Demand for Electricity in a Poor Economy*

Robin Burgess, Michael Greenstone, Nicholas Ryan, Anant Sudarshan†

July 17, 2019

Abstract

Over a billion people do not have electricity in their homes and whether electricity spreads across the developing world hinges on household willingness to pay for it. Our surveys, conducted in 100 villages in India's poorest state, uncover innovations in off-grid solar technologies and government-led grid expansions which imply that households face a growing choice set on where to source electricity from. The main contribution of this paper is to create a demand system founded upon experimental variation in solar prices that allows us to assess household willingness to pay for different sources of electricity. This reveals two key insights. The first is that though off-grid solar is valuable to unelectrified households it is largely a stop-gap with households preferring the grid when that becomes available. The second is that extending the grid to all consumers necessitates rationing and disadvantages off-grid solar whereas raising prices or removing theft limits grid access and advantages off-grid solar. Our paper demonstrates that these trade-offs will be central to electricity policy in poor economies.

*We thank Manoj Sinha and Col. Baljit Singh of Husk Power Systems for their partnership in implementing the project. We are grateful to Liran Einav and seminar participants at Berkeley, Harvard, LSE, MIT, Stanford and the Western Economic Association for comments. We thank Shruti Bhimsaria, Menna Bishop, Catherine Che, Ananya Kotia, Rakesh Pandey, Aditya Petwal, Víctor Quintas-Martínez and Rashi Sabherwal for excellent research assistance. We thank the Shakti Sustainable Energy Foundation, the LGT Venture Philanthropy Foundation, USAID and the International Growth Center (IGC) for financial support. All views and errors are solely ours.

†Burgess [Dept of Economics, LSE: r.burgess@lse.ac.uk]; Greenstone [Dept of Economics, University of Chicago: mgreenst@uchicago.edu]; Ryan [Corresponding author, Dept of Economics, Yale University: nicholas.ryan@yale.edu.]; Sudarshan [University of Chicago, Energy Policy Institute at Chicago (India): anants@uchicago.edu]

Contents

1 Introduction **3**

2 Background and Data: The Electricity Landscape in Bihar **6**
2.1 Data . . . 7
2.1.1 Household panel survey . . . 8
2.1.2 HPS administrative data on microgrids . . . 9
2.1.3 State utility administrative data on the grid . . . 9
2.1.4 Survey of diesel generator operators . . . 9
2.2 Baseline electricity choices . . . 10
2.2.1 Grid electricity . . . 10
2.2.2 Diesel . . . 10
2.2.3 Own solar . . . 11
2.2.4 Microgrid solar . . . 11
2.3 Changes over time . . . 11

3 Empirical Analysis **13**
3.1 The Solar Experiment . . . 13
3.1.1 Experimental Design . . . 14
3.1.2 Results . . . 14
3.2 Model of Demand for Electricity Sources . . . 16
3.2.1 Specification . . . 16
3.2.2 Estimation . . . 18
3.2.3 Results . . . 20

4 The Value of Electrification **22**
4.1 Patterns of Substitution . . . 23
4.2 Policy Counterfactuals . . . 23
4.2.1 Counterfactual 1: The Value of Solar Innovation . . . 24
4.2.2 Counterfactual 2: Improving the Quality of the Grid . . . 25
4.2.3 Counterfactual 3: Reducing Grid Theft . . . 26

5 Conclusion **27**

6 Figures **33**

7 Tables **46**

A Appendix **53**
A.1 Additional results . . . 53
A.2 Robustness of demand model estimates . . . 57
A.3 Construction of instrument for hours of supply . . . 62

1 Introduction

As this paper reveals, the electricity landscape in the rural parts of developing countries is undergoing radical changes. Large-scale attempts to extend access to grid electricity (Dinkelman 2011, Lee, Miguel and Wolfram 2016, Burlig and Preonas 2016, Burgess et al. 2019) have coincided with technological innovations propelling off-grid solar's emergence as a viable alternative (Aklin et al. 2017, Grimm et al. 2016). Diesel generation also often exists alongside or in place of the grid (Allcott, Collard-Wexler and O'Connell 2016). How households value these different choices is the question that lies at the heart of this paper.

This is important as the extent to which households value, and hence are willing to pay, for electricity from different sources will determine whether or not it spreads across the developing world. Currently, about a billion people, mainly located in South Asia and sub-Saharan Africa, do not have electricity in their homes. This is close to the world's population in 1879 when Thomas Edison invented the light bulb (Maddison Project, 2013). Even in regions where electricity is available, last-mile connectivity challenges have left many without access and forced others to suffer from erratic, unreliable supply (Lee, Miguel and Wolfram 2016, Burgess et al. 2019). This stands in strict contrast to the 24-hour electricity which is taken for granted in developed countries.

Combating darkness has become a priority for developing country governments, with extension of off-grid solar technologies often being promoted alongside traditional grid expansions.[1] Whereas the economies of scale available from grid electrification require high fixed costs to be spread over a large population (International Energy Agency, 2017), fixed costs for off-grid solar can be borne by even a single, isolated household. The ready nature and falling costs of solar technology have thus spurred hope of a faster, greener path to universal electrification.[2] Former UN Secretary General Ban Ki-moon proclaimed "Developing countries can leapfrog conventional options in favor of cleaner energy solutions, just as they leapfrogged land-line based phone technologies in favor of mobile networks."[3] However, underlying this view is the assumption that rural households value, and hence are willing to pay for, electricity and will thus benefit from improved access.

An equally compelling, alternative hypothesis is that existing darkness simply reflects low household demand for electricity.[4] Understanding whether electrification will spread across the rural areas of developing countries either via grid or off-grid sources therefore requires us to quantify household willingness to pay for different forms of electricity *depending* on the choices they face. This is the objective of this paper. To do this we combine a randomized control trial where we vary the price and availability of an off-grid solar technology with demand estimation covering all potential sources of electricity in the Indian villages that we study. The main contribution of this paper, therefore, is to create a demand system founded upon experimental variation in solar prices that allows us to assess household willingness-to-pay for different sources of electricity.

We proceed in four steps. First, we field, over three waves in 2013, 2016 and 2017, an innovative survey instrument in 100 villages in India, which allows us to monitor the variety of sources from which households obtain electricity and how this changes over time. Second, we introduce a new off-grid solar technology, experimentally varying its price and availability. This allows us to trace out a demand curve for this widely touted alternative to grid electricity. Third, we broaden the scope of our analysis by estimating a nested logit model of electricity source demand, combining

[1] USAID, for example, launched *Power Africa* in 2013 and DFID launched *Energy Africa* in 2015.

[2] See, for example, "Africa Unplugged: Small-scale Solar Power is Surging Ahead", *The Economist*, October 29th, 2016.

[3] "Powering Sustainable Energy for All," *The New York Times*, January 11th, 2012.

[4] Only in the former case do state-led electrification efforts or subsidization of off-grid solar solutions arguably constitute worthy uses of limited funds.

experimental price variation from the solar experiment with electricity source characteristics and household characteristics data collected over three survey waves. By matching consumers across several datasets, we are able to estimate how demand for *all* electricity sources varies according to household characteristics and the availability and characteristics of other source options. This is in-line with the spirit of differentiated product demand models (Berry, Levinsohn and Pakes, 1995, 2004; McFadden, 1974; Lancaster, 1971). Our approach affords greater insight into what drives demand for electricity in a poor economy vs what would be possible by just employing the solar experiment. Fourth, we construct counterfactuals on how market shares and willingness to pay for different electricity sources vary under three scenarios: (i) technological innovations lead solar photovoltaic panels and associated batteries to become cheaper, (ii) the reach and quality of grid electricity is improved and (iii) theft of grid electricity is reduced so that effective prices cover costs.

Demand estimation underpinned by the solar experiment therefore yields valuable out-of-sample predictions regarding how overall access to electricity (the extensive margin) and which source households will choose (the intensive margin) will respond to future innovations in solar technologies and changes in the functioning of the electricity distribution industry. Given how fast the electricity landscape is changing, gaining these insights will be key to guiding future policy decisions, such as the debate on the levels of public investment and subsidization required to achieve the goal of universal access to affordable, reliable and modern energy services by 2030 (UN Sustainable Development Goal 7).[5]

Our experiment is set in rural Bihar, one of India's poorest and most energy-deprived states where per capita electricity consumption in our baseline survey year (2013) is 1% of that in the US and access to electricity is below that in sub-Saharan Africa (World Bank 2017). Panel A in Figure 1 portrays the electricity access situation at baseline. Though most villages are black, which denotes that no electricity was available (73% of households), 17% had diesel generated electricity (blue), 5% grid electricity (red), 5% own solar (yellow) and 1% microgrid solar (green). Thus, even in this poorest backwater of India, we find a strikingly competitive retail market in which households face a rich array of electricity source choices.

We partnered with Husk Power Systems (HPS), a private company that installs and maintains solar microgrids for a monthly charge, to conduct a randomized experiment on the pricing and availability of off-grid solar power. Panel B in Figure 1 depicts the situation three years later, after we had put in place an experiment whereby (i) 33 villages were offered the solar microgrids at a subsidized monthly price of INR 100, (ii) 33 villages were offered the microgrids at the prevailing market price (initially INR 200, later cut to INR 160) and (iii) 34 villages were kept as controls. Panel C in Figure 1 depicts the situation four years after baseline. Figure 1 reveals an uptick in adoption of microgrid solar relative to baseline as the experiment unfolded, however, what is more striking is the rapid growth in other sources of electricity. If we restrict attention to just the 34 control villages, which were isolated from the effects of the experiment, we find that grid connections jumped from 3% to 37% of households and own solar from 3% to 15%, which drove down the proportion of unelectrified households from 76% to 44%.[6] Expansion of grid electricity was government-led whilst that for own solar (which is privately procured) was driven by falls in the cost and improvements in the quality of home systems. Understanding how demand for *all* sources of electricity as captured by market shares and willingness to pay were affected by these rapid changes in the electricity landscape represents the central challenge of the paper.

Our key results are as follows. We found household demand for microgrid solar to be highly price

[5]Apart from universal access by 2030 SDG 7 also targets increasing substantially the share of renewable energy in the global energy mix.

[6]Interestingly connections to diesel electricity fall from 17% to 4% of households.

elastic, reflecting the competitive nature of the electricity market. Three years after baseline, 8% of households had adopted microgrids at the subsidized price of INR 100, relative to near 0% at the prevailing market prices of INR 160-200. Despite low take-up, households in treatment villages were significantly more likely to own light bulbs, use more electricity, purchase more mobile phones and spend less money charging them. Unsurprisingly, these effects are muted in the normal price treatment arm where adoption was low. We were unable to detect any welfare benefits, measured in terms of household income, children's reading and math test scores and self-reported respiratory problems. As such, solar microgrids appear to be valuable in terms of cellphone charging and lighting but far from transformative.

Our nested logit demand estimates echo the reduced-form result that households are highly price elastic and reveals a complex pattern of substitution. For example, demand for diesel generators, which represented the primary electricity source at baseline, is crushed as the grid and off-grid solar become more accessible and affordable. By contrast, own solar picks up demand as it becomes cheaper and improves in quality across our study period. The model furthermore finds that richer households have much stronger preferences for the grid over other electricity sources, likely driven by their demand for higher-load appliances like televisions and fans which can only be supported by this source.

Our counterfactual analysis concludes that the value of off-grid solar largely rests on the grid being incomplete and dysfunctional. Willingness to pay for off-grid solar doubles in the absence of the grid, but when it is the only electricity source available, its market share is merely 34% and so almost two thirds of households choose to remain in darkness. In this sense, solar is at best a stop-gap until other sources are offered, rather than being a viable permanent substitute. We furthermore estimate that demand for grid electricity would be substantially depressed if subsidies were removed, including the implicit subsidies afforded by tolerance of theft under the existing regime. This allows us to rationalize India's seemingly dysfunctional system of simultaneously subsidizing and rationing electricity, in that feasible alternatives would devastate household electricity access among the rural poor.

Our paper contributes to the literature on the demand for and effects of electrification. Much of the literature on rural electrification has focused on the spread of grid electricity, and found that grid electrification has had large effects on labor supply, productivity and welfare.[7] There are also a handful of experiments on the demand for electricity connections including at least two experiments on demand for off-grid solar.[8] The impact analyses in these papers are broadly consistent with our findings that demand for off-grid solar and its welfare benefits for households are limited. A recent experiment in Kenya finds that grid electrification is prohibitively costly for rural Kenyan households, even at heavily subsidized prices (Lee, Miguel and Wolfram, 2016). This finding agrees with our finding of highly elastic demand for off-grid electricity among rural Bihari households. Our paper takes a couple of steps to unify this literature. First, we estimate how households value both grid and off-grid electricity together, in a single demand system. This contrasts with previous work which considered each source in isolation, and so cannot shed light on substitution between sources, which are increasingly important in the context of rapidly expanding choice sets. Second, we study

[7]Rud (2012) shows that electrification leads to structural transformation. Dinkelman (2011) demonstrates that electrification can lead to increased employment and female empowerment. Lipscomb, Mobarak and Barham (2013) find large effects of electrification on the UN Human Development Index and average housing values. Barron and Torero (2017) find that household electrification reduces indoor air pollution.

[8]Aklin et al. (2017) find that offering off-grid solar power in Uttar Pradesh, India increased electrification rates by 7 pp but had no effect on socio-economic outcomes like expenditures, business creation or time studying. Grimm et al. (2016) find that household willingness to pay is high relative to incomes but that nearly no households are willing to pay market prices.

the supply side to try to rationalize the seemingly dysfunctional rural electricity sector. Our model shows that the playing field for off-grid solar is indeed lopsided, due to massive subsidies for grid electricity, but that these subsidies can be justified by the government that values access for the rural poor.

Our study also contributes methodologically to the development literature by placing a greater emphasis on the external validity of experimental results. Field experiments have lately gotten longer to address the realism and durability of effects.[9] The analysis of experiments often uses structural models to understand how an intervention works and to go from reduced-form treatment effects to policy counterfactuals (Duflo, Hanna and Rya, 2012; Bryan, Chowdhury and Mobarak, 2014). Our experiment offered long-term subsidies for solar technology, on a standing basis for everyone in treatment villages, and tracked take-up and welfare outcomes over three-and-a-half years. The experimental variation therefore closely mimics real-world variation in the price of a technology. We use this variation to recover the price elasticity of demand for electricity sources in a discrete choice demand model.[10] We use the model estimates to study counterfactuals on grid policy that are well beyond the boundaries of the experiment itself.

The rest of the paper is organized as follows: Section 2 describes the data we gathered and provides background on electrification in Bihar, Section 3 presents the empirical analysis, which is comprised of the solar experiment and a demand estimation covering all sources of electricity, Section 4 presents our counterfactual analysis and Section 5 concludes.

2 Background and Data: The Electricity Landscape in Bihar

The experiment is set in Bihar, home to around 125 million people and one of India's poorest and least electrified states. Table 1 juxtaposes the state of Bihar, India, sub-Saharan Africa and the United States on the dimensions of per-capita GDP, per-capita electricity consumption and access to electricity. The electrification rate in Bihar at the beginning of our study was strikingly low at 25%, and the average citizen used just 122 kWh of electricity per year. At this low level of consumption, an individual can power only two light bulbs totaling 60 watts for six hours per day throughout the year. In stark contrast, access to electricity in the United States is universal and energy consumption per capita is over one hundred times larger. However, as is made clear by Table 1, electricity access and consumption in Bihar are also far behind the figures for India as a whole and, perhaps more surprisingly, sub-Saharan Africa. In India, for example, 79% have access to electricity and the average individual consumes six times as much energy as in Bihar. This is in keeping with Bihar's GDP per capita, which, at USD 420, is just 25% of the figure for India overall.

Despite this underdevelopment, Bihari households do not face a binary choice between darkness and the grid. Rather, even in India's poorest rural backwater, there exists a thriving electricity market in which diesel generators, solar and the grid compete for market share, something that surprised us upon entering the field. Bihar's astonishingly low electrification rate combined with the diverse set of electricity sources available to households is precisely what makes Bihar an interesting laboratory for the study of electrification. In particular, by examining how households in such an

[9] De Mel, McKenzie and Woodruff (2013) study firm formalization with a 31-month followup, Dupas and Robinsona (2013) study household savings with a nearly 3-year follow-up. Bandiera et al. (2017) study a 7-year follow-up to an asset transfer program.

[10] We are not aware of prior work that uses an experiment to estimate price sensitivity in a discrete choice demand model. Kremer et al. (2011) is a close precedent that experimentally varies the quality of a good (a local water source) and uses observable variation in walking distance to water sources as a proxy for price.

environment value electricity, we will be able to shed light on which sources will drive electrification in Bihar and other poor economies.

As we document in Burgess et al. (2019), a social norm has developed in India, which is that all deserve power regardless of payment. This is for two reasons. First, technological constraints make it difficult, if not impossible, to target electricity supply to particular (e.g., paying or poor) customers. Grid connections are not easily excludable and almost anyone can access the same electricity wires, regardless of whether they have paid into the system. Second, social and political constraints have rendered politicians and the government tolerant of rampant theft and non-payment. As a result, the average amount paid across paying and non-paying customers is only half of the official grid price of INR 150, creating large losses for the state utility company and a starkly uneven playing field in electricity markets. Such a dysfunctional system is characteristic of India, where populist, rights-based welfare schemes constitute a widespread form of redistribution. Over the past fifteen years, large government programs have provided universal rights to rural unskilled employment, health, and education, just to name a few.

Over the course of our study, these complex market dynamics were further disrupted by a state-led push to electrify rural households, enshrined in ambitious government targets. These were embraced with enthusiasm among Bihar's local leadership. In addition to traditional grid expansion, discussions have increasingly included a potential role for solar energy. This openness towards an alternative, greener path to electrification has been partially fueled by precipitously falling costs due to technological innovation and rising concern about the effects of global warming on Indian citizens.

In Section 2.1, we first provide details on the different data collection efforts we carried out. Based on the analysis of this data, we will then proceed to examining baseline electricity choices in Section 2.2. This reveals an interesting array of choices available to rural households. In Section 2.3, we consider how these choices evolved over our study period, which extends from 2013 to 2017. We find that changes in source shares can be linked to state-led grid expansion, falling costs of solar technologies, our experimental introduction of solar microgrids and the endogenous exit of diesel suppliers. These institutional details and on-ground observations will motivate the form of our empirical analysis of demand for electricity, which is covered in Section 3.

2.1 Data

This section describes the range of sources from which the data used in our analysis was collected. Our aim was to paint a comprehensive picture of Bihar's electricity market, gathering rich information on both the supply and demand sides. This involved various collaborations, including one with HPS and another with the Government of Bihar as providers of microgrids and the grid respectively. We also orchestrated a substantial amount of original data gathering in the form of a three-round household panel survey and a survey of diesel generator operators. This amalgamation of data on the universe of electricity sources available to households underpins our ability to extend our analysis beyond reduced-form estimates of demand for solar microgrids. As will be described in Section 3, we estimate a nested logit demand model which characterizes willingness to pay for each source in the evolving choice sets that households face as a function of choice-specific characteristics, such as price and daily hours of supply. We are also able to construct counterfactuals that model how demand is likely to respond to further innovations in solar technologies and efforts to expand the grid into rural areas.

Figure 2 illustrates how our different data collection efforts were structured. In August 2013, we conducted a consumer identification survey in each of the one hundred villages covered by our study. This was subsequently used to assign villages to treatment and control groups and also served as the basis from which we drew the sample of households for our panel survey. We collected household-level data that covered demographics, wealth, assets and electricity consumption habits in a sequence of three waves: a baseline survey in November-December 2013, a endline 1 survey between May and July 2016 and a endline 2 survey one year later in mid 2017. Microgrid administrative data at the household-level became available after our baseline survey, at which point we experimentally introduced this form of electricity in treatment villages. Grid administrative data from the state electricity company is also available at the household-level, and we have access to monthly billing and collection data as part of our partnership with the Government of Bihar. The consumer identification numbers (which are identifiers used in grid billing and collection) that we collected in the endline 2 survey enabled us to match across datasets and recover the amounts of grid and microgrid electricity that the households in our experiment paid for and consumed. Finally, we completed our dataset construction by merging in diesel operator data at the village-level.

2.1.1 Household panel survey

Our household panel survey covered roughly three thousand households across 100 villages.[11] This sample was drawn so as to represent those with an interest in a microgrid solar connection. This was achieved through a customer identification survey conducted in August 2013 across all sample villages, which elicited household willingness to pay for a solar microgrid connection. A random sample of 30 households per village was selected among those who expressed interest in paying for a solar connection at a monthly price of INR 100. This identification was barely restrictive in practice - households were not required to put down a deposit nor where they held to their initial statement of interest when the product was eventually offered. Over 90% of households without electricity or with just diesel-based electricity said they would be interested in using microgrids. The same was true for over 70% of households with a grid connection or home solar panels.[12] As such, our survey sample is representative of the vast majority of potential microgrid solar users in the population.

The survey administered to our final sample consisted of two thick rounds, which we refer to as baseline and endline 1, and one thin round, endline 2. The baseline survey took place in November and December of 2013, followed by endline 1 which was conducted between May and July of 2016. These two rounds employed approximately the same survey instruments and covered a range of variables that lend us rich insight into the characteristics of Bihar's rural population. The categories of collected data include demographics, such as literacy and number of adults per household, as well as various wealth proxies, such as income, size and structure of house and ownership of agricultural land. We also collected data on electrification status, the varied sources from which households procured electricity, the payments associated with these sources and other characteristics such as capacity and hours of supply. Moreover, we recorded ownership of assets such as mobile phones, bulbs, fans and TVs to understand how households might be using their electricity supply. Finally, we tracked select measures of education and health, namely children's reading and math test scores and self-reported respiratory problems. The final survey round, endline 2, took place one year after endline 1, in May 2017, and was carried out for the specific purpose of collecting data on demand for and characteristics of the different electricity sources over a longer period. As such, this wave did not track the detailed household covariates with which the first waves were concerned.

[11] The total population of households in all 100 villages was 48,979.

[12] As Section 2.2.4 makes clear, the share of households who expressed an interest in using solar microgrids was vastly higher than actual take-up rates.

2.1.2 HPS administrative data on microgrids

The second source of data employed in our analysis is an administrative dataset on microgrid customers from HPS, the company responsible for installing and maintaining solar microgrids with whom we partnered in experimentally introducing this source of electricity in treatment villages. The dataset includes information on pricing and customer payments from January 2014 to January 2016. This data was subsequently matched with that collected in our household surveys in order to estimate demand for solar microgrids over time between our survey waves, exploiting experimental variation in price.

2.1.3 State utility administrative data on the grid

We also held a valuable partnership with Bihar's state electricity company, through which we were able to access three datasets describing the functioning of the electricity grid in Bihar: a consumer database for all paying customers, a billing dataset containing monthly billing prices and quantities and a collection dataset on actual customer payments to the distribution company. We combined this with data collected in our household panel surveys to construct a time-series of electricity consumed by grid customers and the amounts they were billed for and payed. This was achieved by recording household consumer IDs in our endline 2 survey, which were subsequently matched with those listed in the datasets provided by the state utility company. In the environment we study, households living in villages where the grid was available could connect to the grid in two ways. The formal way entailed filing an application with the state distribution company during government "connection camps", or through a linesman, licensed revenue collector, or power contractor. Alternatively, households could connect informally by hiring a local electrician to wire a connection, potentially bribing utility staff to look the other way. Thus, we define formal households as those who provided us with their consumer IDs during endline 2 and that we successfully located in the consumer database. For these formal payers, we observe the bills they received and the payments they made. Informal consumers, on the other hand, are those who either failed to provide us IDs or gave us IDs that could not be found in the administrative datasets. Some consumers received electricity bills but had made no payments over the preceding twelve months; we treat these consumers as informally connected.

Our ability to distinguish between formal and informal households and to observe their respective consumption behaviour is extremely valuable. The full price of grid for bill payers is high at INR 150 per month. However, of the 158 households who reported using the grid at baseline, less than half answered in the affirmative to the question "Do you pay electricity bills?". Including the non-payers in the average reduces the monthly price of grid to just INR 73 (Table 2, row 1), which is on par with the price of own solar.[13] This distinction between the official and the effective price of grid is indicative of the uneven playing field that characterizes Bihar's electricity market. The implications this has for the relative popularity of different sources will be elaborated upon in Section 4.

2.1.4 Survey of diesel generator operators

Our final source of data is a survey of diesel generator operators. We collected monthly information on cost structure, hours of operation, connection plans, and number of customers from January 2014 to 2016. This affords us a valuable insight into the characteristics of this electricity source, which at baseline is the most popular choice among the households we study.

[13] Tariffs are determined for each household by their consumer category. Many households in the sample qualified for the Kutir Jyoti program, a government effort to electrify households below the poverty line by providing connections without installation charges at a flat-rate monthly tariff of INR 55--60 for unmetered households over the experiment duration. Metered KJ consumers receive their first 30 units at a discounted rate.

2.2 Baseline electricity choices

This section describes household choices in terms of where they sourced electricity at the time of our baseline survey in late 2013. We also discuss the characteristics differentiating one source from one another, namely in terms of load, hours of supply and price, as summarized in Table 2. Later, we will construct a nested logit demand model that captures how these attributes map onto consumer demand for each source. We will develop predictions regarding how changes in source attributes, such as further grid expansions and falling costs of solar due to technological innovations, will be reflected in source-specific market shares.

2.2.1 Grid electricity

The grid price of INR 73, reported in Panel A of Table 2, represents the self-reported monthly payment, averaged across formal and informal customers. On this basis, the grid is just about the cheapest source on the market. However, as previously explained, this price stands at just half of the full price of INR 150.[14] Those connected to the grid are able to access a higher load than can be supported by any other source, [15] allowing them to operate larger appliances than would otherwise be possible. Accordingly, 11% of grid users own a television relative, to just 1% of diesel and 4% of own solar users (Panel B). The same disparity exists for fan ownership. By contrast, all surveyed households own small appliances such as mobile phones or a lightbulb.

Low rates of bill collection have meant that electricity rationing, or load shedding, is a common practice by state distribution companies. Grid feeders, each of which typically supplies electricity to several villages, maintain log books with daily hours of electricity supply, and in many cases, the exact hours of the day for which power was supplied. Across the regions we study, mean daily supply stood at 9.7 hours at baseline, still higher than for any other source.

Despite these desirable attributes, baseline take-up of the grid was low at just 5% (Figure 4). This is largely attributable to lack of availability, as the grid was available in just 35% of villages at baseline.

2.2.2 Diesel

Diesel was the most popular of the electricity sources available at baseline, with a market share of 17%. This form of electricity was supplied by village residents who had purchased a generator and offered connection plans with monthly or semi-monthly payments. As displayed in Panel A of Table 2, diesel generators were substantially cheaper than the grid at its full price of INR 150, but about 35% more expensive than the effective grid price of INR 73. Moreover, diesel supports a much lower average load of 133 watts. Of the various plans offered, the modal option consisted of a 100 watt connection for INR 100 per month. This was sufficient for consumers to power one or more light bulbs or charge mobile phones in the evening.[16] Generators run on a predictable schedule in the evening and early night-time for an average daily supply of 3.4 hours. Because diesel operators

[14] Tariffs are determined for each household by their consumer category. Many households in the sample qualified for the Kutir Jyoti program, a government effort to electrify households below the poverty line by providing connections without installation charges at a flat-rate monthly tariff of INR 55--60 for unmetered households over the experiment duration. Metered KJ consumers receive their first 30 units at a discounted rate.

[15] We do not directly observe the load available to grid, diesel, and ownsolar customers, and thus approximate load categories using household appliance stocks recorded in the household survey data.

[16] The solar microgrid provided similar lighting services at a much lower wattage because of the provision of highly energy efficient LED lights with the connection.

require a sufficient number of customers to support their fixed costs, diesel was only available in 62% of villages at baseline.

2.2.3 Own solar

Alongside the grid and diesel electricity, households also had the option of buying solar systems in private markets. We refer to this as "own solar" to distinguish it from the HPS microgrid. Households paid for these panels upfront and would usually have had to travel to a larger market town or city to do so. In principle, all households should be able to access this source of electricity, and so we assumed that it was available in all villages. As shown in Panel A of Table 2, own solar is strikingly similar to diesel in terms of price and supported load. Once purchased, the effective cost of operating own solar is close to zero; for our purchase, the price of own solar needs to be comparable to the price of other sources, which don't have large upfront outlay. Hence, for the price of own solar, we translated the upfront capital costs into an equivalent monthly payment, using an assumed lifetime of seven years and 20% interest rate. Around 5% of baseline households used own solar, for an average of 8 hours per day, which is significantly more than the average supply reported by diesel customers but still behind grid customers'.

2.2.4 Microgrid solar

Solar microgrids are the instrument employed in our experiment. In partnership with HPS, we experimentally varied the price and availability of this form of electricity across the villages we studied. HPS solar microgrid consists of a 240 watt panel shared between six to nine households. Each households receives its own 3.2 volt rechargeable battery and meter, which comes with a key pad to secure access to the battery. Households must purchase codes on a monthly basis to keep using the system. Each household on the microgrid enjoys 25 to 40 watts of power for 5 to 7 hours per day. This represents a small quantity of power, used to supply a high-efficiency light bulb and an electrical outlet, typically used for mobile phone charging. The inability of solar microgrids to power even small appliances such as fans and radio sets distinguishes it from other available sources. Unsurprisingly, microgrid users have the lowest asset ownership rates (Table 2). Customers in the market price treatment arm were initially charged INR 200, making microgrid the most expensive source and twice as expensive as the second most expensive source, diesel. However, twelve months into the experiment, the normal price for microgrid was cut to INR 160 due to insufficient demand. Those in the subsidized price treatment arm were charged INR 100 throughout the experiment, which is comparable to the price of diesel. At baseline, and before microgrids had been experimentally introduced by HPS, take-up was low at just 0.7%.

2.3 Changes over time

The electricity market in Bihar and much of the rest of rural India is undergoing considerable disruption, the biggest change being state-led grid expansion. In his 2015 independence day address, Indian Prime Minister Narendra Modi launched a rural electrification program with an ambitious 1,000-day deadline to electrify 18,452 census villages[17] still without access, at an estimated cost of USD 11 billion.[18] This goal was declared achieved ahead of schedule on April 28, 2018. However,

[17]There are a total of almost 600,000 census villages in India.

[18]This scheme, the Deen Dayal Upadhyaya Gram Jyoti Yojana (DDUGJY) is not the first federal government electrification scheme in India. Earlier programs, such as the Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY) had similar objectives and have been subsumed in DDUGJY (Government of India, 2015).

a village is defined as electrified once a minimum of 10% of its households and public spaces, such as schools and health centres, have access to electricity.[19] Consequently, even though official targets were "met", an estimated 31 million households across India were still without grid access and last-mile grid connectivity remains an enormous challenge.[20]

Federally-funded electrification programs require political will for state-level implementation. Bihar is an exception in this regard, with access to electricity having emerged as a big political issue in the past decade(Kumar, 2019). Nitish Kumar, Bihar's six-time Chief Minister who has ruled the state for almost all of the past 14 years, has made universal access to electricity a trademark of his political campaigns (Business Today, 2017).

Our study was perfectly timed to capture the disruption that this multi-level state push for rural electrification caused in Bihar's electricity markets. We document this in Figure 1and Figure 4, which show the share of households using different electricity sources for each survey wave. At baseline, just 35 of our 100 study villages had access to the grid and 73% were without electricity entirely. This manifests in Figure 1 as a substantial amount of black and a very small amount of red, the latter reflecting grid take-up of just 5%. Diesel, which is shown in blue, was the most popular electricity source at baseline, adopted by over 16% of our study households.

When we returned to conduct our endline 1 survey just two and a half years later, the electricity market had already been entirely transformed. The number of study villages with grid access had almost doubled and take-up of this electricity source had risen four fold (Figure 4). As a result, demand for diesel had been crushed: take-up had fallen to just 3% and the number of villages not using it at all more than doubled to 90, with many operators having exited the market.

The grid continued to expand thereafter, with take-up almost doubling in the twelve months between endline 1 and 2 surveys. This growth occurred not only at the extensive (village) margin, with the number of study villages with grid access having risen to 76, but also at the intensive (household) margin. In particular, whereas no villages in our sample had grid take-up over 50% at baseline, there were 44 such villages by endline 2. Grid shares benefited from a concerted government campaign to electrify households via connection camps which massively increased last-mile grid connectivity.

A second disruption to rural electricity markets has been caused by global declines in the cost of solar technologies. The US National Renewable Energy Laboratory projects a 55% reduction in the cost of solar photovoltaics and a 75% reduction in the cost of batteries by 2022. Our data reflects these trends. In particular, column 3 of Table 2 shows that the cost of own solar systems fell by approximately 11% from INR 74 to INR 66 per month between baseline and endline 2. Take-up for own solar also rose sharply, partly due to this cost reduction but also due to improvements in unobserved qualities, to which we will return later (Figure 4). In Figure 1, the uptick in own solar shares can be seen in the increasing amount of yellow from an almost negligible level at baseline.

In Figure 5, we document how the changes in source attributes interact with household characteristics to change demand for different electricity sources by survey wave. At baseline (Panel A), when grid availability was limited to just 35% of villages, richer households were less likely to be in darkness and more likely to source electricity from diesel or own solar. Microgrids were practically absent owing to low take-up at the unsubsidized price of INR 200. Over the course of our study, we see the grid being preferred by richer households, potentially because they own larger appliances that consume higher loads.

In sum, a state of market upheaval existed in Bihar during the period spanned by our study, and

[19]As of May 2018, only seven Indian states had achieved full household electrification: Tamil Nadu, Kerala, Andhra Pradesh, Punjab, Goa and Gujarat, and West Bengal (Bhaskar, 2018*a*,*b*).

[20]Like the DDUGJY, there are specific federal projects to fund state efforts in this regard. The Primary amongst these is the Pradhan Mantri Sahaj Bijli Har Ghar Yojana (Saubhagya), a USD 2.5 billion federal scheme launched in September 2017.

consisted of a fall in the price of own solar, a dramatic state-led grid expansion and the endogenous exit of diesel generators. The following section first reports the demand for solar microgrids, and then progresses from studying isolated source-specific demand to an analysis that includes the entire demand system of competing electricity sources. The latter then allows us to quantify how much each change in the market has benefited households.

3 Empirical Analysis

The previous section has made it clear that households face considerable choice in where they source electricity from. It also makes clear that the choice set is expanding. In this section we combine a randomized control trial varying the price and availability of microgrid solar (Section 3.1) with a model of demand for electricity connections (Section 3.2) to gain insights into what drives the willingness to pay for different sources of electricity. The randomized control trial allows us to trace out a demand curve for this new form of off-grid solar electricity which is increasingly seen as means of overcoming the darkness deficit in developing countries. And if households adopt this new form of electricity then we can examine how this affects different household outcomes on dimensions such as lighting, phone charging, income, education and health.

However, Figure 1 makes it clear that we need to move beyond the RCT to incorporate rapidly changing market structure that affects both treatment and control villages. To capture this complex and dynamic pattern of household demand and substitution we use a nested multinomial logit to model choice over differentiated electricity sources depending on household characteristics and source characteristics that can vary over time (Berry, Levinsohn and Pakes, 1995, 2004; McFadden, 1974; Lancaster, 1971). Here the RCT holds a deeper value as it enables us to estimate household sensitivity to price in the model using random variation in the price of solar microgrids from the experiment. Estimating a demand system covering all electricity sources affords us a much greater insights into what drives demand for electricity in a poor economy than would be obtainable from the solar experiment alone.

We then build on our analysis of past choices to construct, in Section 4, counterfactuals on how demand for different electricity sources will vary if (i) technological innovations lead solar power to become cheaper, (ii) the quality and reach of grid electricity is improved and (iii) theft of grid electricity is removed. A strength of our approach therefore is that by uncovering the parameters of household electricity choice we can not only conduct a detailed analysis of past demand over all sources of electricity but also make realistic predictions about out-of-sample policies. This is particularly important in a context where solar electricity is undergoing rapid technological change but has to compete against grid electricity which is heavily subsidized both officially and through the countenance of theft (Burgess et al., 2019).

3.1 The Solar Experiment

We partnered with Husk Power Systems (HPS) to vary the availability and price of solar microgrids in a randomized control trial. HPS was founded in 2007 to provide electricity in rural areas using biomass gasifiers as generators to obtain power from agricultural waste, such as rice husks (hence the name of the company). These biomass plants could only serve a village if demand was sufficiently great (e.g., 100 households) and were subject to fuel supply disruptions. HPS made a strategic decision to add a solar microgrid product to its portfolio as a means of reaching a wider set of villages.

3.1.1 Experimental Design

Our experiment sample consisted of 100 villages distributed across three districts in Bihar, as shown in Figure 3. These villages were chosen to fulfill three criteria. First, they were not listed as electrified villages by the government. This implied that either grid connections were entirely unavailable, or very few households were connected to the grid. Second, villages were chosen to be reasonably close to existing HPS operating sites, so that microgrid services could be feasibly expanded to these areas. Third, sample villages had not already been offered HPS microgrids. The total population of households in all 100 villages was 48,979.

The experimental design is a cluster-randomized control trial at the village-level. We randomly assigned villages into one of three arms: a control arm (34 villages) in which HPS did not offer microgrids, a normal price arm (33 villages) in which HPS offered microgrids at the prevailing price of INR 200 per month and a subsidized price arm (33 villages) in which HPS offered microgrids at the reduced price of INR 100 per month. While the prevailing normal price at the start of the experiment was INR 200, HPS later cut this price, but only within this experimental arm, down to INR 160, due to lack of demand at the higher price. Within each treatment village, all households were offered the same HPS connection and pricing, regardless of whether they had previously expressed interest in HPS's product or whether they participated in our baseline survey. Sales of solar microgrid connections began in January 2014, right after the baseline survey.

Table 3 shows the balance of covariates in our baseline survey across treatment and control arms for demographic variables (Panel A), wealth proxy variables (Panel B) and energy access (Panel C). The first three columns show the mean values of each variable in the control, normal price and subsidized price arms, with standard deviations in square brackets. The next two columns show the differences between normal price and control arms and between subsidized price and control arms, respectively, with standard errors in rounded brackets. The final column shows the F-statistic and p-value from a test of the null hypothesis that the differences in means between normal price and control arms and between subsidized price and control arms are jointly zero at baseline.

The joint test rejects the null of equality of treatment and control arms at the 10%-level for three out of twelve variables at baseline. For example, households in subsidy villages are more likely to have solid or *pukka* houses than control households and to have solid roofs. The overall rate of electrification does not differ by experiment arms (variable "Any elec source (= 1)" has p-value 0.54), but households in the subsidy treatment arm are more likely to have electricity from the grid and somewhat less likely to have it from a diesel generator. We address this slight imbalance by including baseline household covariates as controls in both our reduced-form and structural estimates.

Because of the sample criteria, we are working with a population that is poorer than the population of Bihar as a whole. For example if we look at self-reported household incomes in the control group at baseline, we find an average of INR 7,460 per month (USD PPP 2.6 per person per day). Two-thirds of households own agricultural land and about a quarter have a *pukka* house, which is constructed of solid materials like brick. The average household has 3.3 adults living in 2.4 rooms. Only a quarter of households have electricity from *any* source at baseline. Put differently, we are working in poor, largely unelectrified villages in one of India's poorest states. This is an interesting context in which to trace out how households value different forms of electricity.

3.1.2 Results

Our experimental design allows us to estimate a reduced-form demand curve for solar microgrid connections. As previously described, villages were randomly assigned to control, normal-price treatment (initially INR 200, lowered to INR 160 at twelve months into the experiment) or subsidized-price treatment (INR 100). Households then made monthly decisions on whether to pay for electricity from Husk Power Systems. Households that did not pay would eventually be disconnected but from month to month this risk was very low.[21] Consequently, the demand for microgrids can be traced by the household decision to pay for this service, which can change every month.

In Figure 6, we plot the share of households that chose to pay for microgrid electricity at different prices and over different time intervals. Each line on the figure refers to payment at a different horizon: at least once during the experiment, midway through during months 16-18, and at the time of the endline 1 survey at month 29. Three features of consumer demand stand out. First, the demand for microgrid electricity at the normal price prevailing before the experiment is practically zero: only about 2% of households paid even a single month in this arm. Second, demand is highly elastic, as evidenced by the fact that the microgrid share increased to 17% in the subsidized price arm. Third, demand shifted inwards over the course of the experiment, so that even at the subsidized price, only 7% of households were paying by the time of the endline 1 survey.

Figure A1 breaks down the monthly per household cost of supplying HPS. The total monthly cost, in the left bar, is around USD 1.75 (INR 105) per month and comes from an amortization of just the upfront purchase price and installation costs (exclusive of all variables costs related to billing, collection and system maintenance).[22] Therefore, the estimated demand curve implies that demand is near zero at the unsubsidized 2013 price of INR 200. The price sensitivity we find for solar microgrids agrees with previous studies which found that household demand is nearly zero when off-grid solar is priced at cost (Aklin et al., 2017; Grimm et al., 2016).

The demand curve reflects household willingness to pay for off-grid solar electricity. This willingness to pay reflects perceived household benefits from having a connection. There may also be benefits of solar power that are not perceived or valued by the household when choosing whether to buy microgrids. For example, improved lighting can lead to children having more time to study at home, which may or may not be valued by parents. There may also be intra-household spillovers from reduced kerosene consumption and indoor air pollution (Barron and Torero, 2017).

We estimate the impact of access to microgrid electricity on a battery of social outcomes. We begin by estimating a reduced-form model of the following specification:

$$y_{ivt} = \alpha + \beta_N NormalPrice_v + \beta_S SubsidizedPrice_v + \mathbf{x}_i'\delta + \epsilon_{itv} \tag{1}$$

Here, y_{ivt} is the outcome of interest for household i in village v at time t. $NormalPrice_v$ and $SubsidizedPrice_v$ are dummies that take the value 1 when village v was assigned to solar microgrids at regular and subsidized prices, respectively, and $\mathbf{x}_i$ is a vector of controls from the baseline survey. The outcome variables we examined include measures of electricity access, adult and child respiratory problems, reading and math test scores and household income.

Households in both treatment arms got and used electricity microgrids. The microgrid powered two low-wattage LED lamps and one mobile charging point, which was provided with every connection. Table 4 reports regressions of ownership of various appliances and use of electricity on treatment status. Assignment to a subsidy treatment village increases light bulb ownership by 15 percentage points (standard error 4.7 pp) relative to an ownership rate of 32 percentage points in the control group at baseline (column 1). Subsidy treatment households increased hours of electricity use by an estimated 0.94 hours per day (standard error 0.24 hours per day) relative to 1.16 hours in

[21] This is an accurate description of how the contract between the microgrid supplier and consumers played out *in practice*. On paper, Husk Power Systems had an internal rule mandating that households who did not pay for three consecutive months be disconnected, but this was sporadically enforced.

[22] Capital costs as reported by our partner HPS are additionally net of capital subsidies provided by the government, which were of the order of 60% in 2014.

the control group at baseline (column 2). The effects of being assigned to a normal price village are smaller but have the same sign and are also statistically significant. Households assigned to a subsidy treatment village are also more likely to own a mobile phone, by 3.4 percentage points relative to an already high control group ownership rate of 88 percentage points (column 3) (implying that control households are 2.75 times as likely to own a mobile phone as a light bulb). Finally, assignment to a subsidy treatment village also decreases the amount of money spent charging one's mobile phone, which makes sense because households without electricity will typically charge their mobile phones at a shop for a higher per-unit cost of energy.

We therefore find that when solar microgrids are made available, households use more electricity, purchase more mobile phones and light bulbs, and spend less money charging phones. As we would expect, these effects are more pronounced when prices are subsidized. With regards to social outcomes such as income, education and health, we find little evidence that the electricity provided by solar microgrids had a large impact (Table 5). In summary, while the effects of electrification by solar microgrids were not transformative, households still valued off-grid solar for lighting and other energy services.

The next section introduces a demand model to understand how households value off-grid solar relative to alternative sources of electricity. Our experimental variation in the price of microgrids will allow us to estimate household sensitivity to price and thereby gauge willingness to pay under several counterfactual situations in which we vary availability and price of different electricity sources.

3.2 Model of Demand for Electricity Sources

We model consumer demand for electricity sources using a nested logit model. Following Berry, Levinsohn and Pakes (2004), utility from electricity source j depends on household and source characteristics which we have detailed data on from our surveys (see Section 2).

The model combines several features to allow for a rich pattern of household demand and substitution patterns across a range of electricity sources. First, we estimate the model using micro-data on household characteristics. Second, we allow for unobserved heterogeneity in source quality at the technology-by-village-by-time level, and so can study the value of quality improvements over time. Third, we estimate the household sensitivity to price in the model using random variation in the price of solar microgrids from the experiment.

In the estimation of this demand model is where the random variation in price of solar microgrids induced by our experiment becomes most valuable. This is because in discrete choice models with unobserved product quality, price is an endogenous variable. Typical solutions in the literature use either cost shifters (prices in other markets) or mark-up shifters (characteristics of other products in the same market) as instruments to identify the effects of prices. Our experimental setting allows us to use dummies for the two experimental treatment arms (a normal price of INR 200 and a subsidized price of INR 100 for micro-grid) as instruments for the price. Moreover, we can also obtain an instrument for peak and off-peak hours of supply in a given village using supply to other villages nearby. To the best of our knowledge, this paper represents the first instance of an experimental RCT design being used to estimate a discrete choice demand model.

3.2.1 Specification

Utility for household i in village v from electricity source j in survey wave t is is given by

$$U_{ijtv} = \delta_{jtv} + \mathbf{z}'_{it}\gamma_j + \epsilon_{ijt}. \tag{2}$$

The term δ_{jtv} represents mean utility of an electricity source j in village v and survey wave t and is modeled as

$$\delta_{jtv} = \mathbf{x}'_{jtv}\overline{\beta} + \xi_{jtv}, \tag{3}$$

where $\mathbf{x}_{jtv}$ is a vector of observable electricity source characteristics including price and hours of supply (on- and off-peak), which enter utility through an average effect $\overline{\beta}$, and where ξ_{jtv} represents the unobservable electricity source characteristics (including quality).

The vector $\mathbf{z}_{it}$ contains all observable household characteristics including the number of adults, income, land, literacy, the number of rooms and the solidity of the house and the roof. These characteristics enter utility with a choice-specific coefficient γ_j, which captures how household observables affect utility relative to the mean for that source, village and survey wave. For example, as income increases, households may have a greater preference for grid electricity but an unchanged preference for diesel.

We assume that the unobserved idiosyncratic tastes for electricity source j, $(\epsilon_{i1t}, \ldots, \epsilon_{iJt})$, are iid across households and survey waves. Moreover, for the model to have a nested logit specification, we require the joint distribution of $(\epsilon_{i1t}, \ldots, \epsilon_{iJt})$ to be GEV with CDF

$$F(\epsilon_{i1t}, \ldots, \epsilon_{iJt}) = \exp\left[-\sum_{g=1}^{G} \Big(\sum_{j\in\mathcal{J}_g} e^{-\epsilon_{ijt}/(1-\sigma_g)} \Big)^{1-\sigma_g} \right].$$

Each electricity source j belongs to a nest, indexed by g. The parameters σ_g measure nest-specific similarity of sources within a nest. If two sources j and k belong to the same nest g, the correlation between the unobserved idiosyncratic tastes is approximately $\text{corr}(\epsilon_{ijt}, \epsilon_{ikt}) \approx \sigma_g$. For two sources j and k belonging to different nests, we have $\text{corr}(\epsilon_{ijt}, \epsilon_{ikt}) = 0$ as in a standard multinomial logit, and the IIA property holds for those. As σ_g approaches one, idiosyncratic variance in utilities comes mostly from the nest level, not from distinctions between sources within a nest. As σ_g approaches zero, the distinction between nests disappears and the model becomes a multinomial logit.

Let y_{it} denote household i's electricity source choice in survey wave t. Conditional on choosing within nest g, the probability of choosing a source $j \in \mathcal{J}_g$ is

$$\Pr(y_{it} = j \,|\, y_{it} \in \mathcal{J}_g) = \frac{e^{(\delta_{jtv}+\mathbf{z}'_{it}\gamma_j)/(1-\sigma_g)}}{\sum_{k\in\mathcal{J}_g} e^{(\delta_{ktv}+\mathbf{z}'_{it}\gamma_k)/(1-\sigma_g)}}. \tag{4}$$

Notice the similarity with a multinomial logit.

Define the inclusive value of nest g as

$$IV_{igt} = \ln \sum_{j\in\mathcal{J}_g} e^{(\delta_{jtv}+\mathbf{z}'_{it}\gamma_j)/(1-\sigma_g)}.$$

This is equal to the expected indirect utility when maximizing utility across sources in nest g (up to an additive constant). With that, we can obtain the probability that the household chooses within nest g:

$$\Pr(y_{it} \in \mathcal{J}_g) = \frac{e^{(1-\sigma_g)IV_{igt}}}{\sum_{k=1}^{G} e^{(1-\sigma_k)IV_{ikt}}}. \tag{5}$$

Putting the above results together, we have that the unconditional probability of choosing a source j in nest g is

$$\Pr(y_{it} = j) = \Pr(y_{it} = j \,|\, y_{it} \in \mathcal{J}_g)\Pr(y_{it} \in \mathcal{J}_g) = \frac{e^{(\delta_{jtv}+\mathbf{z}'_{it}\gamma_j)/(1-\sigma_g)}}{e^{\sigma_g IV_{igt}} \sum_{k=1}^{G} e^{(1-\sigma_k)IV_{ikt}}}. \tag{6}$$

The expected indirect utility for the household after the maximization of utility across nests is equal (up to an additive constant) to

$$\mathbb{E}[\max_j U_{ijtv}] \approx \ln \sum_{g=1}^{G} e^{(1-\sigma_g)IV_{igt}}. \tag{7}$$

We can use this quantity to study how consumer welfare changes when we restrict the choice set or change electricity source prices and hours of supply.

3.2.2 Estimation

We estimate the model in two stages. The first, nonlinear stage estimates the parameters of equation 2 via maximum likelihood. The second, linear stage estimates uses the $\hat{\delta}_{jtv}$ from the first stage as the dependent variable to estimate equation 3 using two-stage least squares. This two-step procedure is common in the estimation of random coefficients logit models (Berry, Levinsohn and Pakes, 1995, 2004). The key idea is to invert market shares to solve for unobserved mean indirect utilities, allowing for linear IV estimates that are unbiased in the presence of the endogeneity of price to quality (Berry, 1994).

There are two ways in which our estimation differs from common practice. First, since we have very rich household observable characteristics, we use a simpler nested logit model rather than a mixed logit model, which is often essential with aggregate data to model rich enough patterns of consumer substitution. Using the nested logit also means we can estimate the first stage by maximum likelihood for some gain in efficiency.[23] Second, in our application, we have much richer cross-market variation: we observe 100 villages in 3 survey waves. That means that we can estimate the average effects of source characteristics far more reliably than in applications that observe only regional or national markets.

Nonlinear estimation of the first stage In the first stage, we use maximum likelihood to estimate the parameters δ, γ and σ using equation 6. The log-likelihood of the sample is

$$\log \mathcal{L}(\gamma, \sigma \,|\, \mathbf{y}, \mathbf{z}) = \sum_{j=1}^{J} \sum_{t=1}^{T} \sum_{i:\, y_t = j} \log \Pr(y_{it} = j \,|\, \mathbf{z}_{it}; \gamma, \sigma, \delta(\gamma, \sigma)). \tag{8}$$

We write$\delta(\gamma, \sigma)$ to indicate that we concentrate the δ parameters out of the log-likelihood using the contraction mapping proposed by Berry, Levinsohn and Pakes (1995). To do so, we consider each candidate parameter vector (γ, σ) and we solve for the δ that exactly fits the aggregate market shares. This greatly reduces the dimensionality of the problem, as we would have up to 4 sources $\times$ 100 villages $\times$ 3 surveys = 1200 δ parameters if every source were available in all villages.

Because the source characteristic effects are absorbed into the mean utility δ parameters, the maximum likelihood estimates are consistent regardless of the possible endogeneity of price at the market level.

Linear estimation of the second stage We can now use equation 3 to recover the $\overline{\beta}$ parameters via a linear regression of the estimated $\widehat{\delta}_{jtv}$ on the observable characteristics of the electricity source at the survey wave and village, $\mathbf{x}_{jtv}$.

[23] In principle, one could also estimate a mixed logit model using maximum likelihood, but this approach may be biased by simulation error (Berry, Levinsohn and Pakes, 2004). In the nested logit model we specify, we solve for market shares analytically, so maximum likelihood will be preferred to GMM on efficiency grounds.

Because the price is an endogenous variable when estimating demand systems, we use 2SLS, instrumenting price by dummies for the two experimental treatment arms (normal microgrid price of INR160, subsidized microgrid price of INR100). Since we are only estimating an average effect and the disutility of price is common across all sources (we only have one coefficient), we can use the randomized variation in levels of subsidies as a valid, exogenous instrument.

We are also concerned that the hours of supply on the grid may be endogenous to demand. To account for that, in some specifications we also instrument the hours in a village (both on- and off-peak) by a predicted supply using nearby villages. The logic of this leave-one-out estimator is that supply in a given village may be affected by supply in those villages nearby, for instance, due to common rationing rules if they are served by the same substation. In our setting, a supply instrument based on nearby villages is particularly adequate because the structure of the distribution grid does physically connect the supply decisions within a region. The exclusion restriction is that supply of electricity in nearby villages is not correlated to the determinants of demand in a given village. An example where this restriction would be violated is if there are common unobserved demand shocks across nearby villages, conditional on our rich set of household observables. However, this type of supply shocks under the assumption of no common regional demand shocks are usual in discrete choice applications (Nevo, 2000) The detailed construction of the supply instrument is presented in the appendix.

The residuals in the second-stage regression estimate the unobserved component of mean utility:

$$\widehat{\xi}_{jtv} = \widehat{\delta}_{jtv} - \mathbf{x}'_{jtv}\widehat{\overline{\beta}}.$$

These terms allow us to observe how unmodeled characteristics of electricity sources vary across sources, villages and time. While electricity connections may appear homogeneous, allowing for unobserved quality is important to capture aspects such as the load a connection can serve, the ease or difficulty of obtaining a connection, and the marketing and service associated with a given source. These characteristics all vary widely across the sources that compete in the Bihar market.

Simulation Once we have estimated the parameters in our demand model, we can use it to simulate counterfactuals. We are mainly interested in analyzing the effects on market shares and welfare of changes in the availability and the price of different electricity sources.

The aggregate market share of electricity source j belonging to nest g can be simulated by an average of the household-level predicted probabilities in a given survey wave:

$$\widehat{s}_{jt} = \frac{1}{N}\sum_{i=1}^{N} \frac{e^{(\widehat{\delta}_{jtv}+\mathbf{z}'_{it}\widehat{\gamma}_j)/(1-\widehat{\sigma}_g)}}{e^{\widehat{\sigma}_g \widehat{IV}_{igt}} \sum_{k=1}^{G} e^{(1-\widehat{\sigma}_k)\widehat{IV}_{ikt}}}, \qquad \text{where } \widehat{IV}_{igt} = \ln \sum_{j\in\mathcal{J}_g} e^{(\widehat{\delta}_{jtv}+\mathbf{z}'_{it}\widehat{\gamma}_j)/(1-\widehat{\sigma}_g)}.$$

The expected household-level indirect utility can be simulated by

$$\widehat{\mathbb{E}}[\max_j U_{ijtv} \mid \mathcal{J}] = \ln \sum_{g=1}^{G} e^{(1-\widehat{\sigma}_g)\widehat{IV}_{igt}},$$

where$\mathcal{J}$ denotes the set of available choices. Hence, if we consider a restricted set of choices $\mathcal{J}'$, a measure of the willingness to pay to have the full set of alternatives $\mathcal{J}$ is the change in expected utilities divided by the coefficient on price. This transformation converts from units of utility to monetary units. The average willingness to pay is therefore:

$$\widehat{WTP} = -\frac{1}{N}\sum_{i=1}^{N} \left(\widehat{\mathbb{E}}[\max_j U_{ijtv} \mid \mathcal{J}] - \widehat{\mathbb{E}}[\max_j U_{ijtv} \mid \mathcal{J}']\right) \Big/ \beta_{price}.$$

Both the estimated market shares and the willingness to pay will be objects of central interest in the analysis of our policy counterfactuals (Section 4).

3.2.3 Results

This section reports estimates of household demand for electricity sources. We then use these estimates to run counterfactual simulations to value changes in the technology, availability, and pricing of various sources of electricity.

The full demand model has some 1033 parameters: 1000 technology-by-village-by-survey specific mean indirect utility parameters backed out from the first-stage demand model, 29 parameters governing household heterogeneity, 3 parameters on the average effects of source characteristics and a parameter governing correlation of the source-specific utility shocks. We therefore report only select parameters to give a sense of how the model represents household technology choices. First, we present estimates from the non-linear estimation of how household characteristics affect their choices. Second, we describe the linear estimates of the average effects of source characteristics. Third, we present distributions of source quality to characterize quality changes over time.

Table 6 shows the estimated effects of household characteristics on choice probabilities in the demand model. The choice probabilities are derived from the estimated coefficients of the demand model, reported in Table 7. The effects of household characteristics on choice probabilities are non-linear; we evaluate these effects for a poor household, which we define as a household of two adults with a one-room house, without a solid roof or walls, that does not own agricultural land. (The full profile of a poor household's characteristics is in Appendix Table A3.) The table shows how the household's probability of choosing each electricity source (across columns) varies with a change in characteristics, either from zero to one (for discrete variables) or of one standard deviation (for continuous variables). Similar tables for median and rich households are presented in Appendix Tables A4 and A5.

The main finding of the table is that richer households, by any measure, have stronger preferences for grid electricity over all other sources. For example, our representative poor household has a baseline 21 percent probability of choosing grid electricity. If the household had a solid roof, the probability of choosing grid electricity would increase by 11 percentage points (standard error 2.5 pp). Similarly, increases in the number of household adults, household income, land holdings, literacy, house quality or the number of household rooms all have positive, economically meaningful and statistically significant effects on the household probability of choosing grid electricity, and all also reduce the probability a household chooses no electricity (the outside option). Some household characteristics also increase household choice probabilities for other inside goods; for example, households with higher incomes are more likely to choose solar microgrids. The effects of household characteristics on demand for the other inside goods, however, is much less pronounced than on demand for the grid. Table 2 offers a natural interpretation of this finding: grid electricity offers higher load, and many more households on the grid can run a fan or a television. Richer households want the energy services these devices bring.

The above estimates show how household characteristics change household mean indirect utilities, relative to the mean utility for a source for an average household. We now consider how source characteristics change the mean household indirect utility. Table 7 reports estimates of the linear part of the demand model, obtained by regressing the mean household indirect utility, recovered from the fist stage, on source characteristics. Column 1 estimates the linear part of the model by OLS and column 2 instruments for source price using the experimental variation in microgrid prices. Column 3 instruments for price and also instrument for hours of supply, using the imputed hours of supply based on the supply to nearby villages.

The main finding from the linear part is that price has a large, negative effect on mean household indirect utility (column 3). The estimated coefficient associated with a INR 100 price increase is -2.08 (standard error 0.74). To give a scale to this number, the average probability of choosing the grid is 24 percent, and the model estimates imply that a INR 10 increase in the grid price (17 percent of the mean price of INR 59) decreases grid market share by 2.9 pp (12 percent of the average share). The elasticity of grid market share with respect to price is therefore -0.71. We interpret this as a big change in price; INR 10 is enough money to buy two cups of tea or three bananas, but raising the grid price by this amount in a month cuts market share by a noticeable 2.9 pp. A similar calculation finds the elasticity of own solar market share with respect to price is -3.

The experimental variation in subsidy level is critical to identifying the household price response. In an analogous ordinary least squares specification, we find a small, negative and insignificant effect of price on mean indirect utility (column 1). This discrepancy in results suggests that the prices of some electricity sources are endogenously higher in villages with higher demand for those sources. The experimental instrument has a strong first stage despite the fact that the price variation only applies to one electricity source (column 4). As expected, being in the normal (unsubsidized) price group raises price, whereas being in the subsidized price group lowers price.

In Appendix Table A.2 we present the estimated price-elasticities for each source implied by our model. To compute them, we simulate the aggregate market share that each source would have if every household and source covariate was kept at endline 1 levels but the price of that source (and only of that source) was increased by 10% with respect to the endline 1 average. The figure for the price-elasticity of each source is then obtained as the percentage change in aggregate market share of that source divided by 10%. It is worth mentioning that demand for grid electricity is estimated to be much more inelastic than for other sources of electricity. A 1% increase in the price of grid would decrease its market share by 0.74% at the endline 1 levels, whereas a 1% increase in the price of own solar or microgrid solar electricity would decrease their market shares by 3.48% and 2.34% respectively. This seems to indicate that at least some consumers are tethered to grid electricity. A possible explanation is that grid is the only source capable of generating enough power to support large appliances like televisions and fans, which are widely used by richest households.

We also estimate coefficients on the hours of supply that a source offers both on peak (in the five evening hours from five to ten) and off-peak (all other hours). Running OLS, we find a positive but statistically insignificant effect of peak hours on supply and a smaller, negative coefficient of off-peak hours (column 2). The estimate of the value of peak hours is not precise, but agrees with the idea that agricultural households mainly value light in the evening hours. In the column 3 specification, we additionally instrument for hours of supply using imputed hours of supply based on supply to nearby villages. We find that the coefficient on price is unchanged and the coefficient on peak hours is positive but not statistically different from zero. The point estimate is somewhat larger than the estimate without instrumenting for hours, from column 2, but we could not reject that the value of peak hours is the same in the two specifications. We proceed with the column 3 estimates, instrumenting for price and hours, as our main specification for counterfactuals, on the grounds that supply may be rationed in part based on village latent demand, which argues for the instrumental variables approach on *a priori* grounds. We report results from the column 2 specification as a robustness check.

The demand model allows flexibly for changes in ξ_{vtj}, the unobserved mean quality of electricity across villages and time. Figure 7 summarizes these source qualities by plotting histograms of quality by source (each row is one source) and time (each column is a different survey wave). Within each source and wave, the histogram shows the distribution of source quality across villages.

The figure shows how the landscape of electrification in Bihar shifted, with the grid and own solar systems gaining market share, in a relatively short period of time. Stagnant technologies do not improve quality. The distribution of diesel generator quality, for example, is about the same in all three survey waves (there is some truncation at the bottom, due to exit). The microgrid provider, HPS, in the experiment, did not offer their product in many villages at baseline (by design), and did not change their product between our first and second endline surveys. The stasis of the microgrid product is apparent in the figure, as the distribution of quality is similar between the second and third survey waves (HPS row, right-hand two panels). Contrast diesel and HPS microgrids with own solar systems, which shifted up in quality in each survey wave. These improvements could be due to improvements in technical factors such as battery capacity and load, which we do not model directly, or to a broader reach of marketing and distribution of these systems. Finally, we see large improvements also in the quality of the grid, especially between the second and third survey waves. These improvements show the results of a government drive to increase household connections, which may have increased access and therefore the estimated quality of the grid.

The increases in electrification over the period of study were therefore due in part to solar innovation and improvements in quality, in part to falling prices for solar systems and in part to a rapid expansion in the availability and quality of grid electricity. We now use the model to break down the contribution of these factors and to value their contribution to household surplus.

4 The Value of Electrification

The estimated demand system gives us household willingness to pay for electricity from a variety of sources. This enables us to address the central question in the paper which concerns how households value different sources of electricity. This will depend not only on household characteristics (e.g. rich, poor), or source characteristics (e.g. price, hours of supply) but also on the choice set that a household faces. The demand system we built in Section 3 allows us to model the complex pattern of substitution between different sources of electricity as these factors change.

This, in turn, enables us to carry out out-of-sample simulations of how different policies affect the valuation of different sources of electricity. This is an invaluable tool in a situation where there is large-scale change in the electricity market (see Section 2.1). We can begin to look at the effects of solar power becoming cheaper, of the grid expanding or of the state forcing citizens to pay for the electricity they consume. Gaining these insights is critical not only for guiding policy in Bihar which contains around 125 million people but also for finding effective policies to bring electricity to the billion or so people who remain without.

To examine the value of electrification we proceed in to two steps. We begin, in Section 4.1, by examining how households substitute between different sources of electricity as the price of off-grid solar energy falls or grid electricity becomes more available. Both of these factors could be influenced by policy, for example, by subsidising off-grid solar or the extension of the grid to rural areas. Then in Section 4.2, we look at three specific out-of-sample policy counterfactuals. In the first (Section 4.2.1) we model how the price of solar panels and batteries might fall in the future and examine how this influences market shares and willingness to pay for different sources. Working out the magnitude of the contribution that off-grid solar might make to electrification efforts in the developing world represents a focal area of policy interest at the moment. In the second (Section 4.2.2), we vary the availability and hours of supply of grid electricity and see how this affects market shares and willingness to pay for all electricity sources. In the third (Section 4.2.3), we looking at grid pricing policy, specifically at the effects of removing theft and forcing people to pay for the full price of the electricity they consume. Grid electricity in our sample setting is characterised by heavy financial losses, restricted access and rationing. Our second and third policy counterfactuals provide insights into the underpinnings of dysfunctional grid electricity and how these might be remedied. This is a major policy challenge across the developing world (Burgess et al., 2019).

4.1 Patterns of Substitution

How do households choose between different sources of electricity? Figure 8 shows how the market shares of all electricity sources respond to changes in the price of solar. The prices, on the horizontal axis, range from INR 70 per month in the reduced capital cost scenario (see Appendix Figure A1), up through the range of our experimental treatments, to a near-choke-price of INR 300 per month. Own solar capital costs are varied proportionally with HPS costs since these sources have similar cost structures. The vertical axis shows market shares for each source technology.

Further reductions in the price of solar would increase its market share moderately. At the lowest subsidized price from the experiment, 8 percent of households adopt HPS in the model and 12 adopt own solar, thus solar has a total market share of 20 percent. Setting solar prices at 2022 projections (see Appendix Figure A1), solar adoption increases to 32 percent. Most of this gain in market share comes from households that would not otherwise be electrified (solid line at top) rather than substitution from other sources.

For example, if we consider an HPS price cut from the 2013 control group price INR 200 to the 2022 projected price of INR 70, HPS market share increase by 26 pp, and the share of households without electricity declines by 25 pp. New solar customers are therefore not coming from other sources of power but from customers without any access to electricity. If the price of solar were raised to the choke price of INR 300 per month, the share of households without any source of electricity would rise from the 57 observed to 65 percent. Solar power therefore makes a significant contribution to bringing power to unelectrified households.

The market share and value of solar therefore hinge on the extent to which the grid is available as a substitute. This is seen in Figure 9. Panel A shows how the market shares of solar and the grid change as the extent of the grid varies. Solar achieves a 33% market share when the grid is absent from all villages, which is cut by nearly half if the grid is present in all villages. There is a wide range of grid extent spanned in our surveys, from 29% of villages having grid in the baseline to 53% in endline 1 and 72% in endline 2. In the model, holding constant household and source characteristics, the increase in grid presence between baseline and endline 1 alone accounts for an 11 pp (40%) drop in solar market share from around 27% to 16% (Figure 8, Panel A).

Figure 9, Panel B shows that the presence of the grid similarly diminishes the value of solar power. As described in Section 3.2.2, in order to obtain a measure of willingness to pay, we calculate the expected indirect utility from the bundle of offered electricity sources in our model, recalculate this quantity without solar power in the choice set, and then normalize the difference in expected indirect utility by the estimated price coefficient to get household-specific willingness to pay. The figure reports the value of household WTP averaged over all households for four different extents of the grid. Without the grid, household willingness to pay for the presence of solar in the choice set is INR 1032 per year (1.4% of median annual household income); if the grid is everywhere this drops to around INR 544 (0.8%). The full availability of grid electricity cuts willingness to pay for solar power by half.

4.2 Policy Counterfactuals

Table 8 uses the model to address questions on three themes. First, what is the value of innovation in solar for the poor? Second, what is the value of improving the quality of the grid? Third, how would changes in grid policy to deter theft affect electrification and willingness to pay? Every row of Table 8 is a different counterfactual corresponding to a scenario described in the subsections below. Columns

1 through 5 give the market shares of each technology in the row scenario and columns 8 through 10 measure consumer, producer and total surplus. We take consumer and source characteristics as of the first endline survey. Consumer surplus in each scenario is measured relative to the surplus consumers derive from having no electricity, assumed to be zero. We separately report mean consumer surplus for below poverty line (BPL), above poverty line (APL) and all consumers in columns 6 through 8. In our sample 76% of consumers are BPL, which is the government's official designation of poverty. BPL consumers are modestly but significantly poorer on all the household observable characteristics that enter the demand model (Appendix Table A6). Since the model allows for heterogeneity in demand by household characteristics, surplus for BPL and APL consumers may differ despite that we do not condition demand directly on BPL status.

Producer surplus is the absolute surplus we estimate for the grid only, based on its cost structure. It captures profits or losses that accrue to the state from supplying grid electrcity. The producer surplus can be taken as comprehensive if we assume that the other sources are competitively supplied; this is probably accurate for own solar but not for diesel (which, in any case, generally has a small share). All surplus measures are per sample household per year.

Table 8, Panel A shows that the model fits sample market shares. The exact fit is by design since the source-by-village-by-wave specific quality measures are estimated to match market shares. In both the data and the model, 57 percent of households have no electricity connection. The grid is the most popular source of electricity, with a 24% share, but the two types of solar together are not far behind, with a 17% share. Consumers have a mean surplus of INR 1939 (USD 30) per year in this scenario and the state grid a surplus of *negative* INR 497 (USD 8). The grid is heavily subsidized and many consumers steal; the average payment for the grid of INR 59 per month is about 40% of the variable cost of supply. The government therefore loses a significant amount of money on every customer connected to the grid.

4.2.1 Counterfactual 1: The Value of Solar Innovation

Panels A and B of Table 8 show that solar makes a meaningful contribution to electrification and reduces the financial losses of the state grid. The same results are displayed graphically in Figure 10. From the benchmark scenario, if we remove solar altogether (Panel B, row 1), the share of households without electricity increases by 11 percentage points, from 57% to 68%. That increase is less than the endline 1 solar market share since households substitute to diesel and the grid. As the grid market share increases by 5 percentage points the losses from the grid increase in magnitude to negative INR 612 per year (23% larger).

During the sample period the market share of solar power increased by 12.5 percentage points, or about threefold. The price of solar has continued to come down over time and the quality has continued to improve. It is therefore relevant to ask what is the value of innovation in solar in our data and whether, even if households prefer the grid today, further innovation is likely to generate greater willingness to pay for off-grid solar.

Even rapid falls in the price of solar panels bring down the cost of an installed solar system only so much, since an installed system has costs other than capital. To discipline the possible value of innovation we obtained administrative estimates from our partner HPS on the cost of their solar system. Appendix Figure A1 shows the installation cost of the system in units of amortized USD per month broken down by the cost for each component of the system. The total cost is about INR 114 (USD 1.73) per month, which exceeds our lowest subsidized price of INR 100 per month. At the normal price of INR 160, and without allowing for any variable costs of labor and collections, which are certainly not negligible, this cost basis would imply a profit margin of 40%. The capital component of the cost, which we consider the object of innovation, is USD 0.62 per month for the panel (36% of the total) and USD 0.33 for the battery (19%), thus 55% in total.

Counterfactually, we consider reductions in cost for solar photovoltaics and for batteries. For solar PV, we assume a 55% reduction in cost in line with the National Renewable Energy Laboratory's projections for 2022 (Feldman, Margolis and Denholm, 2016). For batteries, we assume that they fall in cost by 75% in accord with the US Department of Energy's 2022 goal (Howell et al., 2016). Since panel capital and batteries only make up part of a system, these changes imply a reduction in price by 34% to USD 1.12, 30% cheaper than our subsidy treatment. We apply the same proportional reduction in price to own solar, conservatively assuming that the 55% capital component of total cost observed for HPS applies to own solar as well. This is conservative since the HPS product involves monthly recharge costs that do not apply to the own solar model. We can also consider a range of other possible solar prices, but take this feasible reduction in cost as a benchmark.

Taking a longer view, with solar nowhere as the baseline, we state three values for solar innovation. The first value is from no solar (Panel B, row 1) to the first endline observed in our data (Panel A, row 2). We estimate that solar increased household surplus by INR 746 (USD 11) per household in our sample (column 8). The increase in surplus due to solar is slightly higher for the poor, though they have lower demand for electricity, because richer households prefer the grid where it is available (the column 6 difference in surplus across rows is INR 762, compared to a difference of INR 695 in column 7). The second value considers further solar innovation if prices were cut along our optimistic path for 2022 innovation. The total solar market share increases to 33 percentage points, surpassing the grid and raising consumer surplus a further INR 374 per year, for a total gain, relative to the absence of distributed solar, of INR 1120 per year (USD 17). To put this gain in perspective, the average *monthly* income in our sample is INR 7,576 and the median is INR 6,000. APL households have consumer surplus about a fifth higher than BPL households in this scenario.

Finally, solar may also improve on non-price dimensions such as the capacity of batteries or the distribution network of own solar systems. To capture possible increases on these dimensions, we take the increase in unobserved quality of solar systems, shown in Figure 7, and project this increase out linearly for three more years (from the 2017 endline 2 survey to 2020). These improvements in solar quality are hypothetical and do not have a clear technological basis as do the reductions in cost. We include this scenario to understand what the upper bound of off-grid solar penetration may look like. The final row of Panel B considers hypothetical increases in solar quality, along the trend observed in our data. Increases in solar quality roughly double the consumer surplus from electrification. However, we consider this scenario an upper bound on possible gains, since it is not clear if solar innovation can maintain the rapid pace we observed in our sample.

4.2.2 Counterfactual 2: Improving the Quality of the Grid

The experiment is based in a setting where the grid is deficient: it is offered only in some villages and even there the average supply is only 11 hours per day. The effects of solar on electrification rates would have been different if the grid was everywhere, so households had better substitutes, or if the grid has non-existent, as is the case in some parts of India and much of rural Sub-Saharan Africa.

We predict demand under these scenarios by removing grid electricity from the choice set or by extending it to all villages. The counterfactual results are presented in Panel C of Table 8, and shown graphically in Figure 11. When the grid is not present in a village, we cannot estimate an unobserved quality for the grid in that village; we therefore impute the unobserved quality as the mean unobserved quality for the grid in villages with the grid in that survey wave.

Improvements to the availability and quality of the grid would increase electrification, with an effect similar in size to that of further innovation from off-grid solar. By the time of our first endline the grid had reached only 53 percent of sample villages. If the grid were removed from all villages, the share of households without electricity, from any source, would rise from 57 to 63 percent, and household consumer surplus from all sources would fall by 43% (Panel C row 2). If the grid were extended to all villages, it would increase electrification by 13 percentage points and household surplus from all electricity sources by 26% (Panel C row 3). Even after this grid extension, we estimate that 44% of households would remain without any source of electricity. This finding shows the challenge of achieving universal electrification in a setting with a high concentration of poor people.

We also simulate improvements in grid supply that increase the duration of power available by two hours per day (or to five hours, the number of peak hours per day, if a two hour increase would exceed five peak hours in total). This allow us to compare intensive and extensive margin policy changes. If the grid were to offer two more hours of power during the peak period, this would bring 9 percentage points more households onto the grid. The poor quality of the power grid therefore has a significant effect on whether households get an electricity connection.

4.2.3 Counterfactual 3: Reducing Grid Theft

If a deficient grid might reduce the appeal of the grid and inflate the market share of solar, the policy towards electricity pricing will do the opposite. Electricity in Bihar is heavily subsidized and the state tolerates a high level of theft, which lowers effective prices further still. If power on the grid was priced at cost improvements in solar may have been even more valuable to households.

Counterfactually we raise the price of grid electricity from INR 59 per month, the sample mean in the endline 1 survey wave, to INR 140, which we calculate would be sufficient to cover the variable costs of supply. This price change will shift surplus from consumers to the government and change household take-up of electricity connections from the grid and other sources. We also consider a budget neutral reform that would increase supply during the peak hours and pay for that increase by increasing prices. We calculate that a price increase to INR 90 per month would be sufficient to pay for the increase in supply. The results are presented in Panel D of Table 8, and displayed graphically in Figure 12.

Raising the price of the grid to cover the cost of supply would devastate grid market share. In Panel D, row 1, we simulate raising the grid price from INR 59 to INR 140 per month, which would cut demand for the grid from 24 pp to only 5 pp, with many household substituting towards solar power instead, and cut consumer surplus by 33%. The fall in consumer surplus from pricing power at cost is not much smaller than the fall in surplus due to removing the grid altogether (43%). Producer surplus, on the contrary, would surge from a loss of roughly INR 500 per household per month to nearly break-even. Total surplus from this price change is therefore moderately lower (9%) lower than in the status quo, though there is a dramatic shift in surplus from consumers to the state and many consumers fall off the grid. This large swing in market share may seem extreme, but it is consistent with our experimental estimates, in particular the high price sensitivity observed for solar around the price (at INR 100 per month) of competing options like diesel power (Figure 6).

Could the government do better than such a pure transfer, by trying to increase prices and quality at the same time? Our fourth main finding is that the scope for this kind of grand bargain appears limited. We estimate that increasing the supply on the grid up to two hours in the peak, and raising prices enough to just cover the cost of this additional supply, would yield about the same total surplus (Panel D, row 2, column 8, compared to panel A, row 2, column 8). This bargain would increase the electrification rate by 5 pp but, paradoxically, not decrease grid electrification marginally. The logic, in the model, is that the increase in peak hours brings some consumers onto the grid, even as the increase in price drives more price sensitive consumers off the grid towards solar. Therefore there is a a negative net flow of consumers from the grid, compared to the baseline scenario in Panel A, but the total electrification rate rises.

The interpretation that the grand bargain roughly breaks even depends on the model specification we use to a greater extent than do the other counterfactual findings. We estimate the value of price to consumers using our experiment, but do not have as strong an instrument for hours of supply. Because our estimate of the value of peak supply is imprecise, the conclusion depends on the exact specification of the demand system. For example, if we instrument for price but do not instrument for hours, in the second, linear stage of the demand model, then the grand bargain increase in quality and price would decrease total surplus, due to the lower household valuation, in that specification, for additional peak hours of supply. The general conclusion is that our main result from the model, that the grand bargain breaks even in surplus terms, is perhaps slightly optimistic.

The value of solar power depends with great sensitivity on grid policy, namely availability and pricing. In particular, the value of solar is much greater the weaker is the grid on availability and price. Solar also appears to be pro-poor: BPL households have a stronger preference for solar and APL households for the grid, and so the value of adding solar in the choice set is relatively high for households that would not have used the higher load the grid offers. On price, we may consider that competition from the grid, which is heavily subsidized through both explicitly low prices and a tolerance for theft, is unfair in this setting. A fairly extensive grid with cheap power and connection fees contrasts with some African countries where the grid is largely absent and prices are higher (World Bank, 2017).

5 Conclusion

Electricity markets in the poorest parts of the world are undergoing radical changes. On the one hand, darkness is still a pervasive problem — roughly a billion households mainly in South Asia and Africa remain without any access to electricity at all. On the other hand, consumers enjoy a growing number of choices on where they source electricity from – technological innovations are making off-grid solar a viable alternative to the grid and governments are prioritizing extending the grid to poorer, rural areas. The main contribution of this paper is to create a demand system founded upon experimental variation in solar prices that allows us to assess household willingness to pay for different sources of electricity. By using demand estimation we can assess how households of different types value different sources of electricity. In this way we can assess whether the darkness problem in the rural areas of developing countries just reflects a failure in expanding access to electricity or rather an unwillingness of poor, rural households to pay for electricity relative to other demands on their budgets.

Access to electricity from the grid is taken for granted in developed countries but vast swathes of communities in the developing world either lack access to it altogether, or receive rationed, intermittent supply. In rural Bihar, one of the poorest parts of the world, our data collection efforts reveal that this scarcity of formal grid has engendered a rich, dynamic parallel market for alternative sources of electricity. Our surveys, conducted in 100 villages in India's poorest state, uncover innovations in off-grid solar technologies and government-led grid expansions which imply that households face a growing choice set on where to source electricity from. Where there is sufficient demand, markets for diesel generators have come up. Still others choose to buy their own solar panels but at a substantial cost. New technologies such as microgrid solar are being introduced into this competitive market. Our study was well timed to capture this period of dramatic change in the electricity market. This region had epitomised the problem of darkness enveloping the developing world — at baseline, just 27% of households within our sample villages had access to electricity and electrification rates were below those in sub-Saharan Africa. Just four years later that number had jumped to 64% with the expansion of the grid and the spread of own solar playing key roles.

In this fast evolving setting we carry out an RCT varying the price and availability of a new off-grid technology - microgrid solar. In partnership with HPS, we randomly assign the availability and price of microgrids across 100 villages and monitor the evolution of demand for this electricity source and its competitors via an innovative survey instrument over three waves. This allows us to trace out a demand curve for microgrids. However, this paper's key innovation arguably lies in our subsequent combination of the experimental price variation created by the RCT with a rich dataset comprising both demand and supply-side information. This enables us to estimate a multinomial demand model which describes the electricity market as a whole, characterizing both market shares and household willingness to pay for the main sources of electricity in rural Bihar (grid, diesel, own solar, microgrid solar). We then employ our model to construct counterfactuals, yielding valuable out-of-sample predictions for how the market is likely to respond to future innovations in solar technologies and potential changes in the functioning of the electricity distribution industry.

Our reduced-form estimates point to the high price sensitivity of consumer demand for microgrids, with demand dropping to near 0 at unsubsidised prices. Moreover, evidence suggests demand shifted left over the course of the experiment, meaning only 7% of households were paying even at the subsidised price at endline 1. This is in line with the strong and increasing competition observed in electricity markets. Despite low take-up, households in subsidised treatment villages in particular were significantly more likely to own light bulbs, use more electricity, purchase more mobile phones and spend less money charging them. We were largely unable to detect any benefits to household income, children's reading and math test scores or self-reported respiratory problems. Overall, this supports the view that solar microgrids are valuable for certain households but not transformative in the way that some have hoped.

One of our multinomial demand model's key result is that richer households, by any measure, have stronger preferences for grid electricity than any other sources. This is likely attributable to the grid's unique ability to support larger appliances such as televisions and fans, which richer households demand. It also substantiates the reduced-form result that households are highly price elastic — raising the price of the grid by just two cups of tea per month is estimated to reduce the electrification rate by 3 percentage points.

The finding of high elasticity implies that anticipated innovation in solar technology and small changes in government policy, which in turn affect the characteristics of different electricity sources, may have dramatic effects on the electricity market. These are explored in our counterfactual analysis. We find that further reductions in solar prices would moderately increase its market share, mostly arising from adoption by households that would not otherwise be electrified. However, when solar is the only electricity source available, its market share is limited to 34% and almost two thirds of households choose to remain in darkness Indeed, willingness to pay for off-grid solar is considerably lower if the grid is available, so that it appears to be more of a stop gap – albeit a potentially important one – for households who do not have access to grid or for poor households who cannot afford full-cost grid electricity. In this sense, solar power is valuable in large part because the grid is incomplete and dysfunctional.

Another key contribution of this paper is to rationalize India's existing system of simultaneously subsidizing and rationing electricity with rampant theft via informal connections to the grid. In particular, our counterfactual analysis predicts that demand raising prices or cracking down on non-payment would devastate household electricity access among the rural poor — households are highly price elastic and would simply opt to live in darkness. This is at odds with the view that households value electricity and benefit from the extensive efforts undertaken in developing countries to expand access to electricity, whether this takes place via on or off-grid mechanisms.

Our demand model suggests an alternative interpretation of enthusiasm for solar power in developing countries for off-grid, small-scale use. Generating electricity via solar power does not emit carbon, whether those solar panels are on the grid or on isolated houses or microgrids. Economies of scale in solar generation and for the distribution of power still suggest that connecting solar panels to the grid would achieve lower costs. Yet the government may have a rationale to subsidize off-grid solar, wholly aside from environmental and cost concerns. Tolerance for theft, as much as being able to serve higher loads, is a large part of the grid's appeal in this setting. Each customer does not consume much but loses a lot of money for the state. Every customer that better solar power takes from the money-losing grid increases the state's producer surplus by reducing losses. These changes are large, with solar's entry saving the government a sum of money about equal to what the households that take-up solar themselves pay.

A broader question that our static analysis cannot answer is what the cost of this dysfunctional electricity supply sector is in the longer run. There is some evidence that electrification has large external returns (Lipscomb, Mobarak and Barham, 2013). It is hard to imagine a large business, for example in manufacturing or services, opening in an area with eleven hours of electricity supply (Allcott, Collard-Wexler and O'Connell, 2016). The combination of these facts implies that, even if rationing electricity is a statically necessary policy to support electricity access, it may limit rural growth. Off-grid solar systems cannot replace the scale economies of a well-functioning grid.

References

Aklin, Michaël, Patrick Bayer, SP Harish, and Johannes Urpelainen. 2017. "Does basic energy access generate socioeconomic benefits? A field experiment with off-grid solar power in India." *Science Advances*, 3(5): e1602153.

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

Barron, Manuel, and Maximo Torero. 2017. "Household electrification and indoor air pollution." *Journal of Environmental Economics and Management.*

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, Steven, James Levinsohn, and Ariel Pakes. 2004. "Differentiated products demand systems from a combination of micro and macro data: The new car market." *Journal of Political Economy*, 112(1): 68–105.

Berry, Steven T. 1994. "Estimating discrete-choice models of product differentiation." *The RAND Journal of Economics*, 242–262.

Bhaskar, Utpal. 2018*a*. "All Indian households to have electricity connection this year." *Live Mint.*

Bhaskar, Utpal. 2018*b*. "Narendra Modi highlights success of rural electrification." *Live Mint.*

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh." *Econometrica*, 82(5): 1671–1748.

Burgess, Robin, Michael Greenstone, Nicholas Ryan, and Anant Sudarshan. 2019. "Electricity is not a Right." *Under preparation, Journal of Economic Perspectives.*

Burlig, Fiona, and Louis Preonas. 2016. "Out of the Darkness and Into the Light? Development Effects of Rural Electrification." Working Paper,(October).

Business Today. 2017. "CM Nitish Kumar orders free power connection for all in Bihar."

De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2013. "The demand for, and consequences of, formalization among informal firms in Sri Lanka." *American Economic Journal: Applied Economics*, 5(2): 122–150.

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

Duflo, Esther, Rema Hanna, and Stephen P Rya. 2012. "Incentives work: Getting teachers to come to school." *The American Economic Review*, 102(4): 1241–1278.

Dupas, Pascaline, and Jonathan Robinsona. 2013. "Why don't the poor save more? Evidence from health savings experiments." *The American Economic Review*, 103(4): 1138–1171.

Feldman, David, Robert Margolis, and Paul Denholm. 2016. "Exploring the potential competitiveness of utility-scale photovoltaics plus batteries with concentrating solar power, 2015–2030." National Renewable Energy Laboratory.

Government of India. 2015. "Prime Minister to Launch Deendayal Upadhyaya Gram Jyoti Yojana in Patna." Press Information Bureau.

Grimm, Michael, Luciane Lenz, Jorg Peters, and Maximiliane Sievert. 2016. "Demand for Off-grid Solar Electricity: Experimental Evidence from Rwanda." *IZA Discussion Paper Series*, , (10427).

Howell, David, Brian Cunningham, Tien Duong, and Peter Faguy. 2016. "Overview of the DOE VTO advanced battery R&D pogram." Department of Energy.

International Energy Agency. 2017. "From Poverty to Prosperity, a World Energy Outlook-2017 Special Report."

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane. 2011. "Spring cleaning: Rural water impacts, valuation, and property rights institutions." *The Quarterly Journal of Economics*, 126(1): 145–205.

Kumar, Manish. 2019. "Bihar Has Power Under Nitish Kumar": Amit Shah's Positive Twist To An Old Topic." *New Delhi Television.*

Lancaster, Kelvin. 1971. "A New Approach to Consumer Demand Theory."

Lee, Kenneth, Edward Miguel, and Catherine Wolfram. 2016. "Experimental Evidence on the Demand for and Costs of Rural electrification." National Bureau of Economic Research Working Paper 22292.

Lipscomb, Molly, Mushfiq A Mobarak, and Tania Barham. 2013. "Development effects of electrification: Evidence from the topographic placement of hydropower plants in Brazil." *American Economic Journal: Applied Economics*, 5(2): 200–231.

McFadden, D. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." *Frontiers in Econometrics*, 105–142.

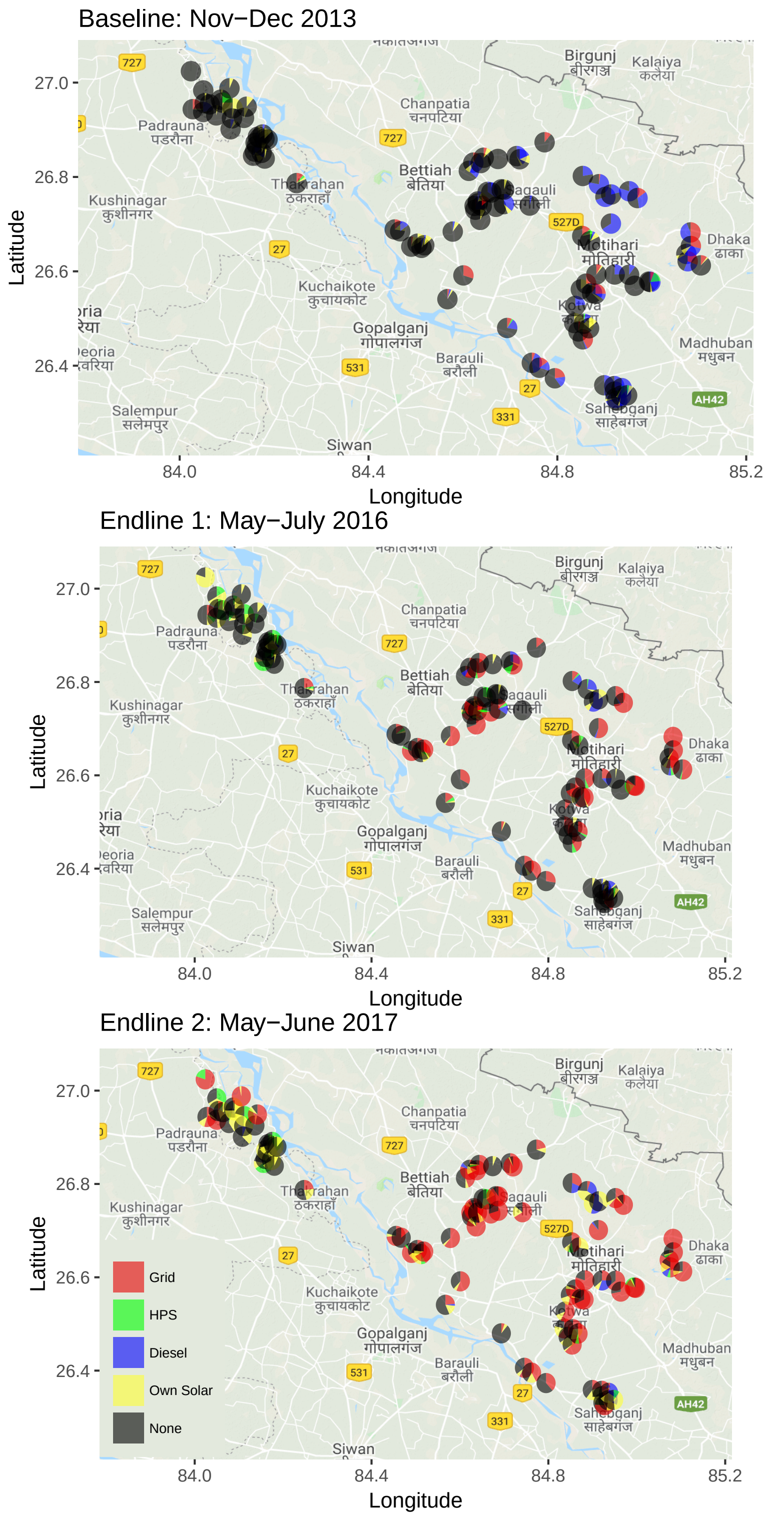
Nevo, Aviv. 2000. "A practitioner's guide to estimation of random-coefficients logit models of demand." *Journal of economics & management strategy*, 9(4): 513–548.

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

World Bank. 2017. "State of electricity access report 2017 (Vol. 2)." World Bank.

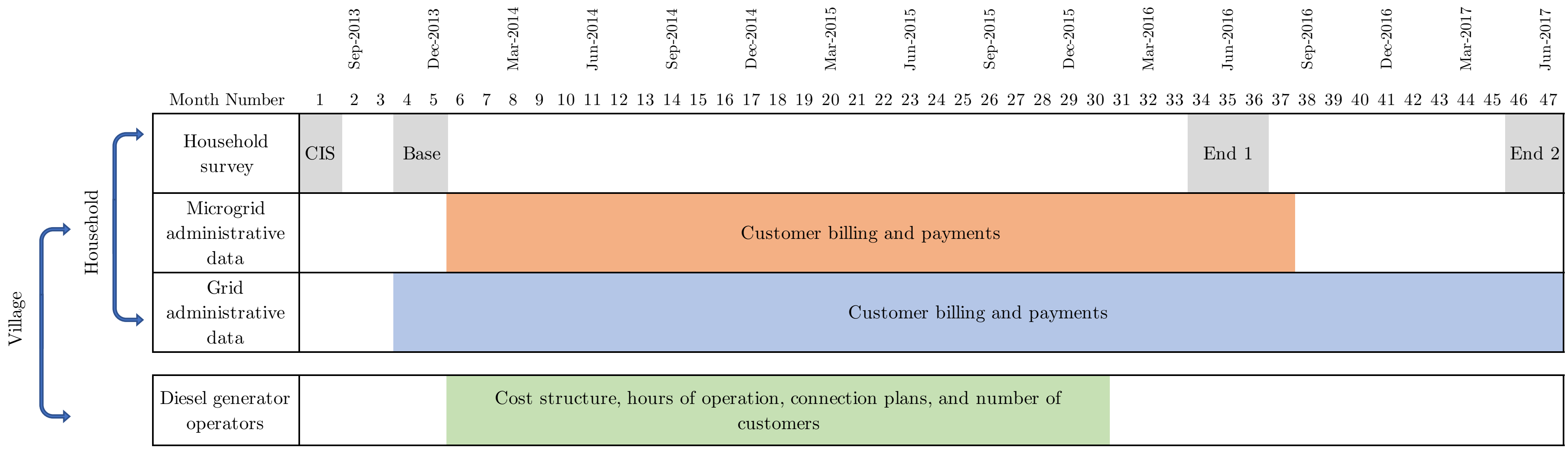
6 Figures

Figure 1: Sources of Electricity for Households in Rural Bihar



The figures show the highly dynamic composition of the electricity market in our sample villages across the three survey waves. There is considerable heterogeneity in the availability of grid, diesel, and solar across space and time. In fewer than four full years, the grid electrification rate in our sample rose from 5% to 41%. At the same time, the share of sample households with electricity from diesel generators fell from 17% to 3%, and the share with their own solar systems leapt from 5% to 21%.

Figure 2: Data Collection Timeline

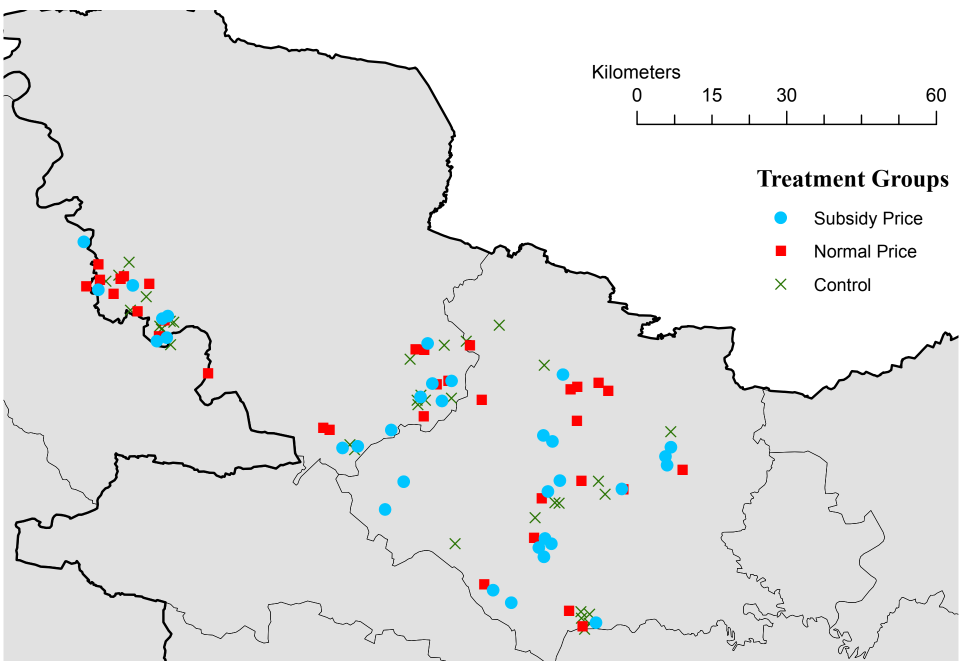


Household survey, microgrid administrative data, and grid administrative data are each at the household-level. We connect these 3 sources together by using unique household identification numbers. Finally, we enrich our dataset by joining in diesel operator data at the village-level. In Aug 2013, we conducted a customer identification survey ("CIS") for the villages in our study, which was subsequently used to assign villages to treatment and control groups and also served as the basis from which we drew our household sample. In each of the household baseline and endline 1 surveys, we collected data that covered demographics (e.g. literacy and number of adults per household), wealth proxies (e.g. income, size and structure of house and ownership of agricultural land), electrification status (e.g. all sources from which the household procured electricity and payments for each source), and quality of electricity sources (e.g. load and daily hours of supply). Moreover, we also recorded ownership of assets such as mobile phones, bulbs, fans and TVs, and tracked select measures of education and health, namely children's reading and math test scores and self-reported respiratory problems. The household endline 2 survey was a shorter survey in which we recorded data on demand for and use of different electricity sources. We also collected household consumer IDs in this survey, which facilitated household matching across datasets and allowed us to distinguish formal (i.e. paying) consumers from informal (i.e. non-paying) consumers.

Figure 3: Maps of Study Area

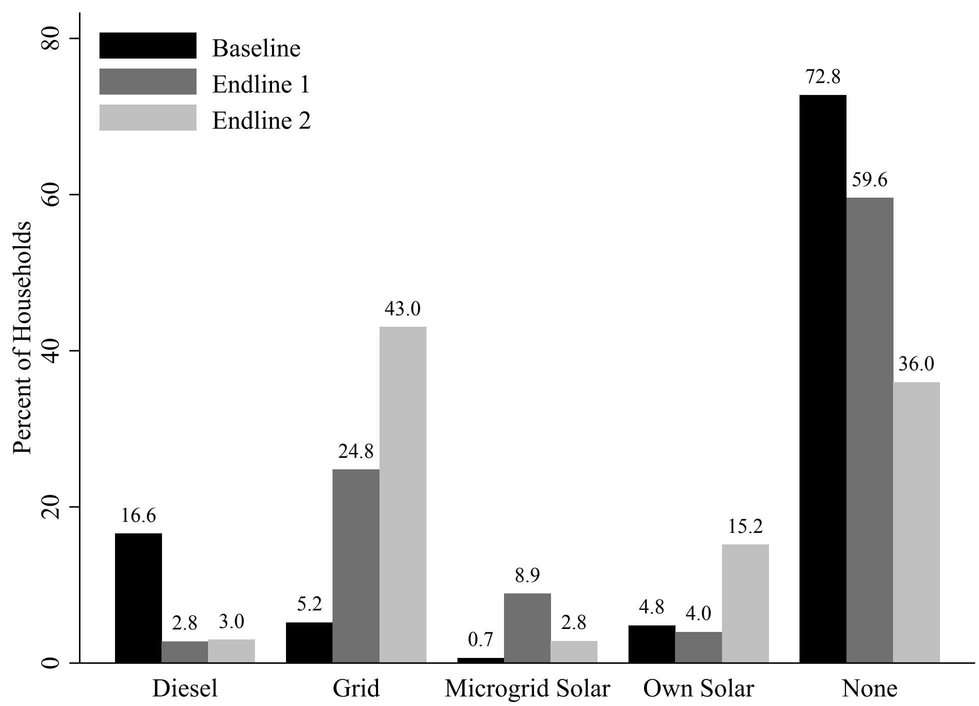
(a) Study districts within the state of Bihar, India

(b) Sample villages within study districts



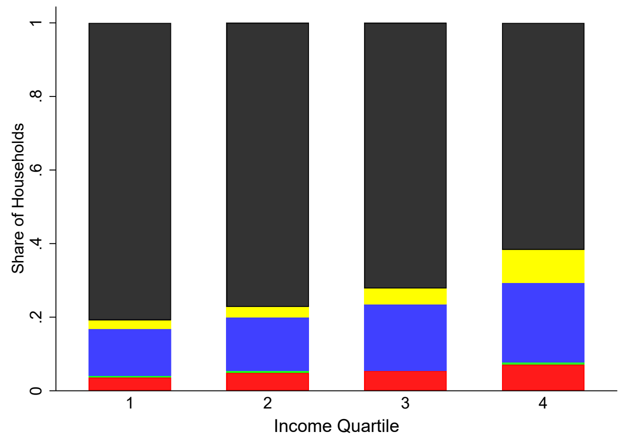
The figure shows the location of the study area. Panel A highlights the two districts of West Champaran and East Champaran in northwestern Bihar state where the study villages are located. Panel B shows, within the two study districts, the locations of sample villages and their treatment assignments. The nearest large towns are Bettiah and Motihari. The river Gandak in the northwest forms the state border with Uttar Pradesh.

Figure 4: Take-Up of Electricity Sources by Survey Wave

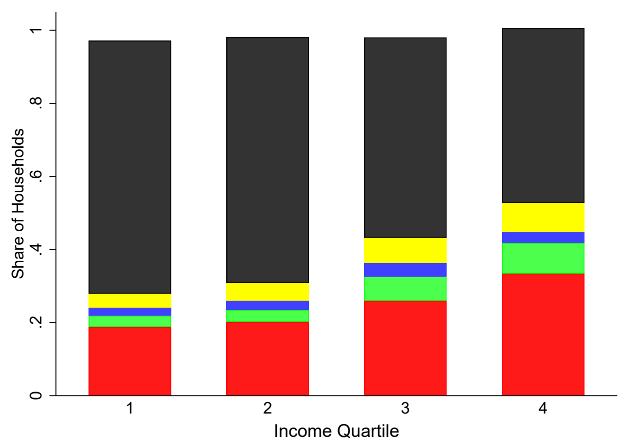


The figure shows the take-up of electricity sources across the three different waves of our household survey. Each group of bars shows the market shares of a given electricity source across sample households. Diesel is diesel generators, grid is the state-run electricity grid, microgrid solar is the HPS solar microgrid offered in the experiment, and own solar is individual household-level solar systems. Within each group of bars the market share is shown for each of three survey waves. The survey waves are: baseline (starting November, 2013), endline 1 (starting May, 2016) and endline 2 (starting May, 2017).

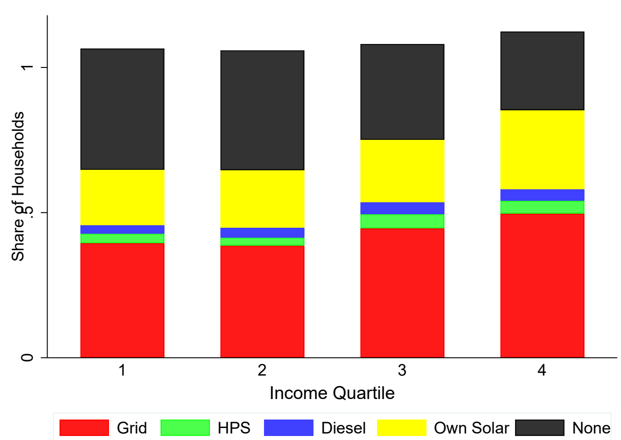
Figure 5: Market Share of Electricity by Household Income Profiles



(a) Baseline Survey

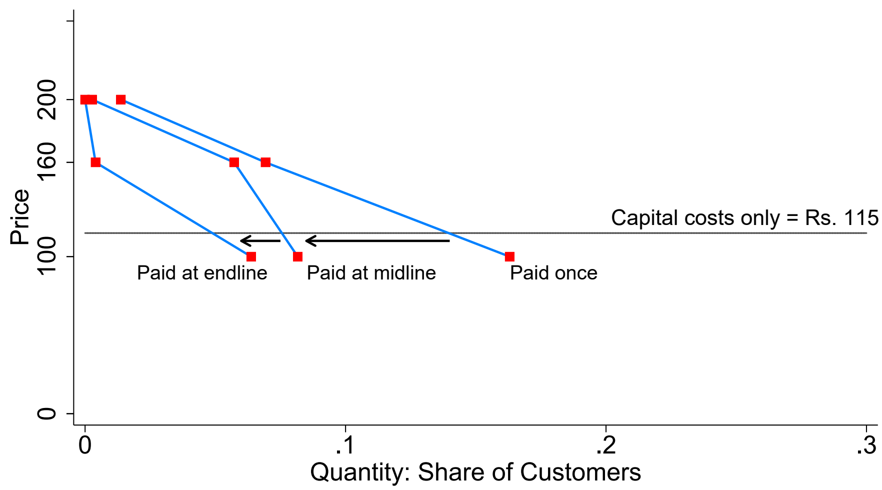


(b) Endline 1 Survey



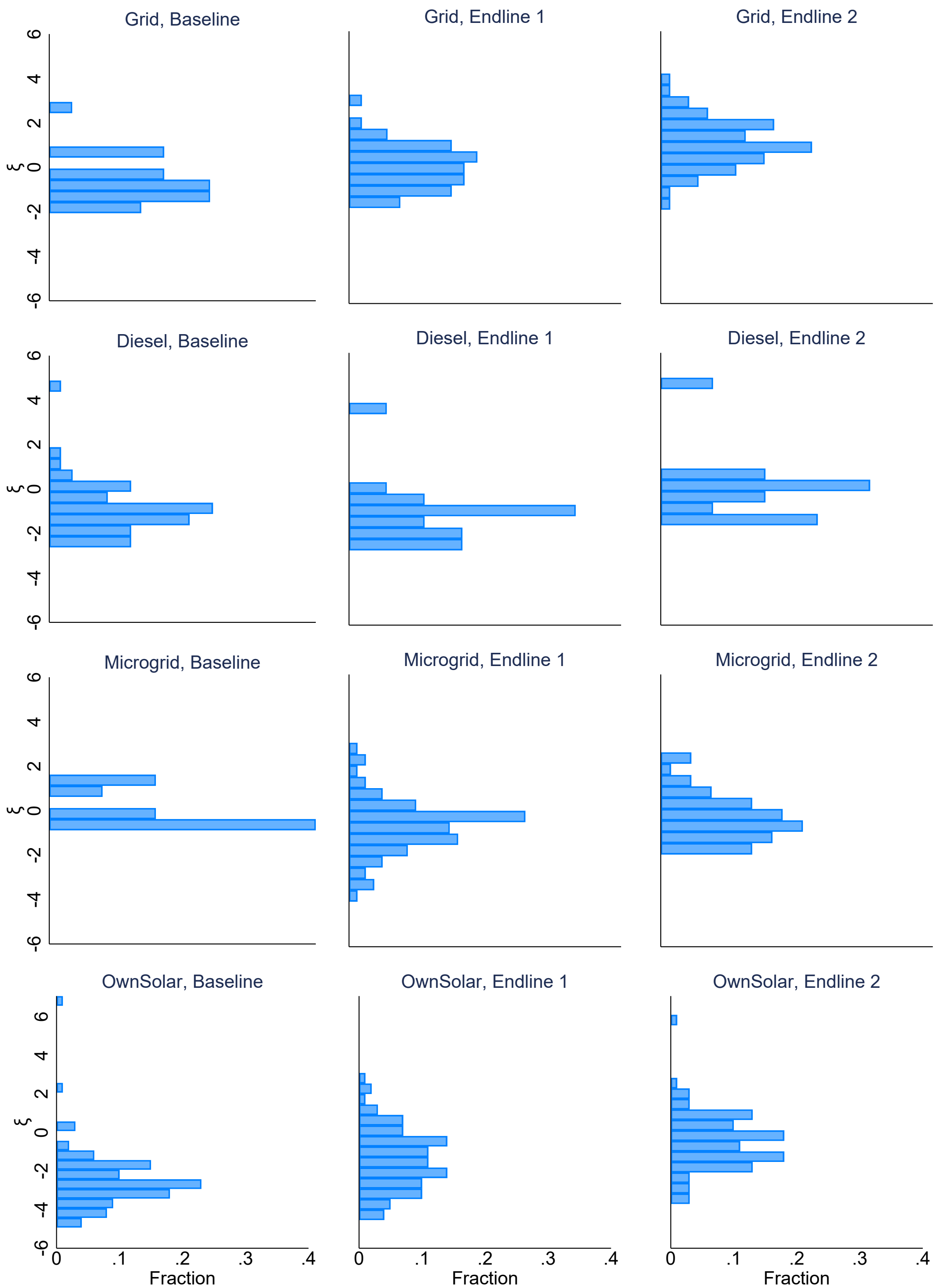
(c) Endline 2 Survey

Figure 6: Demand Curve for Microgrid Solar



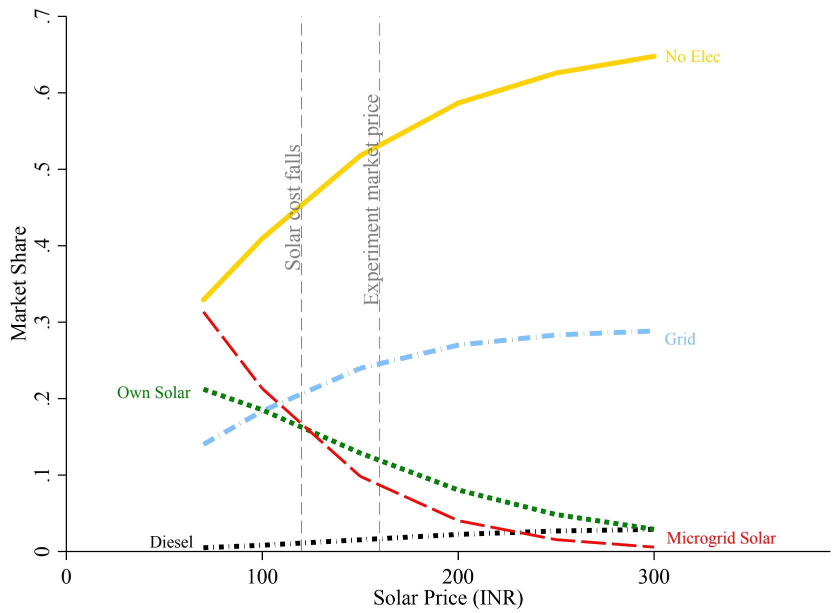
The figure plots the share of sample households that paid for the Husk Power Systems (HPS) solar microgrid at three different times. The horizontal axis is the share of households paying and the vertical axis is the monthly price. Each line on the figure represents the demand for solar microgrids at a different point in time. The outer demand curve is for households who paid at any time during the experiment; the middle curve for households that paid at the midpoint of the experiment (months 16-18); the inner curve is for households that paid during the endline one survey (starting in May, 2016, month 29). At each point in time household demand is shown for three different prices. The horizontal line on the figure is an estimate of the amortized monthly capital costs of the microgrid system per household. This cost does not include costs such as marketing, billing or operations and maintenance.

Figure 7: Distribution of Unobserved Mean Indirect Utility (ξ_{vtj}) by Source and Wave



The figure plots the quality of different electricity sources over time. The four rows are for different electricity sources, from top to bottom: grid electricity, diesel, HPS solar microgrids, and household's own solar systems. The four columns are for different survey waves, from left to right: baseline (starting November, 2013), endline one (starting May, 2016) and endline two (starting May, 2017). Each tiled panel in the figure shows the distribution across villages of source-specific unobserved mean quality ξ_{vtj} for the row source during the column survey wave. The vertical axis is the value of mean unobserved quality, where the outside option is normalized to zero, and the horizontal axis is the density of the historgram. The mean unobserved quality is estimated in the demand model as the residual that fits source market shares given the observed characteristics of each source. The unobserved quality is therefore only recovered, and plotted, for source-village-wave combinations in which a given source was offered (i.e., it is not possible to infer grid quality when the grid is not present in a village).

Figure 8: Market Shares under Varying Solar Prices

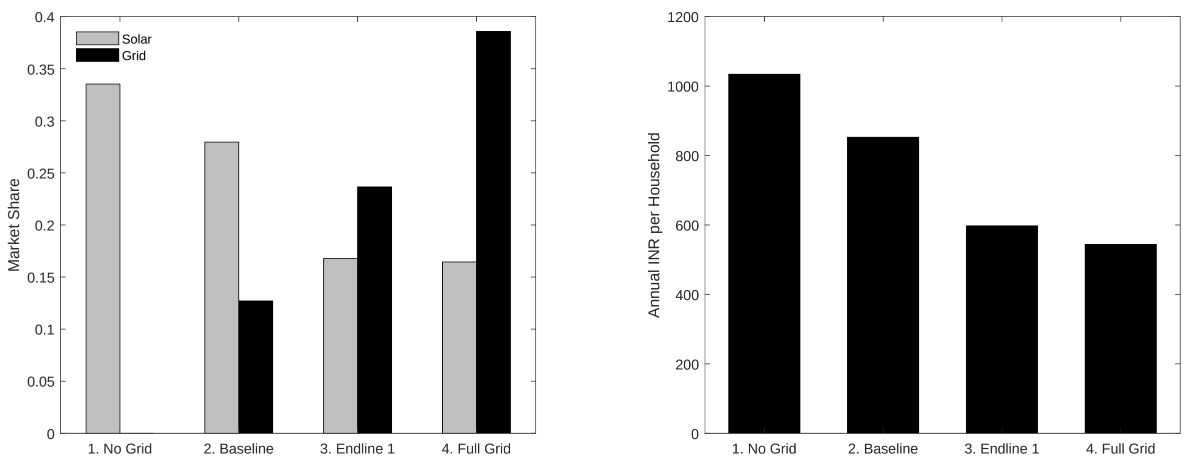


The figure shows demand for all electricity source technologies as the price for solar power varies. Each curve is the predicted market share of an electricity source technology. The horizontal axis gives the price of an HPS solar microgrid. While the horizontal axis shows the price of an HPS solar microgrid only, we vary the price of own solar systems proportionately with the microgrid, on the grounds that capital cost reductions in solar photovoltaic panels or batteries would affect the price of both productions in proportion to their capital share. The microgrid price shown in the figure ranges from INR 70 up to the choke price of INR 300. Household and source characteristics and the availability of all sources are fixed at their endline one (mid-2016) levels.

Figure 9: Solar Market Share and Surplus by Grid Penetration

(a) Solar and grid market shares

(b) willingness to pay for off-grid solar



The figure shows solar market share (own solar and microgrid solar combined) and the household willingness to pay for solar depending on the extent of the grid. The four grid extents are: no grid (0%), grid as of our baseline survey in 2013 (43% of villages with any connection), grid as of our endline one survey in mid-2016 (57%), and complete grid penetration (100%). Panel A shows the market shares for solar and the grid for each extent, holding technology and household characteristics constant at endline one levels. Panel B shows the mean household willingness to pay (in INR '000s) for a choice set that includes solar power relative to a choice set without off-grid solar.

Figure 10: Counterfactual - Value of Solar

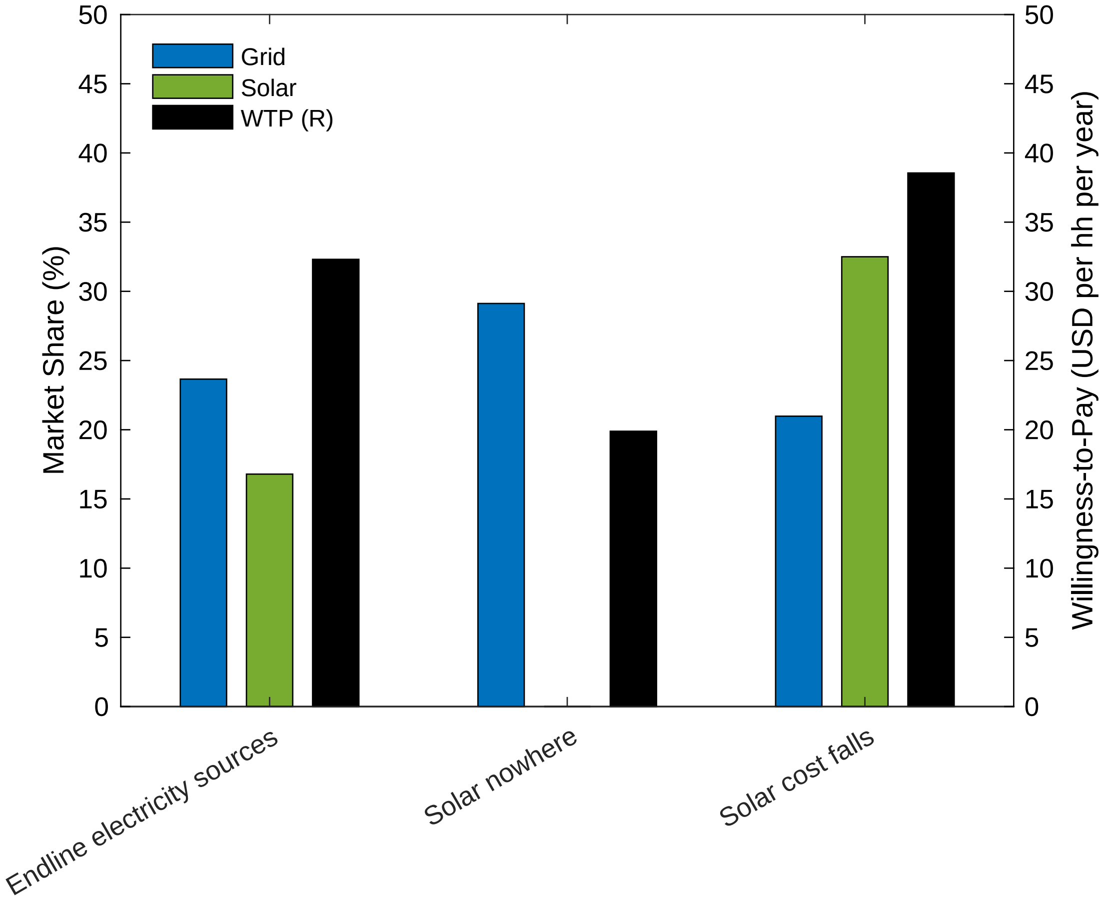


Figure 11: Counterfactual - Value of Grid

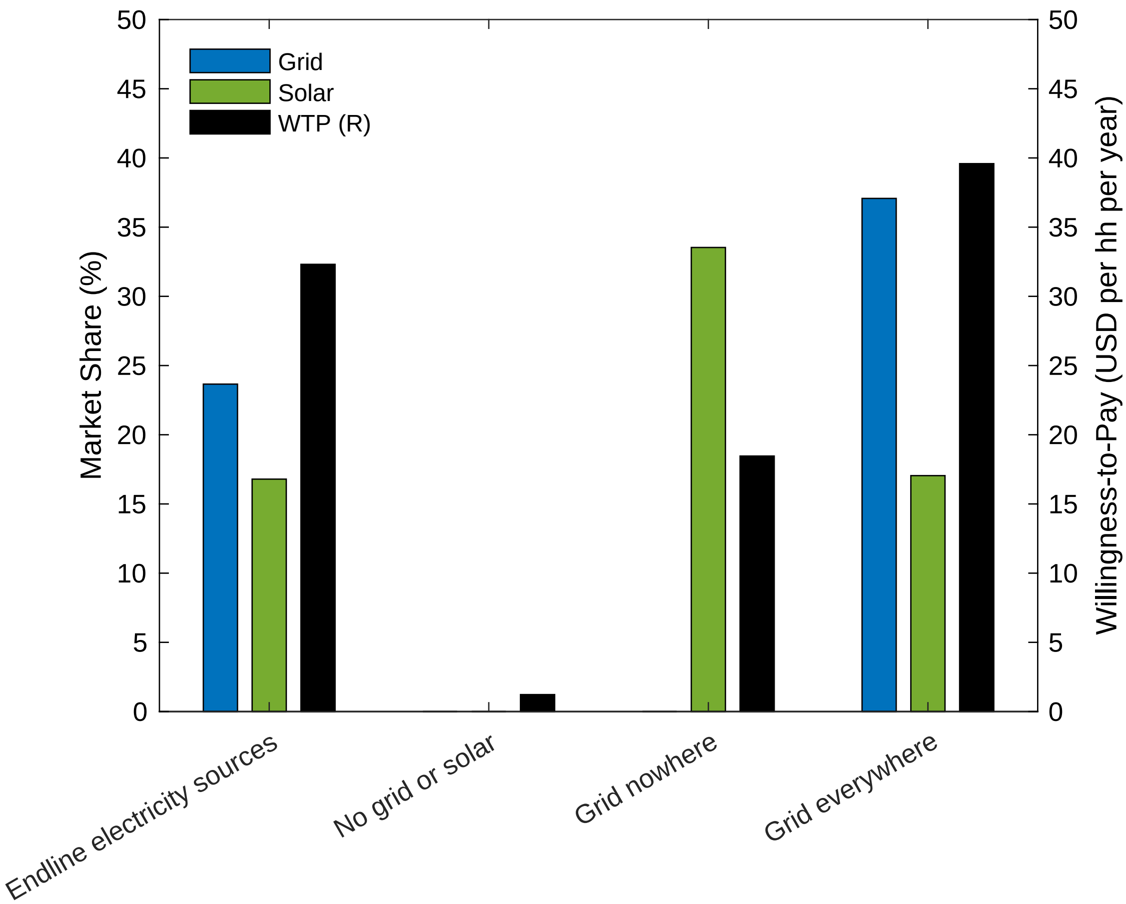
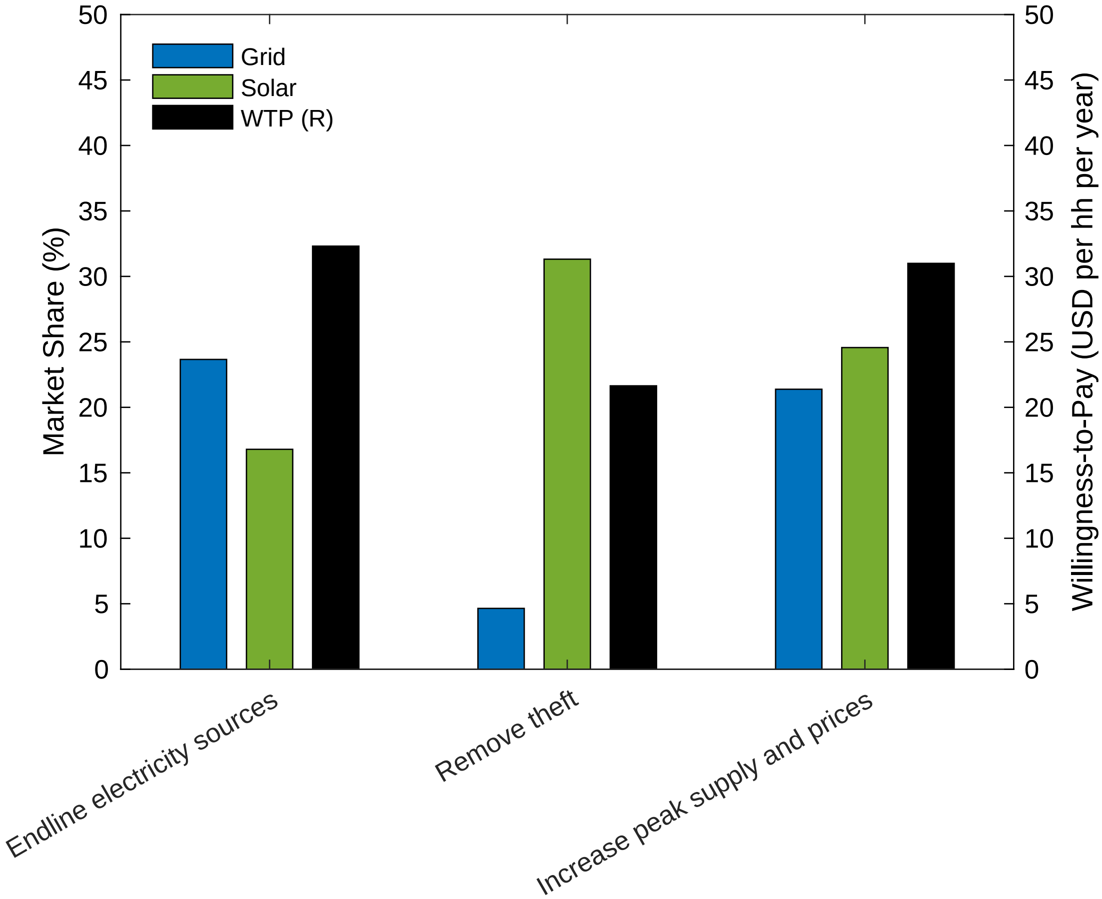


Figure 12: Counterfactual - Reform grid policy



7 Tables

Table 1: Electrification Context Around the World

	United States (1)	India (2)	Sub-Saharan Africa (3)	Bihar (4)
GDP per capita (USD)	57,467	1,709	1,449	420
kWh per capita	12,985	765	481	122
Electricity access (% of population)	100	79	37	25
kWh per capita / US kWh per capita	1	0.059	0.037	0.009

The table places the income and electricity access in the state of Bihar, India, the site of the study (column 4), in the context of other areas of the world (columns 1 through 3). The first row is nominal GDP per capita, the second row is mean electricity consumption per capita, the third row is the electrification rate and the last row is the ratio of mean electricity consumption per capita to mean consumption in the United States. The source of data is (World Bank, 2017).

Table 2: Summary of Electricity Sources, All Panels

	Grid (1)	Diesel (2)	Own Solar (3)	Microgrid (4)
Panel A. Baseline				
Monthly price (Rs.)	73.26	98.88	74.43	200
Load (watts)	322	133	151	28
Supply hours	9.74	3.38	7.92	3.14
Source available (percent)	35.3	62.1	100	.65
Panel B. Endline 1				
Monthly price (Rs.)	59.34	87.50	96.06	164.07
Supply hours	10.01	3.08	5.71	6.16
Source available	58.9	10.1	100	66.4
Ownership of assets				
Mobile and/or bulb	1	1	1	0.93
Fan	0.34	0.04	0.10	0.03
TV	0.11	0.01	0.04	0.02
Monthly Income (Rs.)	9222	8547	8760	8493
Panel C. Endline 2				
Monthly price (Rs.)	58.71	88.90	66.08	170
Supply hours	12	3.08	5.7	6.16
Source available	76.3	15.4	100	65.7

The table shows characteristics of the various sources of electricity that constitute the rural electricity market we study. Load reported here is based on household survey appliance ownership, and household survey reports of own solar watt ratings. In the model, we apply diesel generator survey data to assign load available to households served by each generator, as well as technical specifications from HPS for panel capacity.

Table 3: Baseline Covariate Balance

	Control (1)	Normal (2)	Subsidy (3)	Diff(N-C) (4)	Diff(S-C) (5)	FTest (6)
Panel A. Demographics						
Literacy of household head (1-8)	2.44	2.69	2.60	0.25	0.16	1.33
	[2.04]	[2.15]	[2.10]	(0.16)	(0.15)	(0.27)
	1031	971	989	2002	2020	
Number of adults	3.31	3.50	3.49	0.20*	0.18*	2.19
	[1.58]	[1.75]	[1.78]	(0.11)	(0.11)	(0.12)
	1052	983	1001	2035	2053	
Panel B. Wealth Proxies						
Income (Rs. '000s/month)	7.46	7.32	7.28	-0.14	-0.18	0.068
	[6.91]	[6.93]	[7.09]	(0.57)	(0.51)	(0.93)
	1041	963	983	2004	2024	
Number of rooms	2.40	2.55	2.53	0.15	0.13	1.29
	[1.32]	[1.45]	[1.45]	(0.10)	(0.098)	(0.28)
	1052	981	999	2033	2051	
House type (pukka = 1)	0.24	0.27	0.31	0.035	0.074**	2.79*
	[0.43]	[0.45]	[0.46]	(0.037)	(0.031)	(0.066)
	1052	983	1001	2035	2053	
Owns agricultural land	0.67	0.69	0.67	0.015	0.0022	0.045
	[0.47]	[0.46]	[0.47]	(0.056)	(0.053)	(0.96)
	1052	983	1001	2035	2053	
Solid Roof (=1)	0.42	0.46	0.51	0.042	0.095**	3.08*
	[0.49]	[0.50]	[0.50]	(0.043)	(0.039)	(0.050)
	1052	983	1001	2035	2053	
Panel C. Energy Access						
Any elec source (=1)	0.25	0.31	0.27	0.061	0.022	0.63
	[0.43]	[0.46]	[0.44]	(0.055)	(0.050)	(0.54)
	1052	983	1001	2035	2053	
Uses gov. elec (=1)	0.030	0.036	0.091	0.0052	0.060**	2.53*
	[0.17]	[0.19]	[0.29]	(0.017)	(0.028)	(0.085)
	1052	983	1001	2035	2053	
Uses diesel elec (=1)	0.17	0.21	0.11	0.039	-0.063	1.70
	[0.38]	[0.41]	[0.31]	(0.058)	(0.046)	(0.19)
	1052	983	1001	2035	2053	
Uses own panel (=1)	0.034	0.050	0.061	0.016	0.027*	1.81
	[0.18]	[0.22]	[0.24]	(0.014)	(0.015)	(0.17)
	1052	983	1001	2035	2053	
Uses microgrid solar (=1)	0.0067	0.0081	0.0050	0.0015	-0.0017	0.14
	[0.081]	[0.090]	[0.071]	(0.0078)	(0.0054)	(0.87)
	1052	983	1001	2035	2053	

The table reports the balance of covariates in our baseline survey across treatment arms for demographic variables (Panel A), wealth or demand proxy variables (Panel B) and energy access (Panel C). The first three columns show the mean values of each variable in the control, normal price and subsidized price arms, with standard deviations in brackets. The next two columns show the differences between the normal price and control arms and subsidized price and control arms, respectively, with the standard error of the difference. The final column shows the F-stat and p-value from a test of the null that the treatment dummies are jointly zero at baseline. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0:01$

Table 4: Household Electricity Use Outcomes

	Light Bulb Ownership (=1)	Daily Hours of Electricity Use	Mobile Phone Ownership	Price of Full Charge (Rs.)
	(1)	(2)	(3)	(4)
Subsidy treat village (=1)	0.15***	0.94***	0.034**	-0.67***
	(0.047)	(0.24)	(0.014)	(0.24)
No subsidy treat village (=1)	0.098**	0.52**	0.022	-0.46*
	(0.044)	(0.20)	(0.013)	(0.23)
Baseline Controls	Yes	Yes	Yes	Yes
Control mean	0.32	1.16	0.88	4.72
Observations	3001	2868	3001	964

The table shows regressions of ownership of LED lamps and one mobile charging point and the use of electricity on treatment status. Households in the treatments got and used electricity microgrids and these appliances. The specifications include baseline electricity source indicators, baseline monthly income, and baseline equivalent of outcome variable as controls. Standard errors clustered at the village level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Household Income, Education and Health Outcomes

	Monthly income	Standardized test score		Respiratory problems (=1)	
	(INR '000s)	Reading	Math	Adults	Children
	(1)	(2)	(3)	(4)	(5)
	Panel A. Reduced Form				
Subsidy treat village (=1)	0.18	0.11*	0.095	0.026	0.012
	(0.31)	(0.061)	(0.065)	(0.021)	(0.0082)
Normal treat village (=1)	0.63*	0.020	0.071	0.017	0.0041
	(0.33)	(0.061)	(0.062)	(0.018)	(0.0082)
	Panel B. Instrumental Variables				
Hours of electricity	0.15	0.22	0.21	0.027	0.014
	(0.35)	(0.24)	(0.23)	(0.027)	(0.012)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Control mean	7.5	0	0	0.14	0.024
Observations	2692	646	637	2710	2669

The table shows the effects of provision of solar microgrids on social and economic outcomes, for health, education and test scores. Panel A of the table is the reduced-form or intent-to-treat effect of solar microgrids for these outcomes, and Panel B is the instrumental variable estimate of the coefficient on hours of electricity using the two treatment assignment dummies as instruments. We find no evidence that respiratory problems decrease for adults or children (Panel B, columns 4 and 5). The predominant source of indoor air-pollution comes from cooking, which is unaffected by the provision of microgrids, and we do not find significant declines in kerosene expenditure (not reported). Effects on reading test scores are positive but imprecisely estimated (columns 2 and 3). For example, we estimate that an hour of additional electricity use increase children's reading scores by 0.22 standard deviations (standard error 0.22 standard deviations). This is a fairly large standardized effect but imprecise due to low first-stage take-up and the children tested being only a subsample of the overall experiment. We cannot rule out a zero effect or a significant positive effect of lighting on child test scores. Finally we find that electricity has a null effect on household income of INR 150 per month (standard error 350), which is small compared to baseline income of INR 7,500 per month. Test score results are at the child level. The regressions include baseline electricity source indicators, baseline monthly income, and baseline equivalent of outcome variable controls. Standard errors clustered at the village level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Impact of Household Characteristics on Choice Probabilities

	Grid	Diesel	Own Solar	Microgrid	None
Number of adults	0.036	0.002	0.001	0.005	-0.045
	(0.009)	(0.006)	(0.001)	(0.004)	(0.008)
Household income	0.016	0.002	0.001	0.014	-0.034
	(0.007)	(0.005)	(0.001)	(0.004)	(0.008)
Household owns land	0.049	-0.023	0.003	0.008	-0.037
	(0.018)	(0.010)	(0.003)	(0.009)	(0.016)
Household head literacy	0.026	0.008	-0.001	0.002	-0.036
	(0.008)	(0.005)	(0.001)	(0.004)	(0.007)
Pukka (solid) house	0.077	-0.004	-0.002	-0.006	-0.065
	(0.023)	(0.013)	(0.003)	(0.008)	(0.019)
Solid roof	0.107	-0.007	0.005	-0.007	-0.098
	(0.025)	(0.013)	(0.003)	(0.007)	(0.018)
Number of rooms	0.026	0.011	0.003	0.003	-0.043
	(0.008)	(0.005)	(0.001)	(0.004)	(0.008)

The table shows the discrete effects of changes in household observable characteristics (in rows) on the probability the household will purchase different electricity sources (in columns). The household characteristics are from our survey. The changes in choice probabilities are calculated with the demand model, for which the estimated coefficients are presented in Appendix Table A9. Each cell entry is the change in choice probability for a poor household from increasing the row characteristics. For discrete household characteristics, the increase is from zero to one. For continuous household characteristics, the increase is of one standard deviation. Appendix Table A3 describes the statistical profile of a poor household and Appendix Table A1 shows the magnitude of changes in household characteristics for each variable. Standard errors are constructed using the delta method.

Table 7: Two-Stage Least Squares Estimates for Demand for Electricity

	OLS (1)	Price IV (2)	Price & Hours IV (3)	Price First Stage (4)
Price (Rs. 100)	-0.19*	-2.08***	-2.07***	
	(0.11)	(0.74)	(0.74)	
Hours of supply on peak	0.20	0.10	0.19	
	(0.21)	(0.21)	(0.27)	
Hours of supply off peak	-0.092*	-0.077*	-0.11*	
	(0.047)	(0.047)	(0.060)	
Normal Price				0.045
				(0.035)
Subsidy Price				-0.14***
				(0.031)
Peak Hours Instrument				-0.032
				(0.044)
Peak Hours Instrument				0.0037
				(0.0091)
ξ_{tj} mean effects	Yes	Yes	Yes	Yes
Observations	1000	1000	1000	1000
fstat				

The table presents 2SLS estimates of our demand system. The dependent variable is mean indirect utility at the market $\times$ survey wave level retrived from the non-linear first stage. Peak hours refers to supply of electrricity during the evening (5pm-10pm). The second column presents two-stage least squares estimates where instrument price with the experimentally varied HPS treatment assignment. In the third column we instrument for price and peak, off-peak hours. For grid hours, we use predictions from a random forest model (see appendix for details) as instruments. For off-grid sources, we use the data out instrument matrix. The last column provides first stage regrressions for the specification in column 2. First stage regressions for the specification in column 3 are presented in the appendix Table A9. All regressions control for wave $\times$ source mean effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors cluster at the village level in parentheses.

Table 8: The Value of Electrification under Counterfactual Policies

	Market shares					Surplus (INR per household per year)				
						Consumer				
	Grid (1)	Diesel (2)	Own solar (3)	Microgrid (4)	None (5)	BPL (6)	APL (7)	All (8)	Producer (9)	Total (10)
Panel A: Fit of model										
Data	24	3	7	10	57	-	-	-	-	-
Model	24	3	7	10	57	1849	2208	1939	-497	1442
Panel B: Value of solar innovation										
Solar nowhere	29	3	0	0	68	1087	1513	1193	-612	582
Solar everywhere	23	2	13	10	51	1801	2277	1920	-492	1428
Solar cost falls	21	1	17	16	45	2199	2654	2313	-441	1872
Further solar innovation	17	1	36	8	39	3737	4185	3849	-349	3500
Panel C: Grid Extension										
No grid or solar	0	5	0	0	95	58	120	74	0	74
Grid nowhere	0	3	20	13	63	1039	1312	1108	0	1108
Grid everywhere	39	1	8	9	44	2342	2710	2434	-803	1631
Extra 2 Hours	29	2	12	10	48	2037	2568	2170	-736	1434
Grid everywhere and solar cost falls	34	1	11	14	41	2649	3024	2743	-708	2035
Grid everywhere, increase peak hours, and solar cost falls	45	0	7	12	36	3020	3412	3118	-1115	2004
Panel D: Theft Reductions										
Remove theft by raising grid price	5	3	19	13	61	1217	1543	1299	-52	1247
Increase peak supply and raise prices	21	2	14	11	52	1749	2192	1860	-497	1363

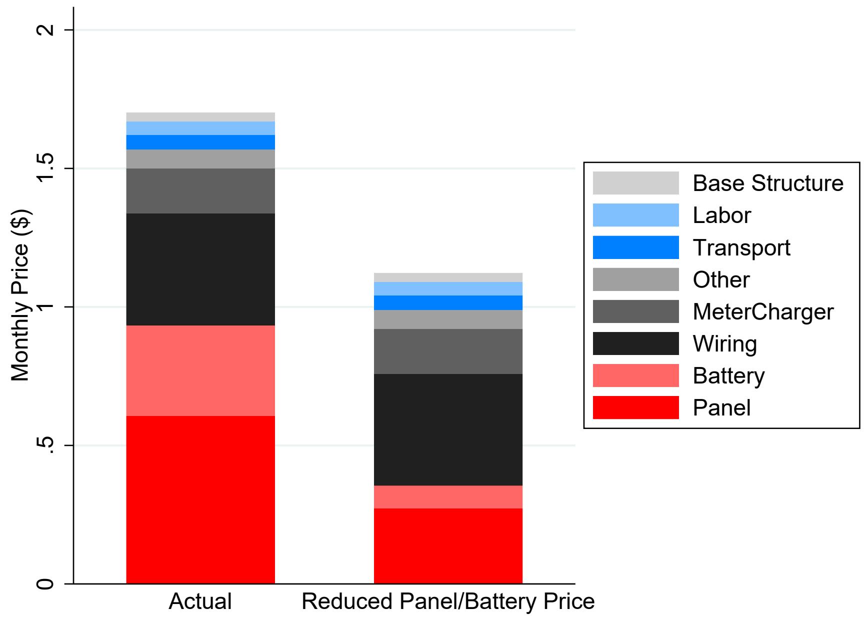
The table presents market shares and surplus under counterfactual changes in the supply side of the electricity market. All counterfactuals are calculated using our demand model estimates. The counterfactual scenarios are laid out in Section 4 of the text and the detailed assumptions behind the counterfactuals are in Appendix Table A11. In Panel A, we compare the market shares in the data at the time of the endline one survey (mid-2016) to market shares in the model. Panel A, row 2 is the baseline scenario for surplus in the status quo. In Panel B we vary the

A Appendix

This section gives additional results and tests the robustness of the results in the text. Subsection A.1 gives results additional to the main text, subsection A.2 shows robustness checks for the demand model, and subsection A.3 describes the construment of the instrument for hours of supply.

A.1 Additional results

Figure A1: Microgrid Solar Price under Current and Counterfactual Capital Costs



In our counterfactuals, we consider reductions in cost for solar photovoltaics and for batteries. We assume a 55% reduction in cost of solar PV in line with the National Renewable Energy Laboratory's for 2022. For batteries, we assume a cost reduction of 75% in accordance with the US Department of Energy's 2022 goal.

Table A1: Definition of Household Characteristics and Magnitude of Marginal Change

Characteristic	Definition	Marginal Change
Income	Monthly income	1 SD (INR 6486)
Land	Indicator for agricultural land	0 to 1
Roof	Indicator for solid roof	0 to 1
Pukka	Indicator for pukka house	0 to 1
Rooms	Number of rooms in the house	1 SD (1.32 rooms)
Adults	Adults in the HH	1 SD (1.82 persons)
Literacy	Literacy of household head (1-8)	1 SD (2.04 years)

The table shows the magnitude of the change in household covarates for which the marginal impact on household choice probabilities is estimated in Table A4 and Table A5. Literacy classification: 1 =not literate, 2= Aanganwadi, 3 = literate but below primary, 4 = literate till primary, 5 = literate till middle, 6 = literate till secondary, 7= literate till higher secondary, 8 = graduate and above

Table A2: Summary Statistics of Household Characteristics

	Mean	Median	Q1	Q3	SD	Min	Max
Number of rooms in the house	2.45	2	2	3	1.32	1	11
Indicator for pukka house	0.32	0	0	1	0.47	0	1
Indicator for agricultural land	0.63	1	0	1	0.48	0	1
Indicator for solid roof	0.51	1	0	1	0.50	0	1
Literacy of household head (1-8)	2.48	1	1	4	2.04	1	8
Adults in the HH	3.67	3	2	5	1.83	1	15
Monthly income (INR)	7575.5	6000	4000	8500	6486.2	0	65000.0
Observations	8822						

Table A3: Household Profile for Marginal Effects

Profile	Rooms	Pukka	Land	Roof	Literacy	Adults	Income (INR)
Poor	1	0	0	0	1	2	3750
Median	2	0	1	1	1	3	6000
Rich	3	1	1	1	5	5	9500

Table A4: Impact of Household Characteristics on Choice Probabilities (Median Household)

	Grid	Diesel	Own Solar	Microgrid	None
Number of adults	0.047	-0.001	0.001	0.003	-0.049
	(0.010)	(0.003)	(0.002)	(0.004)	(0.008)
Household income	0.019	0.000	0.001	0.014	-0.034
	(0.010)	(0.003)	(0.002)	(0.005)	(0.008)
Household owns land	-	-	-	-	-
Household head literacy	0.036	0.004	-0.002	0.001	-0.038
	(0.010)	(0.003)	(0.002)	(0.004)	(0.008)
Pukka (solid) house	0.098	-0.006	-0.006	-0.011	-0.075
	(0.025)	(0.008)	(0.005)	(0.009)	(0.020)
Solid roof	-	-	-	-	-
Number of rooms	0.035	0.005	0.004	0.001	-0.046
	(0.010)	(0.004)	(0.002)	(0.005)	(0.009)

The table shows the discrete effects of changes in household observable characteristics (in rows) on the probability the household will purchase different electricity sources (in columns). The household characteristics are from our survey. The changes in choice probabilities are calculated with the demand model, for which the estimated coefficients are presented in Appendix Table A9. Each cell entry is the change in choice probability for a poor household from increasing the row characteristics. For discrete household characteristics, the increase is from zero to one. For continuous household characteristics, the increase is of one standard deviation. Appendix Table A3 describes the statistical profile of a poor household and Appendix Table A1 shows the magnitude of changes in household characteristics for each variable. Standard errors are constructed using the delta method.

Table A5: Impact of Household Characteristics on Choice Probabilities (Rich Household)

	Grid	Diesel	Own Solar	Microgrid	None
Number of adults	0.036	-0.003	-0.000	0.000	-0.033
	(0.007)	(0.003)	(0.001)	(0.004)	(0.005)
Household income	0.011	-0.001	0.000	0.011	-0.022
	(0.008)	(0.003)	(0.001)	(0.005)	(0.006)
Household owns land	-	-	-	-	-
Household head literacy	0.028	0.002	-0.002	-0.001	-0.026
	(0.008)	(0.003)	(0.001)	(0.003)	(0.004)
Pukka (solid) house	-	-	-	-	-
Solid roof	-	-	-	-	-
Number of rooms	0.026	0.003	0.002	-0.001	-0.030
	(0.009)	(0.004)	(0.002)	(0.004)	(0.006)

The table shows the discrete effects of changes in household observable characteristics (in rows) on the probability the household will purchase different electricity sources (in columns). The household characteristics are from our survey. The changes in choice probabilities are calculated with the demand model, for which the estimated coefficients are presented in Appendix Table A9. Each cell entry is the change in choice probability for a poor household from increasing the row characteristics. For discrete household characteristics, the increase is from zero to one. For continuous household characteristics, the increase is of one standard deviation. Appendix Table A3 describes the statistical profile of a poor household and Appendix Table A1 shows the magnitude of changes in household characteristics for each variable. Standard errors are constructed using the delta method.

Table A6: Summary Statistics by BPL Status

	BPL	APL	BPL-APL
Rooms	2.38	2.58	-0.20***
	[1.18]	[1.50]	(0.054)
Pukka house	0.33	0.42	-0.088***
	[0.47]	[0.49]	(0.020)
Agricultural land	0.58	0.66	-0.088***
	[0.49]	[0.47]	(0.021)
Solid roof	0.51	0.58	-0.069***
	[0.50]	[0.49]	(0.021)
Literacy	2.25	3.01	-0.76***
	[1.87]	[2.31]	(0.085)
Income	0.73	0.87	-0.14***
	[0.58]	[0.74]	(0.027)
Number of adults in household	3.76	3.88	-0.12
	[1.83]	[2.02]	(0.080)
Observations	2186	731	

$*p < 0.10$, $**p < 0.05$, $***p < 0.01$

A.2 Robustness of demand model estimates

Table A7: First stage results choice-specific household characteristics and nest-similarity parameter

β_{rj}	Multinomial Logit	(Grid, diesel, own solar) & (HPS)	(Grid, diesel, HPS) & (Own solar)	(Grid, Own solar, HPS) & (Diesel)	(Grid) & (off-Grid)	(Grid, Own solar) & (diesel, HPS)	(Grid/Diesel) & (Solar)
Grid × Income	0.21	0.19	0.21	0.21	0.19	0.21	0.20
	(0.06)	(0.05)	(0.06)	(0.06)		(0.07)	(0.06)
Diesel × Income	0.14	0.16	0.14	0.14	0.21	0.14	0.14
	(0.08)	(0.06)	(0.09)	(0.08)		(0.14)	(0.08)
Own solar × Income	0.17	0.18	0.17	0.17	0.21	0.17	0.19
	(0.07)	(0.06)	(0.07)	(0.07)		(0.14)	(0.08)
HPS × Income	0.49	0.48	0.49	0.48	0.22	0.49	0.43
	(0.12)	(0.11)	(0.14)	(0.13)		(0.13)	(0.14)
Grid × Land	0.24	0.20	0.24	0.24	0.24	0.24	0.24
	(0.09)	(0.08)	(0.09)	(0.09)		(0.09)	(0.09)
Diesel × Land	-0.15	-0.09	-0.15	-0.15	0.04	-0.15	-0.15
	(0.12)	(0.11)	(0.12)	(0.12)		(0.15)	(0.12)
Own solar × Land	0.15	0.18	0.15	0.15	0.07	0.15	0.15
	(0.11)	(0.09)	(0.11)	(0.12)		(0.17)	(0.11)
HPS × Land	0.22	0.22	0.22	0.22	0.06	0.22	0.16
	(0.17)	(0.17)	(0.17)	(0.17)		(0.17)	(0.14)
Grid × Adults	0.12	0.12	0.12	0.12	0.12	0.12	0.12
	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)
Diesel × Adults	0.09	0.09	0.09	0.09	0.09	0.09	0.09
	(0.04)	(0.03)	(0.04)	(0.04)		(0.06)	(0.03)
Own solar × Adults	0.09	0.10	0.09	0.09	0.09	0.09	0.09
	(0.03)	(0.02)	(0.03)	(0.03)		(0.06)	(0.03)
HPS × Adults	0.09	0.09	0.09	0.09	0.10	0.09	0.09
	(0.04)	(0.04)	(0.04)	(0.04)		(0.04)	(0.03)
Grid × Pukka	0.42	0.36	0.42	0.42	0.42	0.42	0.43
	(0.10)	(0.09)	(0.10)	(0.10)		(0.10)	(0.10)
Diesel × Pukka	0.13	0.20	0.13	0.13	0.10	0.13	0.17
	(0.14)	(0.11)	(0.14)	(0.14)		(0.16)	(0.15)
Own solar × Pukka	0.12	0.16	0.12	0.12	0.10	0.12	0.11
	(0.12)	(0.10)	(0.12)	(0.12)		(0.15)	(0.12)
HPS × Pukka	-0.03	-0.03	-0.03	-0.01	0.11	-0.03	0.07
	(0.19)	(0.19)	(0.21)	(0.21)		(0.20)	(0.18)
Grid × Lit	0.10	0.08	0.10	0.09	0.10	0.10	0.10
	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)
Diesel × Lit	0.09	0.08	0.09	0.09	0.05	0.09	0.09
	(0.03)	(0.02)	(0.03)	(0.03)		(0.06)	(0.03)
Own solar × Lit	0.02	0.05	0.02	0.02	0.05	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.03)		(0.03)	(0.02)
HPS × Lit	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	(0.04)	(0.04)	(0.04)	(0.04)		(0.04)	(0.03)
Grid × Roof	0.58	0.52	0.58	0.57	0.58	0.58	0.57
	(0.10)	(0.09)	(0.10)	(0.10)		(0.10)	(0.10)
Diesel × Roof	0.20	0.28	0.20	0.20	0.26	0.20	0.20
	(0.13)	(0.11)	(0.13)	(0.13)		(0.17)	(0.13)
Own solar × Roof	0.41	0.44	0.42	0.41	0.28	0.42	0.36
	(0.12)	(0.09)	(0.12)	(0.12)		(0.29)	(0.13)
HPS × Roof	-0.00	0.00	0.00	0.02	0.26	0.00	0.06
	(0.18)	(0.18)	(0.19)	(0.20)		(0.18)	(0.22)
Grid × Rooms	0.13	0.14	0.13	0.13	0.13	0.13	0.13
	(0.03)	(0.03)	(0.03)	(0.03)		(0.03)	(0.03)
Diesel × Rooms	0.15	0.15	0.15	0.15	0.16	0.15	0.15
	(0.05)	(0.03)	(0.05)	(0.05)		(0.10)	(0.04)
Own solar × Rooms	0.18	0.17	0.18	0.18	0.16	0.18	0.18
	(0.04)	(0.03)	(0.04)	(0.04)		(0.11)	(0.04)
HPS × Rooms	0.10	0.10	0.10	0.10	0.16	0.10	0.12
	(0.07)	(0.07)	(0.07)	(0.06)		(0.07)	(0.06)
σ_1	-	0.55	0.01	0.06	0.95	0.01	0.17
		(0.20)	(0.28)	(0.33)		(0.50)	(0.34)
σ_2	-	-	-	-	-	0.01	0.42
						(0.39)	(0.54)
σ_2	-	-	-	-	-	0.39	0.54
Number of Observations	8822.00	8822.00	8822.00	8822.00	8822.00	8822.00	8822.00
Log likelihood	-5791.35	-5789.35	-5791.39	-5791.34	-5789.26	-5791.41	-5791.32
LR test statistic	-	4.01	-0.07	0.03	4.20	-0.10	0.08
LR test p value	-	0.05	1.00	0.85	0.04	1.00	0.78

The likelihood Ratio test statistic: $LR = -2\{LL(\theta_{constrained}) - LL(\theta_{unconstrained})\}$ Each of the nested-logit specifications (columns 2 through 7) are tested against the constrained multinomial logit specification in column 1. LR is distributed $\chi 2$ with degrees of freedom equal to the number of constraints on θ. LL is the negative of the optimized objective function in MATLAB (which is defined as the negative of the sum of the individual household contributions to log of the likelihood function).

Table A8: Two-Stage Least Squares Estimates for Demand for Electricity [WILL BE UPDATED]

	(1) (Grid, Diesel, HPS) (Own Solar)	(2) (Grid, Diesel) (Solar)	(3) (Grid) (Off-Grid)	(4) Multinomial Logit
Price (Rs. 100)	-2.051***	-2.082***	-2.116**	-1.918**
	(0.769)	(0.775)	(0.876)	(0.746)
Hours of supply on peak	0.263	0.401	0.439*	0.437*
	(0.243)	(0.259)	(0.261)	(0.255)
Hours of supply off peak	-0.110**	-0.139**	-0.145**	-0.144***
	(0.0534)	(0.0567)	(0.0571)	(0.0555)
ξ_{tj} mean effects	Yes	Yes	Yes	Yes
Observations	1000	1000	1000	1000

The table presents 2SLS estimates of our demand system for different first-stge nest specifications. The dependent variable is mean indirect utility at the market × survey wave level retrived from the non-linear first stage. The first column, uses our preferred first stage nest-specification of grouping grid, diesel, and HPS in one nest and own-solar in its own nest. The estimates in the first column are the same as column 2 of Table ??. The second column uses a first-stage model with grid and diesel in one nest and both solar technologies in another. In the third column, we group grid in its own nest and all off-grid technologies in a second nest. In the last column, we use the mean indirect utilities derived from a multinomial logit first stage. We instrument price with our experimentally varied HPS treatment assignment. Peak hours refers to supply of electrricity during the evening (5pm-10pm). All regressions control for wave × source mean effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors cluster at the village level in parentheses.

Table A9: First-Stage of 2SLS Estimates for Demand for Electricity

	Price First Stage (only price IV) (1)	Price First Stage (price and hours IV) (2)	Peak hours First Stage (3)	Off-peak hours First Stage (4)
Normal Price	0.045	0.045	0.0050	-0.0046
	(0.035)	(0.035)	(0.0050)	(0.030)
Subsidy Price	-0.14***	-0.14***	0.0075	0.014
	(0.031)	(0.031)	(0.0063)	(0.031)
Hours of supply on peak	-0.045			
	(0.045)			
Hours of supply off peak	0.0067			
	(0.012)			
Peak Hours Instrument		-0.032	0.94***	0.19
		(0.044)	(0.063)	(0.15)
Peak Hours Instrument		0.0037	0.032**	0.88***
		(0.0091)	(0.013)	(0.030)
ξ_{tj} mean effects	Yes	Yes	Yes	Yes
Observations	1000	1000	1000	1000
fstat				

The table presents the first stage estimates of the 2SLS estimates provided in column 2 and 3 of Table 7. The construction of the instrument for hours of supply is outlined in Section 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors cluster at the village level in parentheses.

Table A10: Price Elasticity of Electricity Sources

	Price elasticity
Grid	-0.74
Diesel	-1.34
Own solar	-3.48
Microgrid solar	-2.34

The price elasticity for a given technology is calculated by calculating the percent change in its market share induced by a ten percent increase in its price above its mean endline 1 price.

Table A11: Counterfactual Analysis: Assumptions

Scenario	Source availability	Source hours (peak)	Source price
		Panel A: Theft Reductions	
Data Model	Endline 1	Endline 1	Endline 1
		Panel B: Value of solar innovation	
Solar nowhere	Endline 1 for grid and diesel, solar nowhere	Endline 1	Endline 1
Solar everywhere	Endline 1 for grid and diesel, solar everywhere	Endline 1	Endline 1
Solar cost falls	Endline 1 for grid and diesel, solar everywhere	Endline 1	Reduction in HPS price from INR 170 to INR 120, according to the "solar cost falls" scenario in Figure 8. Own solar price is proportionally decreased.
Further solar innovation	Endline 1 for grid and diesel, solar everywhere	Endline 1	Own solar: 55% reduction in solar panel cost (NREL Fig 8, low estimates, pp. 17) and a 75% reduction in batteries (DOE, pp. 2). Panel is 36% of total cost and batteries is 19% of total cost. These numbers imply a reduction in total own solar price by 34%. Mean own solar price is reduced by 34%. All other solar costs (meter charger, wiring, labour, transport, other are assumed constant). Mean HPS price is similarly reduced by 34%. Prices of all other sources are according to endline 1.
		Panel C: Grid Extension	
Grid nowhere	Endline 1 for solar and diesel, grid nowhere	Endline 1	Endline 1
Grid Everywhere	Grid everywhere, endline 1 for diesel, solar everywhere	Endline 1	Endline 1
Grid 2 extra peak hours	Endline 1 for grid and diesel, solar everywhere	Two additional peak hours for grid, endline 1 peak hours for all other sources.	Endline 1
		Panel D: Theft Reductions	
Remove theft by raising grid price	Endline 1 for grid and diesel, solar everywhere	Endline 1	Grid at INR 140, all else at endline 1. The grid price was derived as follows: grid price in the model estimation is presently defined as the reported bill value in the surveys multiplied by payment rate, where payment rate is the mean of the responses to "Do you pay your bill?" in the endline 2 survey. We therefore define the "remove theft" counterfactual by using the reported bill value as full price un-scaled by payment rate. This yields INR 140."
Increase peak supply and raise prices	Endline 1 for grid and diesel, solar everywhere	Grid peak hour = 5 hours everywhere, all else at endline 1	Grid at INR 95 everywhere, all else at endline 1. INR 95 is derived as follows: grid peak supply is set to 5 hours everywhere. Price is set so that total loss per HH in the endline sample is equal to that obtained in row 3 (model with solar everywhere).

Note: Household characteristics and source off-peak hours are unchanged (at their endline 1 levels) across all counterfactual cases and hence omitted from the table.

A.3 Construction of instrument for hours of supply

The imputation of missing feeder supply data is done in two steps.

1. If supply data for a village is missing for a given survey wave but some data is available for that village in the $+/-6$ month window, we replace the missing observation with the temporal mean of available data in this time window.

2. If a village has missing data for all months in the $+/-6$ month time window, we use a random forest (RF) algorithm to impute missing hours of supply for that village.

RF has the advantage that it necessarily yields internal predictions and so imputed hours of supply are sensibly bounded. We include the following predictor variables (features) in the RF model: (1) hours of supply of the three nearest villages for which we do have data, (2) division fixed effects, (3) polynomials upto degree 5 of district-demeaned latitude and longitude of each village, and (4) interactions of division fixed effects with each of the demeaned lat-lon polynomials. Hence, all the features that go into the RF model are plausibly exogenous to our demand model. We exclude unelectrified villages (supply for these is replaced zero) from the sample because we don't want to use data of unelectrified villages to impute missing supply data for electrified villages. For instance, there are 56 electrified villages in the endline survey which have non-mising data. This is our master sample for the RF model. We randomly select 80% of this sample (45 obs) as the training sample and the remaining 20% as the testing sample (11 obs). The RF model is fit on the training sample.

Figures ?? and ?? describe our prediction model. The main parameter to tune in a random Forest model is the number of candidate variables to select from at each split. To do this, we start with 2 variables and increase by a step factor of 1.5 until the improvement in out-of-box (OOB) error is less than one percent. As shown in panel A of Figure ??, for the endline 1 data, this yields 6 variables. Figure ?? shows the most relevant variables chosen by the model.

The RMSE of our prediction model is 1.9 hours. We take the predictions from this model and use them to impute missing observations in administrative grid supply data.

In one our second-stage two-stage least squares specifications (Table 7, column 3), we instrument for peak and off-peak hours of supply of electricity, in addition to price. HPS, diesel, and own solar have constant supply hours in all villages. For HPS and diesel this is 0 for off-peak hours and 5 (maximum) for peak hours. For own solar, it is the global median of the peak and off-peak hours for that source. For grid, supply is observed from administrative log-books at the feeder level, mapped to sample villages. We use our predictions from the above random forest model as the instrument for grid supply.