	(0.006)	(0.022)
Other backward caste	−0.056***	−0.074***	0.045***	0.062***
	(0.005)	(0.009)	(0.006)	(0.011)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Observations	136,154	136,154	136,154	136,154
F-stat		12.08		12.08

Notes: The dependent variable is a dummy taking value 1 if LPG (fuelwood) is adopted as the primary cooking fuel by a household and 0 otherwise. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Baseline – below and above mean expenditures

	Below				Above			
	LPG		Fuelwood		LPG		Fuelwood	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: extensive								
Electricity	0.046***	−0.526**	−0.040***	0.758**	0.163***	−0.091	−0.148***	−0.148
	(0.003)	(0.229)	(0.006)	(0.384)	(0.009)	(0.600)	(0.011)	(0.592)
Observations	92,773	92,772	92,773	92,772	43,381	43,381	43,381	43,381
LIML estimate		−1.260		2.467		−8.095		−240.7
CLR 95% conf. set		$(-\infty; -0.532]$		$[0.913; +\infty)$		$(-\infty; +\infty)$		$(-\infty; +\infty)$
F-stat		6.72		6.72		7.93		7.93
Montiel-Pflueger F-stat		4.86		4.86		2.77		2.77
Panel B: intensive								
Electricity	0.409***	−1.512	−1.080***	12.250**	1.027***	3.017	−1.260***	−1.122
	(0.028)	(1.167)	(0.068)	(6.032)	(0.052)	(3.088)	(0.117)	(4.006)
Observations	92,591	92,590	92,469	92,468	43,200	43,200	43,259	43,259
LIML estimate		−6.954		17.301		5.160		−0.995
CLR 95% conf. set		$(-\infty; -2.313]$		$[-1.795; +\infty)$		$[-5.603; 19.235]$		$[-17.165; 11.515]$
F-stat		7.14		6.53		7.90		7.88
Montiel-Pflueger F-stat		5.04		4.82		2.77		2.78
Year F.E.	yes	yes	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes	yes	yes

Notes: *Below* denotes the sub-sample of households below the mean of monthly household consumption expenditures of 724 Rs. (roughly 10USD) while *Above* denotes the sub-sample of households above the mean of monthly household consumption expenditures of 724 Rs. In *Panel A*, the dependent variable is a dummy taking value 1 if LPG (fuelwood) is adopted as the primary cooking fuel by a household and 0 otherwise. In *Panel B*, the dependent variable is the share of expenditure in LPG (fuelwood) over total monthly expenditures. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: Assets

	Dep. variable: adoption		
	TV	Radio	Fan
	(1)	(2)	(3)
Electricity	0.633**	0.240	0.793***
	(0.253)	(0.292)	(0.241)
Controls	yes	yes	yes
Year F.E.	yes	yes	yes
District F.E.	yes	yes	yes
Observations	124,618	125,192	126,432
F-stat	12.44	13	11.81

Notes: This table reports only instrumental variable estimations for the adoption of TVs, radios and electric fans. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Gender

	LPG			Fuelwood		
	All	Below	Above	All	Below	Above
	(1)	(2)	(3)	(4)	(5)	(6)
<u>*Panel A: Extensive*</u>						
Electricity	-0.493^{**}	-0.491^{**}	0.009	0.469	0.724^{*}	-0.207
	(0.225)	(0.227)	(0.583)	(0.299)	(0.380)	(0.577)
*Gender*Elec*	-0.086^{***}	-0.041^{***}	-0.109^{***}	0.062^{***}	0.041^{**}	0.059^{***}
	(0.012)	(0.016)	(0.026)	(0.014)	(0.019)	(0.022)
Observations	136,154	92,772	43,381	136,154	92,772	43,381
F-stat Electricity	14.61	4.77	13.84	14.61	4.77	13.84
F-stat Gender*Elec	51.09	34.11	56.05	51.09	34.11	56.05
<u>*Panel B: Intensive*</u>						
Electricity	0.508	-1.183	3.451	6.520^{*}	12.348^{**}	-1.308
	(1.195)	(1.154)	(2.995)	(3.550)	(5.988)	(3.892)
*Gender*Elec*	-0.487^{***}	-0.389^{***}	-0.467^{***}	0.354^{**}	-0.121	0.201
	(0.076)	(0.107)	(0.138)	(0.169)	(0.244)	(0.161)
Observations	135,791	92,590	43,200	135,728	92,468	43,259
F-stat Electricity	15.22	14.84	5.16	13.95	14.49	4.64
F-stat Gender*Elec	50.99	34.10	56.07	40.72	50.82	56.01
Controls	yes	yes	yes	yes	yes	yes
Year F.E.	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes

Notes: This table reports only instrumental variable estimations. The *Gender* variable is a dummy for the gender of the head of the household. It takes value 1 for male and 0 otherwise. *All* includes the whole sample, *Below* is the sub-sample of households below the mean of monthly household consumption expenditure of 724 Rs. (roughly 10 USD) and *Above* is the sub-sample of households above the mean of monthly household consumption expenditure of 724 Rs. In *Panel A*, the dependent variable is a dummy taking value 1 if LPG (fuelwood) is adopted as the primary cooking fuel by a household and 0 otherwise. In *Panel B*, the dependent variable is the share of expenditure in LPG (fuelwood) over total monthly expenditures. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Baseline – intensive margin

	Dep. variable: income share			
	LPG		Fuelwood	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Electricity	0.614***	0.094	−1.451***	6.835*
	(0.030)	(1.230)	(0.072)	(3.652)
Price Fuelwood	0.017**	0.018**	0.100**	0.085**
	(0.007)	(0.008)	(0.042)	(0.042)
Price LPG	−0.018***	−0.018***	0.003	0.006
	(0.004)	(0.004)	(0.007)	(0.009)
HH size	−0.038***	−0.033**	−0.187***	−0.274***
	(0.003)	(0.013)	(0.009)	(0.041)
Total land owned	−0.008**	−0.006	−0.060***	−0.095***
	(0.003)	(0.006)	(0.013)	(0.024)
Education head of HH	0.179***	0.193***	−0.338***	−0.572***
	(0.005)	(0.035)	(0.010)	(0.105)
Sex of head of HH	−0.304***	−0.314***	−0.230***	−0.066
	(0.027)	(0.038)	(0.061)	(0.111)
Age head of HH	0.012***	0.013***	−0.012***	−0.030***
	(0.001)	(0.003)	(0.001)	(0.008)
Hindu	0.105***	0.110***	−0.070	−0.156
	(0.033)	(0.037)	(0.073)	(0.105)
Scheduled tribe	−0.458***	−0.533***	1.115***	2.303***
	(0.049)	(0.184)	(0.117)	(0.557)
Scheduled caste	−0.457***	−0.496***	0.877***	1.494***
	(0.028)	(0.095)	(0.071)	(0.275)
Other backward caste	−0.303***	−0.317***	0.345***	0.575***
	(0.025)	(0.042)	(0.052)	(0.114)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Observations	135,791	135,791	135,728	135,728
F-stat		12.30		11.87

Notes: The dependent variable is the share of expenditure in LPG (fuelwood) over total monthly expenditures. The sample is smaller than for the extensive margin because 420 households report using LPG, yet, do not report the amount of their expenditures. Moreover, some households (24) report their LPG expenditures but do not use LPG as the primary cooking fuel. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Reduced form

	Dep. variable: adoption	
	LPG	**Fuelwood**
	(1)	(2)
Hydro	0.265	−0.234
	(0.268)	(0.309)
*Hydro*Initial elec*	−0.062	0.032
	(0.043)	(0.055)
Price Fuelwood	0.004**	−0.011***
	(0.002)	(0.004)
Price LPG	−0.005***	0.005***
	(0.001)	(0.001)
HH size	−0.003**	0.004***
	(0.001)	(0.001)
Total land owned	0.000	0.001
	(0.001)	(0.001)
Education head of HH	0.041***	−0.038***
	(0.002)	(0.002)
Sex of head of HH	−0.070***	0.057***
	(0.008)	(0.009)
Age head of HH	0.003***	−0.002***
	(0.000)	(0.000)
Hindu	−0.005	−0.017
	(0.011)	(0.014)
Scheduled tribe	−0.081***	0.108***
	(0.013)	(0.015)
Scheduled caste	−0.086***	0.090***
	(0.009)	(0.011)
Other backward caste	−0.063***	0.060***
	(0.009)	(0.011)
Year F.E.	yes	yes
District F.E.	yes	yes
Observations	41,486	41,486

Notes: The dependent variable is a dummy taking value 1 if LPG (fuelwood) is adopted as the primary cooking fuel by a household and 0 otherwise. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 9: Baseline specifications excluding information collected during the monsoon months

	Below				Above			
	LPG		Fuelwood		LPG		Fuelwood	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: extensive								
Electricity	0.045***	−0.566**	−0.041***	0.869**	0.169***	−0.242	−0.152***	−0.155
	(0.003)	(0.277)	(0.006)	(0.440)	(0.010)	(0.717)	(0.012)	(0.673)
Observations	74,188	74,187	74,188	74,187	35,083	35,083	35,083	35,083
LIML estimate		−1.377		2.507		−10.069		7.195
CLR 95% conf. set		$(-\infty; -0.508]$		$[0.807; +\infty)$		$(-\infty; +\infty)$		$(-\infty; +\infty)$
F-stat		5		5		5.90		5.90
Montiel-Pflueger F-stat		4.02		4.02		2.11		2.11
Panel B: intensive								
Electricity	0.408***	−1.644	−1.070***	13.915*	1.080***	3.125	−1.300***	0.058
	(0.029)	(1.511)	(0.071)	(7.243)	(0.059)	(3.672)	(0.142)	(4.611)
Observations	74,035	74,034	73,035	73,034	34,923	34,923	34.983	34,983
LIML estimate		−6.996		19.773		7.516		3.057
CLR 95% conf. set		$(-\infty; -2.156]$		$[0.360; +\infty)$		$[-11.561; 32.044]$		$[-21.704; 20.570]$
F-stat		5.36		4.83		5.86		4.83
Montiel-Pflueger F-stat		4.22		3.94		2.11		2.13
Year F.E.	yes	yes	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes	yes	yes

Notes: The sample used in this table excludes households which have been interviewed about their consumption during the monsoon months (July-September). Since NSS asks respondents about their consumption over the last 30 days, we exclude households interviewed between July 31st and October 7th. *Below* denotes the sub-sample of households below the mean of monthly household consumption expenditures of 724 Rs. (roughly 10USD) while *Above* denotes the sub-sample of households above the mean of monthly household consumption expenditures of 724 Rs. In *Panel A*, the dependent variable is a dummy taking value 1 if LPG (fuelwood) is adopted as the primary cooking fuel by a household and 0 otherwise. In *Panel B*, the dependent variable is the share of expenditure in LPG (fuelwood) over total monthly expenditures. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Baseline – interaction

	LPG		Fuelwood	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
<u>*Panel A: extensive*</u>				
Electricity	0.239***	−0.266	−0.236***	0.256
	(0.007)	(0.207)	(0.008)	(0.278)
*Expenditure*Elec*	−0.221***	−0.252***	0.226***	0.224**
	(0.005)	(0.015)	(0.006)	(0.024)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Observations	136,154	136,154	136,154	136,154
F-stat Electricity		14.24		14.24
F-stat Expenditure*Elec		73.26		73.26
<u>*Panel B: intensive*</u>				
Electricity	1.185***	0.853	−3.305***	3.150
	(0.037)	(1.142)	(0.090)	(3.258)
*Expenditure*Elec*	−0.794***	−0.658***	2.606***	3.083*
	(0.031)	(0.076)	(0.068)	(0.285)
Year F.E.	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes
Observations	135,791	135,791	135,728	135,728
F-stat Electricity		14.33		14.13
F-stat Expenditure*Elec		73.32		73.01

Notes: The variable *Expenditure* is a dummy taking value 1 if a household's expenditures are below the mean level of expenditures (724 Rs.) and 0 otherwise. In *Panel A*, the dependent variable is a dummy taking value 1 if LPG (fuelwood) is adopted as the primary cooking fuel by a household and 0 otherwise. In *Panel B*, the dependent variable is the share of expenditure in LPG (fuelwood) over total monthly expenditures. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Caste – extensive margin

	Dep. variable: adoption				
	ST	SC	OBC	All Backward	Other
	(1)	(2)	(3)	(4)	(5)
Panel A: LPG					
Electricity	-0.855^{*}	0.199	-0.597^{*}	-0.495^{**}	-0.235
	(0.470)	(0.274)	(0.315)	(0.227)	(0.395)
Observations	21,811	24,253	52,270	98,419	37,720
LIML estimate	-0.862	0.930	-1.090	-1.247	-1.984
CLR 95% conf. set	$(-\infty; -0.201)$	$[0.832; \infty)$	$(-\infty; -0.361]$	$(-\infty; -0.679)$	$[-2.999; 1.643]$
F-stat	2.31	9.93	4.23	7.81	8.64
Montiel-Pflueger F-stat	2.13	7.34	5.28	7.11	5.30
Panel B: Fuelwood					
Electricity	0.724	-0.568	1.572^{***}	0.581^{*}	-0.339
	(0.492)	(0.445)	(0.610)	(0.321)	(0.409)
Observations	21,811	24,253	52,270	98,419	37,720
LIML estimate	0.818	-2.530	2.370	2.103	-1.240
CLR 95% conf. set	$(-\infty; +\infty)$	$(-\infty; -0.860]$	$[1.135; \infty)$	$[0.843; \infty)$	$[-2.006; 2.369]$
F-stat	2.31	9.93	4.23	7.81	8.64
Montiel-Pflueger F-stat	2.13	7.37	5.28	7.11	5.30
Controls	yes	yes	yes	yes	yes
Year F.E.	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes

Notes: This table reports only instrumental variable estimations. *Panel A (B)* reports results for LPG (fuelwood) adoption and, therefore, the dependent variable is a dummy taking value 1 if LPG (fuelwood) is adopted as the primary cooking fuel by a household and 0 otherwise. *ST*, *SC* and *OBC* represent scheduled tribe, scheduled caste and other backward caste, respectively. *All Backward* is all the aforementioned castes combined into one category and *Other* includes all upper castes. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 12: Caste – intensive margin

	Dep. variable: income share				
	ST	SC	OBC	All Backward	Other
	(1)	(2)	(3)	(4)	(5)
<u>*Panel A: LPG*</u>					
Electricity	−3.373	2.314	−1.646	−1.528	4.359*
	(2.422)	(1.455)	(1.582)	(1.123)	(2.510)
Observations	21,707	24,216	52,152	98,160	37,618
LIML estimate	−3.440	4.111	−2.567	−3.660	5.423
CLR 95% conf. set	$(-\infty; 0.809]$	$[3.184; +\infty)$	$(-\infty; 0.137]$	$(-\infty; -1.797)$	$[-0.827; 11.396]$
F-stat	2.25	10.63	4.16	8.09	8.67
Montiel-Pflueger F-stat	2.08	7.77	5.22	7.25	5.26
<u>*Panel B: Fuelwood*</u>					
Electricity	5.086	−0.796	16.875**	7.975*	−3.589
	(6.422)	(3.449)	(8.067)	(4.669)	(3.611)
Observations	21,727	24,184	52,117	98,113	37,600
LIML estimate	8.044	209.311	21.279	11.156	−8.276
CLR 95% conf. set	$(-\infty; +\infty)$	$(-\infty; +\infty)$	$[2.701; +\infty)$	$[-8.018; +\infty)$	$[-7.719; 9.000]$
F-stat	2.34	9.80	4.20	7.77	8.26
Montiel-Pflueger F-stat	2.16	7.32	5.30	7.15	5.17
Controls	yes	yes	yes	yes	yes
Year F.E.	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes

Notes: This table reports only instrumental variable estimations. *Panel A (B)* reports results for the intensive margin of LPG (fuelwood) and, therefore, the dependent variable is the share of expenditure in LPG (fuelwood) over total monthly expenditures. *ST*, *SC* and *OBC* represent scheduled tribe, scheduled caste and other backward caste, respectively. *All Backward* is all the aforementioned castes combined into one category and *Other* includes all upper castes. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Schooling

	Dep. variable: adoption and income share					
	LPG			Fuelwood		
	All	Below	Above	All	Below	Above
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Extensive						
Electricity	−0.632**	−0.559**	0.079	0.533*	0.753*	−0.301
	(0.253)	(0.249)	(0.557)	(0.317)	(0.388)	(0.560)
Below Primary	−0.285***	−0.163***	−0.254***	0.244***	0.179***	0.182***
	(0.024)	(0.026)	(0.027)	(0.031)	(0.040)	(0.028)
Observations	136,183	92,792	43,390	136,183	92,792	43,390
F-stat Electricity	18.53	14.67	15.65	18.53	14.67	15.65
F-stat Below Primary*Elec	40.74	27.10	64.56	40.74	27.10	64.56
Panel B: Intensive						
Electricity	−0.135	−1.813	3.548	6.852*	12.652**	−2.111
	(1.303)	(1.334)	(2.910)	(3.594)	(6.099)	(3.711)
Below Primary	−1.084***	−1.022***	−0.861***	2.304***	2.361***	0.958***
	(0.124)	(0.152)	(0.137)	(0.356)	(0.637)	(0.187)
Observations	135,820	92,610	43,209	135,755	92,486	43,268
F-stat Electricity	18.60	14.84	15.68	18.60	14.92	15.71
F-stat Below Primary*Elec	40.75	27.19	64.63	40.72	27.06	64.73
Controls	yes	yes	yes	yes	yes	yes
Year F.E.	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes

Notes: This table reports only instrumental variable estimations. *All* includes the whole sample, *Below* is the sub-sample of households below the mean of monthly household consumption expenditure of 724 Rs. (roughly 10USD) and *Above* is the sub-sample of households above the mean of monthly household consumption expenditure of 724 Rs. The variable *Below Primary* is a dummy coded 1 if the head of the household complete at most primary education and 0 otherwise. In *Panel A*, the dependent variable is a dummy taking value 1 if LPG (fuelwood) is adopted as the primary cooking fuel by a household and 0 otherwise. In *Panel B*, the dependent variable is the share of expenditure in LPG (fuelwood) over total monthly expenditures. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

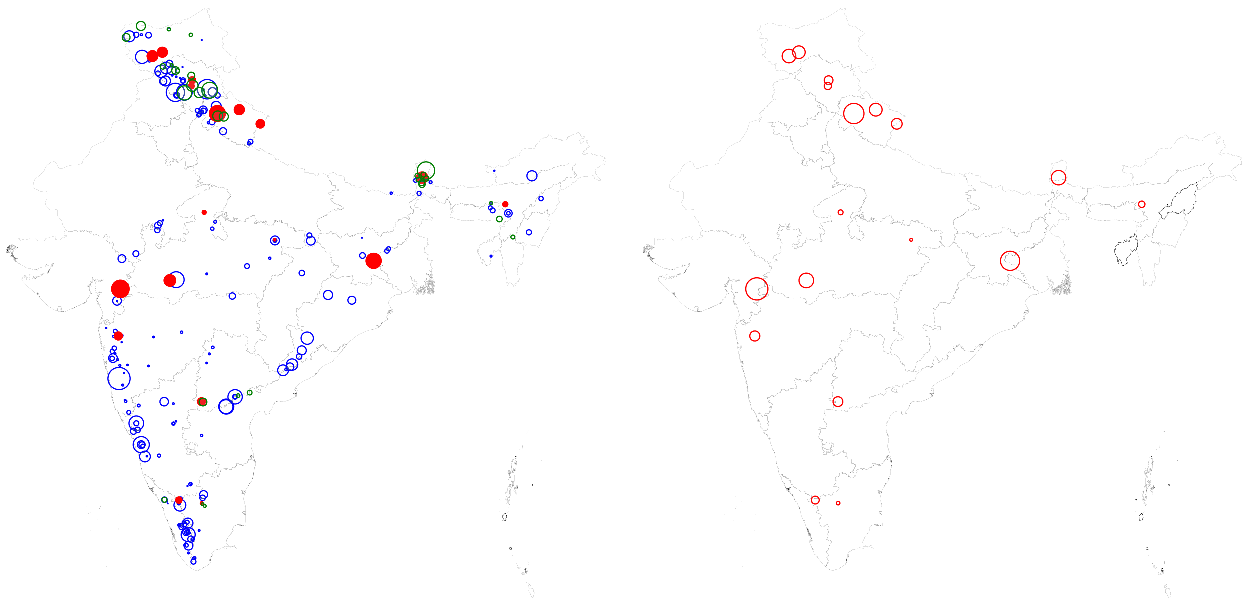
Table 14: Share of fuel expenditures

	Dep. variable: share of fuel expenditures					
	All		**Below**		**Above**	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Electricity	1.436***	17.631***	2.053***	22.676***	1.713***	−13.048
	(0.093)	(5.389)	(0.096)	(8.263)	(0.165)	(9.103)
Price Fuelwood	0.126**	0.096	0.104	0.058	0.194***	0.179**
	(0.058)	(0.061)	(0.068)	(0.078)	(0.066)	(0.076)
Price LPG	−0.015*	−0.010	−0.019**	−0.015	0.001	−0.013
	(0.008)	(0.013)	(0.009)	(0.017)	(0.008)	(0.015)
HH size	−0.661***	−0.831***	−0.860***	−1.193***	−0.804***	−0.641***
	(0.014)	(0.062)	(0.017)	(0.138)	(0.020)	(0.104)
Total land owned	−0.118***	−0.186***	−0.065**	−0.138**	0.008	0.019
	(0.024)	(0.040)	(0.027)	(0.057)	(0.014)	(0.016)
Education head of HH	−0.346***	−0.803***	−0.161***	−0.733***	−0.160***	0.067
	(0.011)	(0.156)	(0.012)	(0.234)	(0.013)	(0.142)
Sex of head of HH	−1.215***	−0.903***	−1.617***	−1.367***	−0.842***	−1.154***
	(0.086)	(0.174)	(0.097)	(0.201)	(0.102)	(0.242)
Age head of HH	0.000	−0.035***	0.015***	−0.025	0.024***	0.037***
	(0.002)	(0.012)	(0.002)	(0.016)	(0.003)	(0.009)
Hindu	0.048	−0.115	0.024	−0.182	0.075	0.104
	(0.083)	(0.152)	(0.086)	(0.194)	(0.106)	(0.145)
Schedued tribe	0.785***	3.099***	0.352***	3.323***	0.341*	−0.826
	(0.135)	(0.833)	(0.133)	(1.246)	(0.195)	(0.798)
Scheduled caste	0.982***	2.185***	0.533***	2.011***	0.122	−0.566
	(0.091)	(0.424)	(0.090)	(0.605)	(0.120)	(0.454)
Other backward caste	0.358***	0.804***	0.101	0.693***	0.118	−0.167
	(0.068)	(0.179)	(0.078)	(0.263)	(0.079)	(0.210)
Year F.E.	yes	yes	yes	yes	yes	yes
District F.E.	yes	yes	yes	yes	yes	yes
Observations	136,070	136,070	92,725	92,724	43,345	43,345
LIML estimate				28.076		−24.606
CLR 95% conf. set				[7.107; 65.625]		[−56.862; −3.365]
F-stat		12.04		6.72		7.93
Montiel-Pflueger F-stat				4.86		2.83

Notes: *All* includes the whole sample, *Below* is the sub-sample of households below the mean of monthly household consumption expenditure of 724 Rs. (roughly 10 USD) and *Above* is the sub-sample of households below the mean of monthly household consumption expenditure of 724 Rs. The dependent variable is the share of expenditure in energy over a household's total monthly expenditures. All regressions contain a constant. Standard errors in parentheses are robust and clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

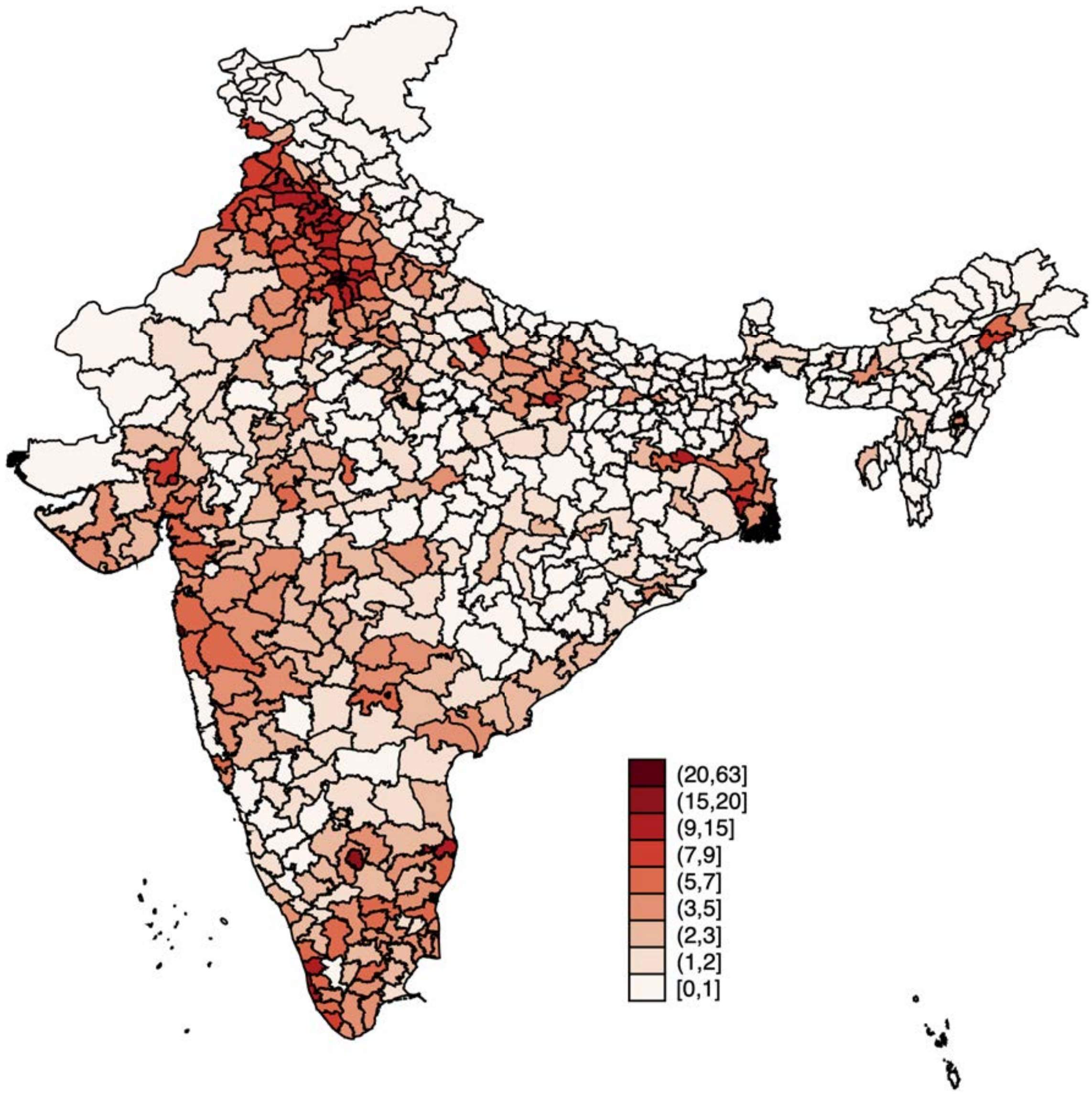
9 Figures

Figure 1: Hydro power plants



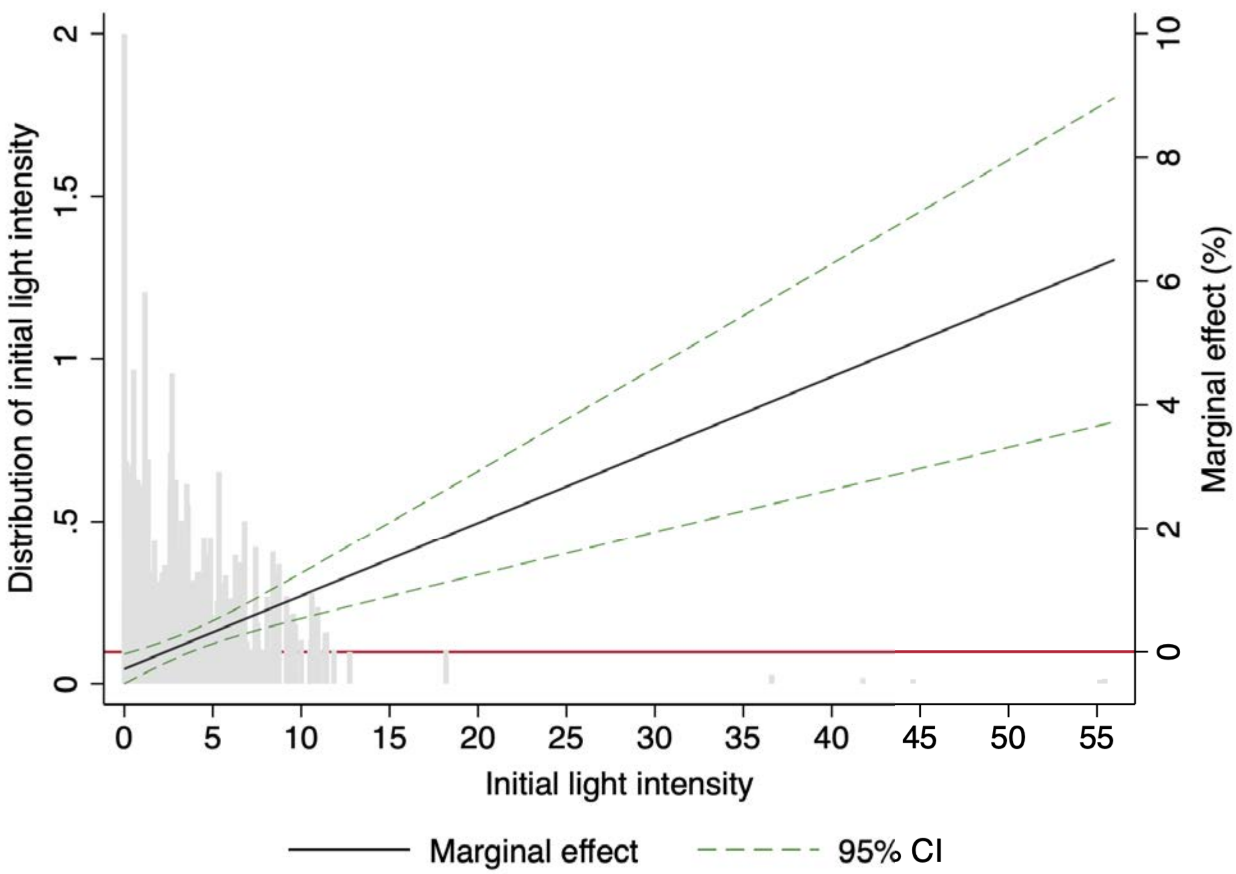
Note: The map on the left shows the location of all the hydro power plants commissioned in India between 1927 and 2018, the blue circles represents the plants commissioned up to 2004, the red dots the plants commissioned between 2005 and 2009 and, finally, the green circles represent the plants commissioned between 2010 and 2018. The size of the circles/dots is proportional to installed capacity. The map on the right only shows the plants commissioned between 2005 and 2009. The size of the dots is proportional to installed capacity.

Figure 2: Night light intensity by district, 1993



Notes: Night light intensity varies between a minimum of 0 (no light output) and a maximum of 63 (maximum light output). Darker shades of red indicate a high light output, while lighter shades indicate a smaller light output. The figure represents the situation of 1993. Source: `http://india.nightlights.io/#/nation/2006/12`

Figure 3: Marginal effect of a supply shift



Notes: The graph represents the marginal effect of the instrument in the first stage estimation. The marginal effect is computed using the coefficients from equation (3) in the following way: $\frac{\partial E_{hdst}}{\partial H_{st-1}} = \beta_1 L_{sd0} + \beta_2$. The gray histogram represents the distribution of districts across the initial level of electrification (1993).