# Demand for Electricity on the Global Electrification Frontier

Robin Burgess, Michael Greenstone, Nicholas Ryan and Anant Sudarshan

January 2023 No: 1445

Warwick Economics Research Papers

ISSN 2059-4283 (online)
ISSN 0083-7350 (print)

# Demand for Electricity on the Global Electrification Frontier*

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

January 6, 2023

**Abstract**

Falling off-grid solar prices and an expanding grid are revolutionizing choices for nearly a billion people without electricity. Using experimental price variation, we estimate demand for all electricity sources in Bihar, India, during a four-year period when electrification leapt from 27% to 64%. We find that: (i) household surplus from electrification increased five-fold; (ii) both solar and the grid boost electrification but households gain more surplus from the grid; (iii) grid investments and subsidies strongly reduce demand for off-grid solar. When we extend the model to eight African countries where grid infrastructure is weaker and subsidies lower we find that off-grid solar often provides higher surplus than the grid.

JEL Codes: O13, Q41, Q21, C93

*We thank Manoj Sinha and Col. Baljit Singh of Husk Power Systems for their partnership in this project. We thank Jennifer Allard, Shruti Bhimsaria, Menna Bishop, Catherine Che, Ananya Kotia, Rakesh Pandey, Aditya Petwal, Victor Quintas-Martinez, Johanna Rayl 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. The project is filed under AEA RCT registry #AEARCTR-0000132. We thank seminar participants at Berkeley, Duke, Harvard, LSE, MIT, and Stanford, as well as Nikhil Agarwal, Liran Einav, Steven Puller, Joe Shapiro, Reed Walker and Catherine Wolfram, for their comments. All errors are our own.

†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 [Department of Economics, University of Warwick: anant.sudarshan@warwick.ac.uk].

# 1 Introduction

The global electrification frontier is the collection of places in the world where households are getting electricity for the first time. The steady movement of this frontier, in the United States from 1935 onwards, Brazil from the 1960s, South Korea in the 1970s and China from the 1980s has been inseparable from economic growth. Nevertheless nearly a billion people, mainly in South Asia and sub-Saharan Africa, remain without electricity and major investments in distribution infrastructure, newer decentralized technologies, and subsidized connections are being made in order to reap the economic gains from pushing the frontier outwards (International Energy Agency, 2017).

In the traditional mode of electrification, the global electrification frontier was a literal boundary, defined by the extent of the grid, with households filling in behind it (Lee et al., 2014). More recently however, the rapid decline in the cost of solar power has transformed the shape of the frontier. Solar panels can supply the grid, but, unlike other sources of power, they can also generate efficiently at a small scale, on the roof of a single household. This has meant that rural households can choose whether to get solar power, regardless of whether the grid has reached them. Innovation in solar has, in effect, expanded the electricity choice set for millions of households.

The advent of off-grid solar has spurred hope of a faster, greener path to universal electrification.[1] This optimism has a real justification: off-grid solar's market share has skyrocketed. Figure 1 shows household solar ownership across time based on household data sets that we have put together for Indian states and African countries. In these places on the global electrification frontier off-grid solar now powers 30% to 60% of rural homes having contributed next to nothing just a decade ago.

The confluence of "big push" grid expansions and rapid diffusion of off-grid implies households in many developing countries now have a choice between competing electricity sources supplying power at different prices and quality. It is this choice, which is key for the nearly billion people on the electrification frontier, that we study in this paper. This requires that we estimate the demand for electricity, over *all* available sources of power in order to understand how households are make their choices. Doing this allows

[1] 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." ("Powering Sustainable Energy for All," *The New York Times*, January 11th, 2012. See also "Africa Unplugged: Small-scale Solar Power is Surging Ahead", *The Economist*, October 29th, 2016.) UN Sustainable Development Goal #7 is to "ensure access to affordable, reliable, sustainable, and modern energy" and targets increasing the share of renewable energy in the global energy mix in particular. Nearly all large-scale aid programs in the power sector include significant on-grid and off-grid components. USAID, for example, launched *Power Africa* in 2013 and DFID launched *Energy Africa* in 2015, both of which invest in off-grid renewable electricity.

us quantify household surplus from electrification and to attribute this value to different sources. It also allows us to study how changes in technology and policy that are taking place on the frontier will affect these choices. Innovation driven falls in solar panel prices are, for example, behind the dramatic market penetration we observe in Figure 1 but at the same time government subsidies are bringing the grid within the reach of rural populations.

Our main empirical setting is Bihar, an Indian state containing over 100 million people, which is a typical outpost on the global electrification frontier. Between 2000 and 2016, India contributed over 80% of the total gain in the number of households in the world connected to the grid (International Energy Agency, 2017). In our rural study areas between 2013 and 2017 off-grid solar was also diffusing at breakneck speed (see Figure 1). Thus while the grid has expanded rapidly, electricity in the state of Bihar, as in many developing countries, has become a differentiated product. There are several sources of electricity, including both the public grid and private off-grid sources, which differ in price, the load they can support, hours of supply and other features.

As the choice between competing technologies has become a defining feature of household electrification on the global frontier we model the demand for electricity with a discrete choice, random coefficients demand model (McFadden, 1974; Berry, Levinsohn and Pakes, 2004). Households choose between four electricity sources—the grid, diesel generators, solar microgrids, and their own off-grid solar systems—and an outside option of no electricity. We allow for observed heterogeneity in household and source characteristics using an original household panel survey. We also allow for unobserved variation in household preferences and in the quality of different electricity sources across villages and time (Berry, 1994).

This last feature is critical to capture rapid changes in new goods like off-grid solar power. Unlike the grid, data on decentralized private sources of power is generally unavailable in administrative records. We therefore draw upon multiple field surveys conducted over four years. A contribution of the paper is to get careful measurement of the characteristics of *all* sources of electricity as well as household choices over them during a period of rapid change. This required close collaboration with the Government of Bihar that operates the grid as well with a private solar provider.

With the latter we ran a randomised experiment that introduced a new product, solar microgrids, and varied its price across 100 village-level markets for two and a half years. We were then able to use this experimental variation in price to estimate a discrete choice demand model. The availability of experimental variation in price to estimate a discrete choice demand model is extraordinarily rare and removes the need to rely on traditional assumptions, about market conduct or the structure of demand shocks, to generate instrumental variables (Berry, Levinsohn and Pakes, 1995; Hausman, 1996; Nevo, 2001). Indeed in our setting we find that the experimental variation is necessary for unbiased and precise estimates of the price elasticity of demand. Our experiment also varies price over a multi-year period, which lessens external validity concerns that arise with very short-run discounts. Overall we estimate highly elastic demand for solar power ($\varepsilon \approx -2$).

To place the experiment within a broader context, we collect comprehensive data on the demand and supply sides of the village electricity markets in our sample over the four years from 2013 to 2017. During this period electrification increased by almost *40 percentage points*. As a basis of comparison, this surge in electrification occurred in half the time it took to increase electrification rates by the same amount in the rural United States, after the passage of the landmark Rural Electrification Act in 1936.[2] The gains in electrification in Bihar were due to the same two factors, the advent of off-grid solar and a "big push" on the grid, that are reshaping electrification around the world. Our setting therefore allows us to use experimental estimates of demand to value historic changes in energy access that are externally relevant for the experience all along the global electrification frontier.

We draw four main conclusions. First, choice matters, in that there is huge heterogeneity in the electricity sources households choose across markets and time. The dramatic expansion of choice we observe in Bihar is happening across the developing world (Figure 1). Solar market shares, across a range of African countries and Indian states, recently range from the low single digits to up to 60% (Figure 1). Within our data in Bihar solar is the leading source of electricity in 35% of village-years, diesel the leading source in 18% and the grid in 43%[3]. Households fluidly substitute between alternative sources that provide similar energy services. These facts underscore the importance of treating electricity as a differentiated product as the choice set for households expands across the developing world. Frameworks and data sets need to be developed to allow us to look at how households make these choices in different contexts.

Second, the increases in energy access observed during our sample period increased the household surplus from electrification by roughly 5×. Both off-grid and grid sources drive these gains. During our sample period off-grid solar in Bihar emerges as potent force for increasing electrification pulling millions into ba-

[2]The same increase for rural (farm) households in the United States took 9 years, during and after World War II, from 1939 to 1948 (Bureau of the Census, 1975).

[3] Percentage for solar is a combination of micro grid and off-grid solar. Across village-years, 5% of villages had no electricity available.

sic levels of electricity usage as solar prices fall. However, it is clear that when available and subsidized consumers prefer the grid. The advent of off-grid solar alone would have increased consumer surplus by 1.5× but of the grid alone by 4.2×. This underlines how important it to jointly quantify the contribution of *all* sources something that is often missed in traditional impact evaluations which focus on single sources. The advantage of the grid is due in part to policies that subsidize energy access. In our study period, the government of Bihar built out the grid from 29% to 72% of villages and subsidized connection costs for households below the poverty line to zero. In addition, once households are on the grid, the price of energy is roughly 27% of cost[4]. The gains in social surplus from grid extension are therefore smaller than the gains in household surplus (2.6×, as compared to 4.2×), because grid extension implies additional transfers from the government to poor households.

Third, our model suggests that much of future growth in electrification in Bihar will come from the grid, since households prefer grid electricity as they become wealthier. This finding follows from our demand estimates, in which households with a solid roof, a common indicator for wealth, have roughly *double* the probability of choosing grid electricity as households without. Wealthier households' preference for the grid likely stems from their wish to run higher load appliances, like fans and televisions, which off-grid sources usually do not support. To forecast the future path of electrification in Bihar, we run a counterfactual where solar costs continue their decline, the grid reaches all villages, the number of hours electricity is available on the grid increases and all households achieve at least the median income. In this scenario, despite the projected fall in the price of solar, most households move to the grid.

Fourth, household choices depend fundamentally on national energy policies. This is because policy determines not just the extent of the grid but also the characteristics of the power it provides - hours of supply and price. To probe the external validity of our results for Bihar we applied our model to household data sets which are representative of a rural population of 500 million in 8 African countries on the global electrification frontier. Here the contribution of off-grid solar to household surplus is larger pointing to the limited extent, high costs and low quality of the grid. These characteristics push rural populations to rely more on solar energy and help us understand both rising solar shares (Figure 1) and the unwillingness to pay to connect to grid electricity found in some African studies (Lee et al., 2014; Lee, Miguel and Wolfram, 2020*b*). At the supply conditions observed in the data, in each country, Bihar has a greater surplus

[4] Assuming 3.88 INR per kWh average procurement rate and average consumption of 60 kWh per month as indicated by out administrative data.

from electrification than even richer African countries (Kenya, Nigeria). In several parts of Africa, solar contributes more to the household surplus from electrification than does the grid (Ethiopia, Uganda, Mali). However, counterfactually, if households in Africa faced the Indian grid—with near universal extent and heavily subsidized energy prices—the grid would dominate energy choices and surplus there, just as it does in Bihar.

Our finding on the path of electrification running mainly on the grid therefore depend on a country's willingness to bear energy subsidies to achieve grid electrification. Taking everything together we see that off-grid solar can help to drive electrification in particular when the grid is not present, of low quality or unsubsidised. It is in this sense that this innovation is important for the world. But when the quality and extent of the grid get better then reliance on off-grid seems to fall away. So there not a sense in which it will leap-frog the grid. However, in places where governments are unwilling or unable to pour money into improving the extent and quality of the grid it is likely to remain a key and growing source of electricity as solar prices fall. Solar will continue to serve the poor where the grid is priced at cost, at least until households reach a significantly higher level of income.

Our paper makes two contributions to the literature on electricity access in developing countries. First, we study household demand for electrification, a revealed-preference measure of the value of electricity, whereas most of the literature has measured the impact of access to electricity on a range of economic and welfare outcomes.[5] Second, we estimate how households value both grid and off-grid electricity together, in a single demand system, which allows us to study substitution between sources. In doing so we pay close attention to how household choices depend on the characteristics of electricity from different sources. Prior work has estimated the demand for individual sources of electricity, rather than electricity generally, leaving the choice set and the pattern of substitution unspecified.[6,7]

More broadly, this study joins a methodological movement in the development literature that combines

[5]Prior work has found that electricity access causes large increases in labor supply (Dinkelman, 2011), industrial output (Rud, 2012; Allcott, Collard-Wexler and O'Connell, 2016), manufacturing productivity (Kline and Moretti, 2014), agricultural productivity (Kitchens and Fishback, 2015), land values (Lewis and Severnini, 2019), and proxies for household welfare, such as the human development index and indoor air quality (Lipscomb, Mobarak and Barham, 2013; Barron and Torero, 2017). See Lee, Miguel and Wolfram (2020*a*) for a review of the impacts of electrification.

[6]In contemporaneous experiments, Lee, Miguel and Wolfram (2020*b*) estimate demand for grid connections in Kenya and Grimm et al. (2020) estimate demand for off-grid solar technologies in Rwanda. Aklin et al. (2018) study how household characteristics predict solar take-up in India.

[7]A couple papers have hypothesized that solar and grid electricity are imperfect substitutes for the rural poor. Fowlie et al. (2019) suggest that a promise of future grid connections, in Rajasthan, India, may have reduced the take-up of off-grid sources like microgrids. Lee, Miguel and Wolfram (2016) report the results from a household survey in Kenya showing that grid users own more high-load appliances than solar users.

structural models with experimental variation to aid in the interpretation and increase the external validity of experimental results.[8] Our study combines experimental price variation with a structural demand model allowing rich observed and unobserved heterogeneity. Our finding on the large gap between the value of electrification and the value of one electricity source shows the value of a structural model for placing experimental estimates of demand, for any given product, in the context of the broader product market. Our paper is therefore related to the literature in industrial organization on estimating the value of new products (Hausman, 1996; Petrin, 2002; Goolsbee and Petrin, 2004). Experimental work in development economics tends to estimate the demand for one product at a time, which may, as in our case for electricity, greatly understate the demand for the categories to which products belong.[9]

# 2 The Changing Electricity Landscape in Bihar

This section introduces the policy context and our data sources. We then use the data to describe the transformation of the electricity market in Bihar during our study.

## 2.1 Context

Bihar, with a population of 104 million (Census of India, 2011), is one of India's poorest states. In 2012, the year before our baseline survey, the electrification rate in Bihar was only 25%, below the rate of 37% in sub-Saharan Africa and about one-third of the all-India rate of 79%. The average Bihari used just 122 kWh of electricity per year, less than one percent of the level in the United States. At this level of consumption, which is an average, including many households with no electricity at all, a person can power two light bulbs totaling 60 watts for six hours per day. The low level of consumption is an equilibrium outcome. Demand for electricity is low because many households are poor. Supply of electricity is limited, on both the extensive margin, since many villages are not on the grid, and the intensive margin, since supply is rationed.

[8] Examples include studies that use experiments to help estimate structural models of fertility, education, labor supply, migration and human capital, and enforcement of plant emission standards, though not demand for electricity (Todd and Wolpin, 2006; Attanasio, Meghir and Santiago, 2012; Duflo, Hanna and Ryan, 2012; Bryan, Chowdhury and Mobarak, 2014; Galiani, Murphy and Pantano, 2015; Duflo et al., 2018; Attanasio et al., 2020).

[9] Experimental estimates of demand have been an enormous area of growth in development economics and are used both to test theories of behavior and to consider optimal policy. Preventive health products (Berry, Fischer and Guiteras, 2020; Peletz et al., 2017; Dupas, 2014) and financial services (Bertrand et al., 2010; Karlan and Zinman, 2018, 2019) are two prominent markets in which experiments have been used to estimate demand. Though many products in these markets arguably have close substitutes, few studies that experimentally estimate demand explicitly model substitution. Kremer et al. (2011) is a close precedent that experimentally varies the quality of a good, a local water source, and estimates a demand model using observable variation in walking distance to water sources as a proxy for price.

Our study was well-timed to capture two big changes in the electricity market. First, beginning before and carrying on through our study period, a decline in the price of solar panels made off-grid solar a feasible alternative to grid power. Second, in response to low rates of electrification in states like Bihar, the Government of India funded large campaigns for grid extension and household connections.

The first force changing the electricity landscape therefore is a fall in the price of solar power. Innovations in this sector has meant that the price of solar power has been declining rapidly for several decades. But only in the last decade has it reached a level low enough to make off-grid solar a viable choice for poor households and this is reflected in rapid uptake of the technology both in India and Africa (Figure 1). Our data reflects these trends. The price of own solar systems fell 10% during our data collection from INR 80 at baseline to INR 72 at follow-up. This lower price is likely not for the same energy service, but a better one since solar panels have grown more efficient and batteries and other systems more reliable. Solar vendors also entered smaller towns closer to villages, effectively lowering connection costs.

The second force was a "big push" for grid electrification at both the national and state levels. In his 2015 independence day address, Indian Prime Minister Narendra Modi launched a rural electrification program with a thousand-day deadline to electrify the remaining 18,452 census villages still without access, at an estimated cost of USD 11 billion.[10] Even when the grid reaches a village, poor households may not connect, or may take a long time to do so (Lee et al., 2014). The Government of India therefore started a complementary USD 2.5 billion program to subsidize states to give out infill connections in electrified villages at no cost to households.[11]

In Bihar, the state government made electricity access a priority (Kumar, 2019). Nitish Kumar, Bihar's six-time Chief Minister, invested heavily in grid electrification, using both central and state funds, and promised universal household electrification as part of his reelection campaign (Business Today, 2017). During our four-year study period, the state's own data report over 7 million new electricity connections, representing a staggering 51 pp increase in the statewide household grid electrification rate. The government not only invested in grid extension but also held camps to sign-up households and heavily subsidized connections, including by offering connections for free to all households designated as Below the Poverty Line

[10] The village-level goal was declared achieved ahead of schedule on April 28, 2018. The government definition of a village being electrified requires that public spaces such as schools and health centers have access to power and a minimum of 10% of households are connected to the grid. The target is out of a total of almost 600,000 census villages in India. This program, the Deen Dayal Upadhyaya Gram Jyoti Yojana (DDUGJY), is a continuation, under a new name, of the prior government's Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY), which had similar objectives but fell short of reaching all villages (Government of India, 2015; Burlig and Preonas, 2016).

[11] The Pradhan Mantri Sahaj Bijli Har Ghar Yojana, known as Saubhagya, launched in September 2017.

(BPL). Heavy state investment in this period allowed the grid to reach progressively poorer households.

## 2.2 Study Sample and Data

Our study sample consists of 100 villages in two districts in Bihar (Figure 2). We worked in these villages in partnership with a solar microgrid provider called Husk Power Systems (HPS) (see Section 3.1). These villages were sampled from a broader population with poor access to electricity at baseline. Poor access implied meeting two criteria. First, eligible villages had to be listed as unelectrified by the government. Second, they could not have been previously served by HPS although we restricted attention to villages that were nevertheless close enough to existing service areas for it to be logistically feasible for the company to expand into them. We selected 100 villages that met these criteria, totaling 48,979 households. A number of the study villages, in West Champaran district, are clustered near the border between Bihar and Uttar Pradesh, with one village being part of Uttar Pradesh.

We collect data from both the demand and supply sides of the market over a nearly four-year period. We use four main sources. First, on the demand side, a household-level panel survey on the sources and uses of electricity. Second, on the supply side, household-level administrative data on customer enrollment and payments for solar microgrid connections from HPS. Third, on the supply side, village-level survey data from the operators of common diesel generators, an off-grid source of electricity. Fourth, on the supply side, household-level administrative data from the state utility on customer billing and payments, as well as village-level electricity supply. We describe the survey here and the rest of the data sources in Appendix A.

Our household panel survey sampled 30 households per village to cover about 3,000 households, containing about 18,000 people, across the 100 sample villages. The sample was drawn to represent those with an interest in a microgrid solar connection, but, because this screening for interest was loose, in practice the sample is nearly representative of the population as a whole.[12]

The survey has three rounds: two thick rounds, which we call baseline and endline, and one thin round, which we call follow-up. The baseline survey took place in November and December of 2013, the endline from May to July of 2016, and the follow-up in May 2017 (Appendix Figure A1 shows the timing of survey rounds). The two thick rounds used nearly the same survey instrument and covered demographics, the

[12] We ran an initial customer identification survey 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, because households were not required to put down a deposit, nor were they held to their initial statement of interest when the product was later 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.

sources and uses of electricity, and welfare outcomes likely to be influenced by electricity use. The follow-up round took place one year after the endline for the experiment. This round was not part of our original plan but was added to update household electricity sources and choices, in light of the massive changes we observed on the supply side. All told, the three surveys cover a three and half year period of extraordinary activity on the energy frontier during which time electrification rates increased sharply, and different energy technologies gained or lost market share.

### 2.3 Characteristics of electricity sources

In developed countries, electricity is the archetype of a homogeneous good: power is available from the grid $24 \times 7$ and can run all kinds of appliances. In many developing countries electricity connections are differentiated products. This part describes the characteristics of different electricity sources in our sample.

There are four electricity sources used by households in our sample including grid electricity and three off-grid sources. *Grid electricity* is provided by a state-run distribution company. Grid electricity is only available to households if the grid has reached their village and if they apply for a connection. There are three off-grid electricity sources: microgrid solar, own solar, and diesel generators, all provided in private markets. A *solar microgrid* is a system composed of a solar panel and batteries that serves a small group of six to nine households.[13] An *own solar* system is a panel and battery bought and operated by a single household. A *diesel generator*, in our context, is a generator set up by an entrepreneur and run with diesel fuel to supply electricity to a large group of households in a single village. Diesel generators serve 100 customers on average, with a range from 60 to 200 in our sample. The outside option is not to have electricity from any of these sources which means relying on kerosene, candles, or battery torches for lighting.

Table 1 gives summary statistics on the characteristics of electricity sources at baseline (columns 1 through 5), endline (columns 6 through 10) and follow-up (columns 11 through 15). Sources differ on a number of dimensions including energy services, price and reliability. Panel A reports on source characteristics: monthly price, the total connected load of appliances a household using each source has plugged in, hours of supply (which we disaggregate into peak and off-peak hours), and the share of villages in which a source is present.[14] Panel B reports appliance ownership conditioned on the household owning the column

[13] The microgrids in our context are offered by Husk Power Systems, our partner in the experiment. The HPS microgrid consists of a 240 watt panel and a separate, 3.2 volt rechargeable battery and meter for each household. Households have a key pad to secure access to the battery and must purchase codes on a monthly basis to keep using the system. Each household on the microgrid gets 25 to 40 watts of power. To compensate for the small load, the system is bundled with two high-efficiency LED bulbs and an electrical outlet, typically used for mobile phone charging.

[14] Properly, the connected load of appliances is not a characteristic of a source, but depends on household appliance purchases.

source. We highlight four findings that characterize the trade-offs households face in choosing an electricity source.

First, the grid can support higher loads and therefore a wider range of energy services than other sources. Most households connected to any source use power for lighting and mobile phone charging (Table 1, panel B). Among grid-connected households, in addition, 22 percent own a fan and 15 percent a television, whereas fewer households with other sources of electricity own these appliances. Households on the grid have the largest connected loads, by far, in all survey waves (panel A, comparing columns 1 through 4).

Second, the pricing in the market is fairly tightly clustered.[15] At baseline, three sources have average monthly prices from INR 72 to INR 99 per month (Table 1, panel A, columns 1 to 3). The grid is the cheapest energy source, due to subsidies and non-payment of bills, but would be among the most expensive if grid electricity were priced at cost and households were forced to pay their bills in full to receive power.[16] The highest-priced product at baseline, above this tight cluster, is microgrids, with a price of INR 200 per month. Our experiment varied the price of this product (Section 3 3.1).

Third, the grid is not as reliable as other sources during the evening peak, when households most want electricity. The mean grid supply in the peak hours, from 5 to 10 pm, was only 2 hours per day at baseline and endline, increasing to 3 hours at follow-up. Even this low average understates the trouble with grid supply, since on one day out of four there is no grid supply at all (Appendix Figures A2 and A3 show the distributions of hours of supply for the grid, in total, off-peak and on-peak). All other sources of power provide more supply during the peak hours in all survey waves.

Fourth, the availability of different sources changed dramatically over the nearly four years of our data collection. The grid was present in 29% of all villages at baseline (Table 1, panel A, column 1), 53% at endline (column 6) and 72% at endline. The availability of diesel *fell* from 57% (column 2) to 13% of villages (column 12) in the same span, predominantly because of providers exiting as they lost market-share

We describe connected load as if it were a source characteristic, because the connected load for all sources except the grid is effectively capped by the maximum power the source technology can provide.

[15] Households pay up front for home solar systems, so we have amortized the cost of these systems into a monthly price equivalent. For own solar, household systems, once purchased, have no operating costs. To make the price comparable to other sources, which are paid monthly, we amortize the capital costs of own solar using an assumed lifespan of seven years and a 20% interest rate. For the grid, we take the monthly price to be the self-reported monthly payment for grid electricity, averaged across formal and informal households on the grid. Grid electricity is in principle charged on a volumetric tariff; however, a minimum monthly payment and infrequent meter reading imply that many poor consumers are *de facto* billed at a flat monthly rate.

[16] The *de facto* grid price is INR 72 per month at baseline and INR 60 at endline. Informality acts as a large price cut for the grid. Of the 158 households using the grid at baseline, only 47% answered yes to the question "Do you pay electricity bills?" The posted electricity tariff would imply payments of INR 153 per month at baseline. This would make the grid amongst the most costly sources. With incomplete payments, monthly costs fall to INR 72 per month making the grid one of the cheapest options. We calculate the break-even price for grid electricity payments to cover variable supply costs to be INR 233.

to the grid. We assume that own solar systems are available in all villages, since households can travel to buy these systems at nearby markets.

The picture of the electricity market in Bihar is therefore variegated. Grid electricity serves all loads and is cheap, from a household's perspective, but is not widely available and has unreliable service. The hypothetical ability to run a fan is not valuable if a household cannot turn on the light. Off-grid electricity sources provide more limited energy services but more reliably and at a fairly low price. There are surely also other unobserved factors, like the difficulty of obtaining a connection, or whether the household or the operator is responsible for maintenance, that will influence households' choice between electricity sources.

### 2.4 Market shares of electricity sources in Bihar

The two disruptions of solar and grid expansion transformed electricity access during our study. Figure 3 shows the market shares of all electricity sources over time. Each stacked bar gives the share of households, from bottom to top, that use grid electricity, diesel generators, solar microgrids, own solar systems or no electricity. Market shares are calculated with respect to the total sample, regardless of whether a source is available in a village or not; in a village where the grid is not present the grid necessarily has a zero share. There are three groups of bars for shares in the baseline, endline and follow-up survey waves. Within each group, the three bars from left to right give the market shares amongst all households, households that do not have a solid roof, and households that do have a solid roof, respectively. Whether a household has a solid roof is commonly used to measure wealth (Alatas et al., 2012; Haushofer and Shapiro, 2016).

We find that all sources of power contribute to electrification in Bihar. The transformation of Bihar's electricity sector during our study period has three main features.

First, plainly, a surge in the overall electrification rate. Consider the left bar in each group, for all households. The electrification rate from any source, the sum of the colored bar stacks, increased 37 pp, from 27% to 64%, in somewhat less than four years.

Second, a compositional shift, away from diesel and towards solar power and especially grid electricity. The net gain in electrification conceals a churning of market shares across sources, all of which contributed significantly to electrification rates at some point. Diesel generators, the black bar segment (second from bottom), were the most popular source of electricity at baseline, with 17% market share (despite being available in only 57% of villages). By endline, diesel had all but disappeared. Grid electricity (the bottom bar segment, in brown), by contrast, surged, with market share rising from 5% to 25% and then 43%, in successive surveys. No village in our sample had a grid take-up of over 50% at baseline, but 44% did by the follow-up survey. Solar grew steadily as a category. Solar microgrids (third from the bottom, in yellow) increased their share, from nothing to 9% at endline, when subsidies were still offered as part of our experiment, but fell back down a year later. Own solar systems (top colored bar, in orange) picked up the slack, rising from a 5% share at baseline to a 15% share at follow-up, with all of their growth coming between the endline and follow-up rounds.

Third, heterogeneity in household demand, with richer households more likely to have electricity from the grid at any given time. At baseline, the electrification rate among households without a solid roof is little more than half that for households with a solid roof. The two disruptions increased electrification rates for both groups and narrowed this divide, though a gap in electrification rates of 15 pp remained at follow-up. The heterogeneity across households also extends to technology choice. Households with a solid roof are much more likely to have grid electricity, whereas they are somewhat less likely, compared to households without a solid roof, to have off-grid solar.

# 3 Demand for Solar Microgrids

This section describes our experiment and uses the experimental variation to estimate demand for solar microgrids. Quantifying demand for microgrids is important, in its own right, because off-grid solar has emerged as a widespread substitute for grid electricity on the global electrification frontier. We use the demand estimates to calculate the contribution of microgrids to household surplus.

While we start by estimating demand for this new good, microgrids are only one of several competing electricity sources in Bihar (Section 2). Therefore, Section 4 will use the same experimental variation that we introduce here to estimate a richer model of demand over all electricity sources. This model will allow us to estimate how much surplus households get from *all* sources of electricity and to examine how this changes as off-grid solar gets cheaper and the extent and quality of the grid improves.

## 3.1 Experimental design

The falling price of solar has made solar-as-a-service a newly viable business. Husk Power Systems (HPS), a social venture company that supplies off-grid power to villages in Bihar, decided to add the solar microgrid product to its portfolio as a means of reaching a wider set of customers.[17] HPS was the only microgrid provider in our sample and so we treat HPS and microgrids as synonymous hereafter.

We partnered with HPS to vary the availability and price of solar microgrids in a cluster-randomized control trial at the village level. We randomly assigned sample villages into one of three arms: a control arm (34 villages) where HPS did not offer microgrids, a normal price arm (33 villages) where HPS offered microgrids at the prevailing price, initially INR 200 per month, and a subsidized price arm (33 villages) where HPS offered microgrids at a price of INR 100 per month. The normal price arm provides microgrid service at, or slightly above, cost and the subsidized arm at perhaps 40% below cost.[18] Within each treatment village, all households were offered the same HPS connection and pricing, regardless of whether they had previously expressed interest in a microgrid or participated in our baseline survey. Sales of microgrid connections began in January 2014, right after the baseline survey.

The treatment assignments set the initial prices in all villages. Prices of microgrids thereafter changed for two reasons. First, the prevailing or normal price arm was not rigid, but was meant to capture the price at which HPS would offer microgrids, if there had not been an experiment. Husk Power, due to low demand at the initial price of INR 200, chose to cut prices to INR 160 in 11 villages in the normal price arm. Second, the experiment ended with our endline survey, in May 2016, but our data collection runs beyond this survey. After the completion of the experiment and our endline, but before the follow-up survey, Husk Power set the price in all 66 treatment villages to INR 170 per month.[19] HPS did not enter the control villages at any point during our study period. In the demand analysis, we use treatment assignments, and their interactions with survey wave indicators, as exogenous instruments for price.

Table 2 shows the balance of household covariates in our sample including 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. Our rural sample is poorer than the population of Bihar as a whole. Self-reported household

[17]HPS was founded in 2007 to provide electricity in rural areas using biomass gasifiers as generators, fueled by agricultural waste, such as rice husks (hence the name of the company). These biomass plants were subject to fuel supply disruptions and could only serve a village if demand was sufficiently broad.

[18]We estimate the capital and installation costs of a microgrid to be INR 105 per household per month (Appendix Figure D1). This figure is net of capital subsidies provided by the government, which were on the order of 60% in 2014. The service of the system would include additional costs for billing, collection and maintenance. It is therefore reasonable to estimate costs in the range of INR 160 to INR 200 per month, the range of prices offered in our normal price arm.

[19]This price adjustment meant that 22 normal price villages experienced price declines of INR 30 (from 200 to 170); 11 normal price villages experienced a INR 10 increase; and all 33 subsidized price villages saw a substantial increase of INR 70 (from 100 to 170).

incomes imply mean per capita daily income of INR 43 (USD PPP 2.5) at baseline, compared to mean per capita daily income of INR 99 (USD PPP 5.8) across the state.[20] Two-thirds of households own agricultural land and less than half have a solid roof (panel B, column 1).

Table 2, columns 4 and 5 show the differences between normal price and control arms and between subsidized price and control arms, respectively, with standard errors in parentheses. 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. We address this slight imbalance by including household covariates as controls in our demand estimates.

### 3.2 Demand estimates

Table 3 presents estimates of aggregate microgrid demand at the village level. The first three columns show intention to treat (ITT) estimates that regress microgrid market shares $s$ in village $v$ in period $t$ on the experimental treatment assignments:

$$s_{vt} = \beta_0 + \beta_1 T_v^{Subsidized} + \beta_2 T_v^{Normal} + \varepsilon_{vt}. \tag{1}$$

The coefficients in the first two rows report the change in market shares for solar microgrids due to the subsidized and normal price treatments, respectively, and the constant gives the market share of microgrids in the control group. Columns 1 through 3 report estimates for different periods: the baseline (November 2013), endline (May 2016) and follow-up surveys (May 2017), respectively.

The first finding in Table 3 is that the experiment increased solar microgrid penetration. We expect there should be zero take-up at the baseline, because microgrids were a new product, about to be launched. At baseline, in column 1, the estimated constant, representing take-up in the control group, and the estimated normal price and subsidized treatment coefficients are very small and statistically not different from zero. At endline, in column 2, the estimated constant was 2.3 pp (standard error 0.5 pp). The coefficient on the subsidized price treatment shows that it increased solar microgrid take-up by 19.3 pp (standard error 4.9 pp). The coefficient on the normal price treatment is considerably smaller (6.0 pp, standard error 2.8 pp), showing the sensitivity of household take-up to microgrid prices. We find a similar gap in estimated demand

[20] Using a Gross State Domestic Product (GSDP) of Rs 36,143 for year 2014-15 (Bihar State Government, 2015), and a INR per USD PPP rate of 17, per OECD Data for India for 2014.

when using administrative measures of household payments, rather than surveys, to measure take-up.[21]

The second finding in Table 3 is that solar microgrid shares fell sharply between endline and follow-up, after experimental subsidies were withdrawn (column 3). By the follow-up survey, relative to the experimental endline one year prior, the solar microgrid market share in the subsidized price villages had declined by more than 11 pp (58%), and in the normal price villages by 4 pp (67%). In the subsidized treatment arm, the increase in price after the experiment ended would be expected to cut market share. However, the proportional decline in market shares was large in the normal price treatment arm as well, which did not experience a large change in price after the experiment. This suggests that the expiration of subsidies does not explain the entire fall in microgrid market shares, which we investigate further in Section 4.3.

The last two columns of Table 3 give instrumental variables estimates of microgrid demand, where we instrument for the price level (or log of price) using the experimental treatment assignment. For example, the column 4 (linear) IV specification of demand consists of the two stages

$$s_{vt} = \beta_0 + \beta_1 Price_{vt} + \varepsilon_{vt} \tag{2}$$

$$Price_{vt} = \alpha_0 + \alpha_1 T_v^{Subsidized} + \eta_{vt}. \tag{3}$$

A corresponding log-log specification is used in column 5. The sample for these columns is limited to the two-thirds of villages in which microgrids were offered. Consistent with the ITT estimates, we find large, negative and highly significant effects of price on microgrid market share in both linear and log-log specifications of demand. The linear demand estimates imply a choke price, at which demand for the product is zero, of INR 270, with demand increasing by a 0.129 share (standard error 0.052) for each INR 100 cut in price.

### 3.3 Surplus from microgrids

We use these experimental demand estimates to calculate the contribution of microgrids to household surplus. The value of a new good is the consumer surplus it creates. Let $P(Q)$ be inverse demand, $Q_{vt}^* =$

[21] We have administrative data from Husk Power that contains the monthly payment history of all eligible households. Appendix Table C9 repeats the demand analysis from Table 3 with these administrative data at baseline and endline, as well as for a separate measure of whether a household ever paid for a Husk solar microgrid. At the endline, we observe that about 18 pp (standard error 5.2 pp) of subsidized treatment households and 1.3 pp (standard error 1.0 pp) of normal treatment households are recorded as customers for solar microgrids. We believe the demand estimates from the administrative data are slightly smaller than in the survey, in the normal price treatment arm, because there was a lag between the time when households stopped paying, and hence removed from the administrative records as a customer, and when they were physically disconnected. The baseline results in the administrative data are also similar to the survey baseline results. We do not have access to the administrative data at the time of the follow-up.

$P^{-1}(Price_{vt})$ be the quantity purchased and $\underline{Q}$ be the small quantity that would be purchased at the choke price (= 0 for linear demand, and taken as $P^{-1}(\text{INR } 500)$ for isoelastic demand). We calculate annual consumer surplus at a monthly price $Price_{vt}$ as

$$CS = 12 \int_{\underline{Q}}^{Q^*_{vt}} (P(Q) - Price_{vt}) dQ. \quad (4)$$

With $Q$ measured in market shares, this yields surplus per household over all households, regardless of whether or not they purchased microgrid services.

Figure 4 reports estimates of the value of microgrids. Each group of bars uses a different demand specification. Groups 1 and 2 report estimates of surplus using the Table 3, column 4 (linear) and column 5 (log-log) demand specifications, respectively. (We will discuss groups 3 and 4 in Section 4, when we compare the results from these village-level demand specifications with those from our full demand model.) Within each group, the two bars show surplus at either the endline microgrid price (1/3 of villages at INR 170, 1/3 at INR 100, and 1/3 not offered) or a uniform, subsidized price (INR 100).

Microgrids are a new means of electricity access, but their limited market shares and our elastic demand estimates imply that they generate only modest gains in surplus. At the subsidized price, microgrids increase surplus by INR 222 or INR 242 per household per year (groups 1 and 2, second bar), depending on the demand specification used. At the actual prices and availability (groups 1 and 2, first bar), as of the endline survey, microgrids give surplus of INR 91 or INR 129 per household per year. The surplus of INR 91, calculated from the linear demand curve estimates, is 1.6% of household energy expenditure in our sample. Because roughly one in ten households purchased microgrids, surplus per microgrid user is higher by about a factor of ten.[22] The surplus estimates are similar for the two different specifications of demand.

The demand for one source of electricity will be a bad proxy for the demand for electricity, on the whole, if there are close substitutes available for any given source. We have argued that many sources are close substitutes in Bihar's electricity market (Section 2). The availability of substitutes affects both the interpretation and the external validity of our estimates. On interpretation, internally-valid estimates of microgrid demand cannot tell us household willingness to pay for the product category *electricity*, even within the context of the experiment, when close substitutes are available. Households may value electricity, but have elastic demand for microgrids, if, when microgrid prices rise, they can buy another source of

[22] The surplus numbers for the hypothetical removal of microgrids are also understated in the sense that they give the effect of removal relative to the status quo at the time of the endline survey. In this status quo, microgrids are not present in the control group, one-third of sample villages, to begin with. Thus the removal of microgrids, by design, has no effect on surplus in those villages.

electricity they prefer. On external validity, household demand for microgrids may have been drastically different in a different policy or supply environment, for example, if the government had not made a big push for the grid, or if the price of alternatives like own solar had not declined. In the following sections, therefore, we will specify and estimate a demand model that covers *all* electricity sources.

# 4 Model of Demand for All Electricity Sources

We model consumer demand over electricity sources using a discrete choice, random coefficients demand model (Berry, Levinsohn and Pakes, 1995, 2004). Several aspects of our empirical setting allow for an especially rich specification of the model and credible estimation of its key parameters. First, our data is a household panel survey, so we specify demand to depend on a rich set of observable characteristics at the household level. Second, we allow the unobserved quality of all electricity sources to vary without restriction across villages and time. Third, we use the experimental variation in microgrid prices across markets and time to generate instruments.

## 4.1 Specification

**Indirect utility.—**Household $i$ in village $v$ at survey wave $t$ can choose one out of $j = 1,\ldots,J$ sources of electricity. We do not allow households to choose bundles of sources as we see very little bundling in our data, perhaps because households are too poor.[23] The outside option of no electricity is indexed $j = 0$ and normalized to have indirect utility zero. The indirect utility for each inside source $j$ is given by

$$u_{ijvt} \quad = \quad \delta_{jvt} + \mu_{ijt} + \varepsilon_{ijt}. \tag{5}$$

The term $\delta_{jvt}$ represents the mean utility of source $j$ at survey wave $t$ in village $v$

$$\delta_{jvt} = x'_{jvt}\overline{\beta}^{o} + p_{jvt}\overline{\beta}^{p} + \xi_{jvt}. \tag{6}$$

The vector $x_{jvt}$ of source characteristics includes hours of supply on-peak (from five to ten pm) and hours of supply off-peak and $p_{jvt}$ is the price of the source. Unobserved source quality $\xi_{jvt}$ is known to households but not the econometrician. It may include both unmeasured physical characteristics, such as the capacity of

[23] At the time of our endline survey, only 1.4% of households held multiple sources (Appendix Table A1). For these few cases, we set a priority order in which households are assumed to have chosen the grid if it is part of their chosen bundle. In other settings, for example in cities, households may choose to bundle by having diesel generators or solar power to provide power during grid outages.

a solar system battery, as well as characteristics of the service, such as hassle costs to obtain a connection.

The term $\mu_{ijt}$ gives the household-specific part of utility from each source

$$\mu_{ijt} = d'_{jvt}\beta^o z_{it} + p_{jvt}\beta^p \nu_i. \tag{7}$$

The first term is the effect of household characteristics on preferences for each source. The vector of dummies $d_{jvt}$ $(J \times 1)$ indicates the source $j$ and the vector $z_{it}$ $(R \times 1)$ of household characteristics includes: whether the household has a solid roof, the number of adults, whether the household owns agricultural land, the education level of the household head and household income. The coefficient matrix $\beta^o$ (J $\times$ R) gives the effect of each household characteristic on preferences for each source. This specification allows, for example, richer households to have a greater preference for grid electricity. The $\nu_i$ is a random preference shock for the household's disutility of price, assumed to be constant within a household over time.

We do not allow the utility of a source to depend on what source a household bought in the past. The model allows for persistence in choices through household observable characteristics and the persistence of $\nu_i$, but the adoption of a source does not change preferences, impose any switching cost or yield any residual asset value. There are two main reasons why this specification is appropriate in our context. First, three of the four sources we study are paid for on a monthly basis, own solar being the exception, and so households do not have any asset value from holding these sources. Second, empirically, it does not appear that households are tied to sources they used in the past (Figure 1). We see total disadoption of diesel, and adoption and then disadoption of microgrids, within our study period, and massive changes in shares from one survey wave to the next. These fluid aggregate movements suggest that households do not show a stickiness in their connection to a particular source.

**Choice probabilities.—**Gather the model parameters as $\beta = (\overline{\beta}^o, \overline{\beta}^p, \beta^o, \beta^p)$ and $\delta_{vt} = (\delta_{1vt}, \ldots, \delta_{Jvt})$ and the characteristics of goods $x_{vt} = (x_{1vt}, \ldots, x_{Jvt}, p_{1vt}, \ldots, p_{Jvt})$. The probability that a household chooses product $j'$ is given by

$$\Pr\left(j' \mid z_{it}, x_{vt}, \beta, \delta_{vt}\right) = \int_{\nu_i} \Pr\left(j' \mid z_{it}, x_{vt}, \beta, \delta_{vt}, \nu_i\right) f\left(\nu_i\right) d\nu_i \tag{8}$$

$$= \int_{\nu_i} \frac{\exp\left(\delta_{j'vt} + \mu_{ij't}\right)}{1 + \sum_{j=1}^{J} \exp\left(\delta_{jvt} + \mu_{ijt}\right)} f\left(\nu_i\right) d\nu_i. \tag{9}$$

Here we make two distributional assumptions. First, in writing (9), we assume, as is common in the discrete choice literature, that the idiosyncratic utility shock $\varepsilon_{ijt}$ is independently and identically distributed across households and time periods with an extreme value type-I (Gumbel) distribution. Second, we will let $f(\nu_i)$ be a triangular distribution with support on $[-1,1]$. We assume the distribution is triangular, rather than normal, because a normal distribution with infinite support would imply that some households have a positive taste for price for any estimates of $(\overline{\beta}^p, \beta^p)$. The parameter $\beta^p$ scales the difference between the mean coefficient on price $\overline{\beta}^p$ and the coefficient on price for a maximally price sensitive household (7).

**Consumer surplus.—**Given the coefficients in the model we can calculate the consumer surplus for any choice set, consisting of a set of goods, their characterisitcs and the characteristics of households. Let $\widehat{V}_{ijvt} = \hat{\delta}_{jvt} + \hat{\mu}_{ijt}$ be the estimated strict utility for any good, which will depend on good and household characteristics via (6) and (7). The consumer has price coefficient $\beta_i^p = \overline{\beta}^p + \beta^p \nu_i$. Conditional on the choice set and a draw of $\nu_i$, expected surplus is

$$E[CS_{it}|\nu_i, J] = \frac{1}{\beta_i^p} \ln \sum_{j \in J} \exp(\widehat{V}_{ijvt}), \tag{10}$$

as in the logit model. We integrate over $\nu_i$ to get $E[CS_{i,t}|J]$. We use Gauss-Hermite quadrature to approximate

$$\mathbb{E}[CS_{it}|J] = E_\nu[E[CS_{i,t}|J]] = \int_{\nu_i} E[CS_{i,t}|\nu_i, J] f(\nu_i)\, d\nu_i \tag{11}$$

The willingness-to-pay between choice sets, as well as for differences in household or source characteristics, can be calculated as the expected difference in consumer surplus generated by these choice sets.

## 4.2 Estimation

We estimate the model in two steps. In the first step we estimate the mean indirect utilities $\delta_{jvt}$ and the parameters $\widetilde{\beta} = (\beta^o, \beta^p)$ that enter household-specific choices. In the second step, we use the $\delta_{jvt}$ to estimate the average effects $\overline{\beta} = (\overline{\beta}^o, \overline{\beta}^p)$ of product characteristics and recover the quality terms $\xi_{jvt}$.

This two-step estimation follows a typical micro-BLP approach (Berry, Levinsohn and Pakes, 2004). However, the richness of our data and experiment allow some innovations relative to a typical application of mixed logit. First, with household panel data for many villages $v$, we recover $V \times T \times J \approx 1200$ unobservable village-time-product mean characteristics. Second, households are observed repeatedly in the panel, which we use to generate moments to help estimate the random coefficients. Third, the experiment provides exogenous variation in price to estimate the average coefficient on price via two-stage least squares.

**Non-linear estimation of the first step.—**We estimate the first step by Generalized Method of Moments. Here we briefly describe the three sets of moments used in estimation. Appendix B gives the formal definition of the moments and the GMM objective function.

The first set of moments are product-survey-village market shares. We solve for the $\hat{\delta}_{jvt}(\widetilde{\beta})$ that match observed market shares, given any candidate household preference parameters $\widetilde{\beta}$. This step ensures the model fits observed market shares exactly.[24] Concentrating out the $\delta$ vector in this way greatly reduces the dimensionality of the non-linear search. The second and third set of moments are used to estimate $\hat{\tilde{\beta}}$. The second set of moments are the covariances between household characteristics and the characteristics of the electricity source a household chose. We use all interactions of $R$ household characteristics with dummy variables indicating each of the $J = 4$ inside products to match the characteristics of households that chose each source in each wave. Third and finally, we use as moments the transition probabilities between household choices of different sources of electricity across waves of our panel survey. For example, a moment would be the probability a household moved from off-grid electricity at endline to grid electricity at the follow-up survey. We expect these transitions over time help identify the variance of household taste shocks. A household with a strong, persistent dislike of price would be expected to transition from one low-priced product to another as prices change.

**Linear estimation of the second step with experimental instruments.—**The second, linear step is to estimate equation (6) to recover the mean effects of product characteristics on utility. Regressing the $\widehat{\delta}_{jvt}$, from the first step, directly on product characteristics is likely to yield biased estimates of $\overline{\beta}$ because unobserved source quality $\xi_{jvt}$ may be endogenous to product characteristics including, in particular, price. For example, diesel operators may charge more in villages where they offer higher loads.

[24] We use a Laplace correction to adjust market shares if a source is available but not purchased by any household in our survey sample. This correction is needed because the model will always predict a strictly positive, though small, share for a given source, while exact zero shares are observed in finite samples. For a sample of size $n$, this correction replaces observed market shares $s_j$ with $\tilde{s}_j = (ns_j + 1)/(n + J + 1)$, which has the effect of giving small, positive shares to any source with a precise zero share, while slightly deflating the shares of other sources. Since we observe availability on the supply side for the grid, microgrid and diesel, separately from whether any household in our sample used a given source, we do not apply this correction if a source was not available in a village. Instead, we remove that choice from the choice set for that village.

We therefore specify a system to be estimated via two-stage least squares

$$\widehat{\delta}_{jvt} = x'_{jvt}\overline{\beta}^{o} + p_{jvt}\overline{\beta}^{p} + \overline{\xi}_{jt} + \tilde{\xi}_{jvt} \tag{12}$$

$$\begin{aligned} p_{jvt} &= \pi_1 T_v^{Normal}\mathbf{1}\{Endline\} + \pi_2 T_v^{Subsidized}\mathbf{1}\{Endline\} + \\ &\quad \pi_3\widehat{Peak}_{vt} + \pi_4\widehat{OPeak}_{vt} + \alpha_{jt} + \eta_{jvt}. \end{aligned} \tag{13}$$

Equation (12) gives mean indirect utility, where we split $\xi_{jvt} = \overline{\xi}_{jt} + \tilde{\xi}_{jvt}$ into the sum of a source average quality $\overline{\xi}_{jt}$ in each survey wave and the deviation $\tilde{\xi}_{jvt}$ of source quality in a village from that average.

Equation (13) gives the first-stage specification for price. The first stage constructs instruments from the experiment. Our solar microgrid experiment offers instruments that are excludeable and likely to be powerful, given that the microgrid treatment changed market shares (Table 3). Equation (13) uses interactions of the village-level treatment indicators $T_v^{Normal}$ and $T_v^{Subsidized}$ and an indicator $\mathbf{1}\{Endline\}$ for the endline survey, when the experiment was ongoing, as instruments for price. The $\alpha_{jt}$ are source-by-wave fixed effects.

The hours of supply on the grid may also be endogenous to product quality. We have worked with the distribution company and it does not knowingly set supply in response to the characteristics of competing sources. However, to account for this possibility, in our preferred specification we also instrument the supply hours in a village (both on- and off-peak) using predicted supply hours $\widehat{Peak}_{vt}$ and $\widehat{OPeak}_{vt}$, where the predictions are made using supply hours in nearby villages. We expect that villages nearby in the electricity grid, for example that are served by the same substation, will be similarly affected by the distribution companies' power supply rationing rules. The exclusion restriction is that supply of electricity in nearby villages is not correlated with the determinants of demand in a given village, after conditioning on our rich set of household observables. Appendix A details the construction of the instrument.

As a basis for comparison, we will also report results using ordinary least squares and using traditional price instruments from the industrial organization literature. We have two sets of alternate instruments for source-village-wave prices. First, the average hours of supply and load from the other products in the same village, which will affect source mark-ups and prices under oligopolistic competition (Berry, Levinsohn and Pakes, 1995). Second, the average price for a given source in the nearest three villages where that source is available, which will covary with source price due to common supply shocks (Hausman, 1996; Nevo, 2001).

### 4.3 Results

This subsection reports estimates of the electricity source demand model. The full demand model has 723 parameters: 699 source-by-village-by-survey wave mean indirect utility parameters[25], backed out from the first-stage demand model; 20 parameters governing household heterogeneity; 3 parameters on the average effects of source characteristics; a parameter for the dispersion of the random coefficient on price. We therefore report only select parameters, to give a sense of how the model represents household choices. First, we report the linear estimates of the average effects of source characteristics, from the second step. Second, we present estimates from the non-linear first step of how household characteristics affect choice probabilities.

**Second stage estimates: Mean effect of source characteristics.—**We begin with estimates of the first stage of the linear part (second step) of the broader model. Our preferred specification (13) instruments for both price and supply hours. We find that the experimental treatment assignments have a highly statistically significant effect on source prices (Appendix Table C2, column 2a). The first-stage $F$-statistic for a test of the null that the instruments do not affect price ranges from 21 to 42, depending on whether we instrument for price and hours simultaneously (column 2a), or only for price (column 1). In addition, our supply instruments strongly predict hours of supply both during peak and off-peak hours (columns 2b and 2c). Alternative instrument sets that do not use the experiment lack the power to predict price in the first stage.[26]

Table 4 reports estimates of the linear part of the demand model (12). Column 1 reports results from ordinary least squares estimates. Columns 2 and 3 report instrumental variables estimates using the first stage from the experiment, instrumenting either for price or for both price and hours. Columns 4 and 5 replace the experimental variables in the instrument set with the alternative instruments for price.

The experimental instrumental variables estimates show a high degree of price sensitivity. We find a coefficient of $-5.05$ (standard error 1.96) on price (column 2). This estimate is essentially unchanged if we additionally instrument for hours of supply (column 3). The magnitude of the coefficient on price is two-and-a-half times greater than in the OLS estimates (column 1), which suggests that OLS is biased upward, towards zero, from some combination of endogeneity and measurement error.

[25] Although there are four products and hundred villages across three time periods, resulting in 1200 possible values for mean indirect utility parameters, product availability varied across villages and survey waves.

[26] Neither the BLP ($F$-statistic 0.4) nor Hausman ($F$-statistic 1.0) instruments have much predictive power for the endogenous price variable (Appendix Table C2, columns 3 and 4, respectively). One interpretation of this result is that the assumption of oligopolistic conduct that underlies the BLP instruments is not appropriate in this setting. Power sources like own solar are perfectly competitively supplied and the government's objective, in pricing grid electricity, is clearly not to maximize profits.

Estimates of the price coefficient using alternative instruments drawn from the literature are imprecise. The point estimates with both the BLP and Hausman instruments are large and positive (columns 4 and 5). We cannot reject the equality of either of these estimates with any of the experimental estimates, the OLS estimates, or a zero coefficient on price. The experimental variation is necessary to recover unbiased and precise estimates of the price coefficient in our setting.

We also estimate the effect of supply hours on household mean utility. We find a positive and statistically significant effect of peak hours of supply on mean utility (Table 4, column 3). We also find a smaller and negative coefficient for off-peak hours. Our estimate for the value of peak hours is not precise, but agrees with the idea that agricultural households, who may be away during the day, mainly value power in the evening hours. For this reason, private diesel generators typically operate only in the evening. We proceed with the column 3 estimates, instrumenting for both price and hours, as our main specification for counterfactual analysis.

**First stage estimates: Heterogeneity in demand across households.—**Table 5 reports the effects of household characteristics on choice probabilities from the demand model. We estimate two instances of the model. First, to provide a simple univariate proxy for wealth, we estimate a model that includes as covariates only the number of adults in the household and a dummy variable for whether the household has a solid roof (columns 1 through 5). Second, we estimate the full model, which includes three additional observable proxies for household demand: whether the household owns agricultural land, the education of the household head and household income.

The effects of household characteristics are non-linear. The table therefore reports marginal effects evaluated for a household that has the median value of each household characteristic (see Appendix Table C4 for summary statistics on the characteristics that enter demand). The marginal effects are not strictly marginal; for binary variables we report the effect on each given choice probability of changing the value from zero to one and for continuous variables the effect of a one standard deviation increase.

The main finding of the table is that richer households have much stronger preferences for grid electricity than all other sources. Consider the simple model specification (columns 1 to 5). The baseline probability of grid choice is 24 percent. On top of this base, a household with a solid roof is 23 pp (standard error 1.7 pp) more likely to choose grid electricity (column 1). Nearly all of this effect comes from a reduction in the choice of the outside option of no electricity (column 5).

In the full model we add additional covariates (columns 6 to 10). A Wald test, reported at the bottom of Table 5, easily rejects the simple demand model in favor of the full model with additional covariates ($p$-value $< 0.001$). The effect of having a solid roof on grid choice declines only slightly, since the solid roof dummy is correlated with other measures of wealth, but remains large (17.4 pp). Increases in the number of household adults, the ownership of agricultural land and the education of the household head are all associated with significantly higher demand for the grid. For example, a household that owns agricultural land is 9.2 pp (standard error 3.1 pp) more likely to choose the grid. These demand proxies have much smaller effects on the choice probabilities for other sources, though some do significantly affect demand; for example, higher-income households appear slightly more likely to choose microgrids (column 9). A natural interpretation of this heterogeneity is that grid electricity supports higher loads, so more households on the grid can run a fan or television (Table 1).[27] Richer households want the energy services that these devices bring.

**Implied price elasticities.—**We calculate the price elasticities implied by our demand estimates. The aggregate own- and cross-price elasticities by source are shown in Table C3. We find that demand for the grid is much less own-price elastic than demand for other electricity sources. In panel A, elasticities are evaluated at the prices observed in our endline survey. We find the demand elasticity for grid electricity is $-0.20$, whereas the own-price elasticities for off-grid technologies are all near $-2$. A plausible reason for this difference is that the grid is heavily subsidized so that most households who have access to the grid in their village will choose the grid, if it is decently reliable. In panel B we evaluate the same elasticities at break-even prices for the grid. The own-price elasticity of grid demand increases in magnitude by a factor of roughly four, to $-0.82$. The cross-price elasticities of off-grid source market shares with respect to the grid price triple. As the grid price approaches the cost of supply households more readily substitute to off-grid sources.

The pattern of elasticities, at both actual and break-even prices, suggests that the public distribution company follows a very different pricing rule than a profit-making firm would. A monopolist maximizes profit by pricing on the elastic part of the demand curve. Here, the natural monopolist distribution company sets much lower prices, on the inelastic part. As off-grid sources can be thought of as competitively supplied there is no clear prediction as to their own-price elasticities when evaluated at market prices. We

[27] In our baseline survey, very few households reporting using microgrids. As a result, the appliance ownership summary statistics conditional on using microgrids at baseline are imprecisely estimated.

find households to be quite elastic with respect to the prices of off-grid sources, consistent with the rapid expansion of solar power as its costs have fallen in recent years (Figure 1)

# 5 Competition between Sources and the Value of Electrification

The model estimates now allow us to measure the surplus households gain from electricity and to study how that surplus depends on the competition between different electricity sources. We do this in three steps. First, we use the model to compare the surplus from electrification to the surplus from microgrids alone. Second, we use the model to value the two disruptions that Bihar went through during our study period, the advent of off-grid solar and a big push for grid supply. Third, we study counterfactual policies that project recent shifts in supply and demand forward, to understand the medium-run future of electrification.

## 5.1 The value of microgrids versus the value of all electricity sources

We start by returning to the estimates of the value of microgrids in Figure 4. With the structural model, we can calculate the surplus from any electricity source, by raising the price of that source to a high level and calculating the decline in total surplus (equation 11).

The value of microgrids in the structural model is nearly the same as calculated previously with the linear model of demand. Figure 4 shows four pairs of bars. In the first three pairs, we report the surplus from microgrids. Within each pair, the left bar reports surplus at endline prices and the right bar at the subsidized price of INR 100. We find that the structural model yields a surplus of INR 204 per year, as compared to a value of INR 222 under a linear demand model and INR 242 under a constant elasticity model. The consistency of valuations across the different models is reassuring since demand is estimated with the same experimental variation in each model.

The second finding is that the surplus from microgrids is a relatively small part of the total surplus from electrification. With the full structural model, we can calculate the surplus from *all* sources of electricity together. We find that the total household surplus from electrification, at endline and with subsidized micro-grids, is INR 1128 per year, greater than the surplus from microgrids alone by a factor of roughly five (right group of bars).

This large difference shows that the modest value of microgrids found in Section 3 does not reflect a low valuation of electricity but rather the availability of other sources of electricity that are similar in appeal to microgrids but are available at lower prices. Studying the demand for any single electricity source, without considering other sources of power, therefore significantly understates household willingness to pay for electricity. Having a framework and data that enables one to estimate demand for all sources is critical in a world where the electricity choice set is expanding quickly.

### 5.2 The two disruptions in Bihar's electricity market

Here we apply our model to measure the contribution of the two disruptions in Bihar's electricity market, the advent of off-grid solar power and the big grid, to the household surplus from electrification.

Table 7 reports a range of counterfactual results based on the demand model estimates. Every row represents one counterfactual model run. Columns 1 through 4 report source market shares and column 5 the electrification rate, the sum of the market shares of all inside sources. Columns 6 to 8 report consumer, producer and total surplus. All surplus measures are per household per year across the entire population, including households who choose the outside option of no electricity. The producer surplus is the surplus from the grid only, which is approximately equal to total producer surplus in the market.[28] Producer surplus is typically negative, because the state distribution company loses money on every grid customer. In some scenarios we eliminate grid energy subsidies by setting prices to breakeven with variable costs (see footnote 16).

There are three main conclusions we draw from Table 7.

First, the household surplus from electrification increased by a factor of roughly five over a short 3.5 year period. At baseline, the grid had low penetration and solar was expensive and unpopular (Table 7, panel A, row 1). Households valued all electricity sources at only INR 380 per *year*. By the follow-up survey this had risen to INR 1,866 (panel A, row 3, column 6). To put this gain in context, at baseline, household expenditure on electricity and lighting in our sample was INR 2,029 per year and on all energy INR 6,024 per year. The increase in surplus from electrification is therefore 73% of baseline electricity and lighting expenditures and 25% of all energy expenditures. Part of the gain in household surplus is due to increased producer losses (panel A, row 3, column 7). The state loses money as more households choose to connect to the extended, and heavily subsidized, grid.

Second, gains in household surplus come both from the grid and solar and are driven by dramatic im-

[28] Producer surplus for the grid is a measure of variable profits: the profits or losses that accrue to the state from supplying grid electricity, after accounting for the cost of energy supplied. Losses must be covered by tax collection from Bihar and from other states, due to central government transfers. Producer surplus for the grid can be taken as capturing producer surplus from the whole market, if we assume that the other sources are competitively supplied. The assumption of zero profits is probably accurate for own solar but not for diesel, which, in any case, has a small market share at endline.

provements in price, quality, and availability over our 3.5 year study period. Household surplus gains from the grid, however, tend to be larger.

To illustrate this we consider two hypothetical pathways to electrification that improve the characteristics of only one of these technologies, in turn, holding the other constant at baseline values. In panel B, we measure the surplus due to the advent of solar power holding constant the extent and quality of the grid. Unlike grid electricity, which has been available for decades, solar is a new electricity product with costs rapidly falling and quality improving. Hence, we model the advent of solar as changing the price and quality of both own solar systems and micro-grids from initially not being available to being available at their estimated values in the follow-up survey. We find that the arrival of improved solar power alone would have increased the share of households with any source of electricity from 24 pp to 50 pp (panel B, row 2 versus row 1, column 5) and the surplus from electrification by 1.6× (panel B, row 2 versus row 1, column 6). Therefore even without state investment in the grid, off-grid solar has the potential to significantly increase electrification rates. This is an important finding for places beyond the electrification frontier where the scope for government investment in the grid is limited.

In panel C, we instead measure the surplus due to improvements in grid extent and quality holding the extent and quality of off-grid solar constant. We find that the grid improvements on their own would have increased electrification from 31 to 51 pp and surplus by 4.2× (panel C, row 2 versus panel A, row 1). This is a large part of the total increase in the value of electrification from all sources. So though increases in electrification rates are similar to those from improved solar the increase in the household surplus from the improved grid is larger. About half of this is purely a transfer from the government via subsidies for energy (panel C, row 2, column 7). However, if we eliminate subsidies by pricing grid electricity at cost and compare the improved grid (Panel C, Row 3) with improved solar (panel B, row 3), the household surplus is about twice as large when grid is improved than when solar is improved. This is even though electrification is actually higher under improved solar (49% vs 40%). This points to a higher valuation of grid electricity.[29]

Third, the value of solar is contingent on the state of the grid, since the two sources are substitutes. We find that the extent to which the state invests in the grid affects how much households value off-grid solar. To illustrate this Panel D values simultaneous improvements in both technologies. We estimate that improvements to solar increase consumer surplus by INR 182 per household per year when grid availability and quality are held at the baseline levels (compare panel B, row 2, column 6 to panel A, row 1, column

[29]This higher valuation may emanate, in part, from grid electricity being to support higher load appliances.

6). In contrast, when grid availability and quality are improved, improvements to solar increase consumer surplus by INR 111 per household per year (compare panel D, row 1, column 6 to panel B, row 2, column 6) - only 60% as large when the grid is left in its baseline state. The progress of solar is therefore less valuable to households when the government is making large investments in the grid.

This quantification of the value brought by the two great disruptions to the electricity supply in Bihar makes clear why a model is needed. Without the model we could not estimate: 1) the total value of electrification to households, 2) how much different electricity sources contribute to this, and 3) how policies enhancing one source of electricity influence how much households value other sources. The demand model allows us to step back and measure the combined effects of improvements in both the grid and off-grid power sources.

### 5.3 Counterfactual policy reforms and the future of electrification

**Scenarios.—**Our model also allows us to move beyond our study period to project how demand for different electricity sources will evolve in Bihar as market conditions, government policies and household characteristics change.

*Falling solar prices.* One of the key disruptions in the global electricity markets has been the dramatic fall in solar energy prices. Solar panel and battery prices are projected to continue to fall (Feldman, Margolis and Denholm, 2016; Howell et al., 2016). To capture this we assume a 50% reduction in solar costs and allow this decline to pass through completely to the prices of both stand-alone solar systems and microgrids.[30]

*Improving grid.* As of the follow-up survey, the grid is still only present in 72% of villages and supplies on average 14 hours of power a day, with only about 3 hours during the 5-hour evening peak. The government continued to invest in grid extension after our surveys and has increased supply to rural areas. To capture improvements in the grid we construct a counterfactual where the grid is extended to all villages and peak supply hours have been extended by two hours a day up to a maximum of five hours, the full duration of evening peak demand.

*Growing incomes.* Bihar is a relatively poor state but among the fastest growing in India, with an average annual growth rate in state product of 11% from 2012 to 2018. To model demand growth, we weakly raise

[30] This assumption is fairly aggressive. Cost reductions for solar PV are projected at 55% (Feldman, Margolis and Denholm, 2016). For batteries, cost reductions are projected at 75% (Howell et al., 2016). Since the panel and batteries only make up a part of the system, however, these reductions are larger than the implied total reduction in system costs (See Appendix Figure D1 for a breakdown of costs). We assume that this decline in cost passes through completely to microgrid prices, thereby lowering prices from a market price of INR 170 to INR 85.

all households to the maximum of their current observables and the median profile (Appendix Table C4). These are large relative changes but small in absolute terms; the 80th percentile of the income distribution in our sample reaches parity only with the per capita GDP of Malawi, one of the world's poorest countries.[31]

**Results.—**Table 7, panel E reports on the results of the forward-looking counterfactuals that carry out the changes described above. The panel begins from the state of the market in our follow-up survey, in row 1, and cumulatively adds improvements to the supply and demand sides of the market.

We begin by simulating a "big push" for electrification on the supply side of the market: improving solar (row 2), extending the grid to all villages (row 3), and increasing grid supply (row 4). Improving solar increases household surplus by 10% of the follow-up value (panel E, column 6, row 2 over row 1) and electrification via solar rises from 22% to 42%. Improvements in solar will therefore be an important force for increasing electrification.

However, if the government is willing to invest in improving the extent and quality of grid then grid electricity becomes the dominant choice of households. In row 3 of Panel E we see that if, in addition to improving solar, we extend the grid to all villages then household surplus increases by an additional 15%. If we also also improve grid availability by extending hours of peak supply then we get an additional 84% increase in surplus and the fraction of household electrified by solar falls back to 19% (row 4). Now the grid dominates supplying electricity to 70% of customers.

Future improvements in the grid is therefore roughly 10$\times$ as valuable to households even when there has been large price cuts in off-grid solar. Although the grid improved greatly during the span of our data, the further gains in this counterfactual "big push" show that an incomplete and low-quality grid remains a major hindrance to electricity access at our follow-up survey.

Income growth will tilt electrification even further toward the grid. In row 5, we boost households to at least the median household profile on their observable characteristics such as income. Taking the improved supply side as the benchmark (panel E, row 4), household surplus improves a further 50% (INR 1961, row 5 less row 4). This growth in household incomes pushes off-grid market shares down from 20 pp to 6 pp, increases the grid market share by 22 pp, and achieves near-universal electrification, with 98% of households

[31]The median reported household per capita income in our sample is INR 12000 per year (USD 656 at 2011 PPP) and the 80th percentile is INR 14250 per year (USD 779 at 2011 PPP). At purchasing power parity rates, the 80th percentile in our sample is therefore about in line with per capita income in Malawi (USD 1143 at 2011 PPP) (World Bank). Income measurement is difficult for rural, agricultural households with multiple sources of income, and this comparison should only be taken as roughly indicative of the level of economic development in our sample.

choosing some electricity source.

Households abandon off-grid sources as the grid improves. Comparing to the follow-up survey as a benchmark, the "big push" and income growth together increase the overall electrification rate by 33 pp (panel E, column 5, row 5 less row 1) but the grid electrification rate by 52 pp (panel E, column 1). This projected dominance of the grid, in response to income growth, was foreshadowed by our demand model estimates, which showed that moderately richer households have much stronger preferences for grid electricity.

Through these scenarios, we assume that the government continues to subsidize electricity as it does today. The "big push" on the supply side and growth in demand, therefore, increase producer losses substantially, to INR $-1908$ (row 5, column 7). If the state instead increased grid prices to break even, the household gains from increases in supply and demand would be 55% of the gain with subsidies (panel E, column 6, row 6 versus row 5 as compared to row 1). In this case grid electrification falls by 15 pp and 16 pp of households remain on off-grid sources, for affordability, even as the quality of the grid improves and income grows.

Solar therefore remains an important fall-back for poorer populations who are unable to afford the grid when priced at cost. Indeed, its appearance as a choice may provide more room for governments to increase grid prices in the future. However, even when government raise prices to break even an improved grid remains the dominant source of electricity supplying 77% of customers (panel E, row 6, column 1). This points to its ability to support higher load appliances leading it to be preferred by customers as incomes grow.

**Discussion.—** We find that in Bihar: (i) household surplus from electrification increased five-fold; (ii) both solar and the grid boost electrification but households gain more surplus from the grid; (iii) grid investments and subsidies strongly reduce demand for off-grid solar.

Off-grid solar has emerged as a powerful force for driving up electrification particular where the grid is not available or of low quality. Given that close to a billion remain without electricity this is an important finding. Where governments cannot afford the large grid investments required to reach the whole population including the poor then off-grid solar represents an important means of access to basic levels of electricity provision.

There is however little suggestion that off-grid solar will leapfrog the grid. The evidence we have mustered indicates that households (and particularly richer households) value the grid electricity considerably more than off-grid solar electricity in part because it can power higher load appliances. And this preference becomes ever more stark as the grid extends, as peak hours of supply increase and customers become wealthier.

A central finding from our counterfactuals is that the choices households make are contingent on energy policy. The value of different energy sources may be different in places where the state is not willing or able to subsidize grid investment, connection charges, and energy prices to the same extent. Prior work in Kenya, for example, has found that many rural households do *not* connect to the grid when it arrives and have low willingness-to-pay for grid connections, as measured via experimental variation in connection subsidies (Lee et al., 2014; Lee, Miguel and Wolfram, 2020*b*).

We therefore want to step out of Bihar and use our model to value electrification for rural populations in different African countries which are at different points on the global electrification frontier. The expansion of choice we observe in Bihar is happening across the developing world as innovation makes off-grid cheaper and governments sink money into extending the grid. By estimating our model in data sets from different countries we can examine how household valuation of off-grid solar and grid electricity lines up with what we observe in Bihar.

Observing how households make these choices in a much wider set of contexts will tell us whether and how what we observe in Bihar generalizes. In particular, it can help us understand whether the different choices that households are making across the world can be rationalized as responses to differences in energy policy. As our model and results for Bihar emphasize the valuation of new sources such as off-grid solar may depend fundamentally on how much the state is willing to subsidise the extent and the quality of the grid.

# 6 The value of electricity sources along the global frontier

The present section uses our demand model estimates to value electrification outside of our sample. It is striking that off-grid solar has become dominant in parts of Africa, reaching more than half of households in some countries, while in Bihar we find it is less valuable than the grid and across India take-up remains muted (Figure 1). We wish to use our model to understand this contrast.

## 6.1 Data and methodological approach

Our approach is to gather data on household choices and the supply side of the electricity sector in Africa and then apply our demand model to calculate how households value different energy sources. This part briefly explains the data we use and the modifications we make to re-fit our model to the African data.

**Data.—**The main source of data we use are Living Standards Measurements Surveys (LSMS) collected by the World Bank. LSMS surveys are not specific to energy but most recent surveys include household energy sources and assets. We screened LSMS surveys to find those that record ownership of solar systems among electricity sources. We use the most recent survey waves from Niger 2014, Malawi 2016, Ethiopia 2018, Uganda 2019, Mali 2017, Tanzania 2019 and Nigeria 2018 as well as the Kenya Continuous Household Survey Programme (2019). These surveys are nationally representative of the rural population in each country. The surveys we use have a total of 48,800 households representing a population of 512 million.

To apply our model, we use the data to assemble characteristics of both households, on the demand side, and of electricity sources, on the supply side. All LSMS surveys cover the household characteristics $z_{it}$ in our demand model, with the exception of household income. BM>NR: This is true, but one caveat - for Kenya (KNBS), which is not part of LSMS, we don't have data on land ownership. Also, we estimate income in Kenya using individual income, and not consumption. We use household consumption to replace income in the model. On the supply side, we define markets using enumeration areas in each survey or aggregations over nearby enumeration areas. We then infer the availability, pricing and reliability of electricity sources available in a given market from the survey responses of households in that market using a given source. Annex 1.3 describes both the market definition and how we set supply characteristics.

**Methodology for refitting model to African data.—**While we seek to apply the model out of sample, it would not be appropriate to apply our model estimates without modification to Africa. The main reason is that the demand model estimates include both household preference parameters $\beta$ and source-village-wave specific quality terms $\xi_{jvt}$. While the model infers the quality terms from household choices, they are not a structural feature of household preferences, but would depend on the supply side in a given market. For example, if the connection charges in a market were high, or it was hard to find solar panels in a given area, the estimated quality of a source would be worse.

We therefore take the preference parameters $\hat{\beta}$ to be as estimated in Bihar and refit the $\xi_{jvt}$ to household choices in each African survey wave. Substantively, this means we hold fixed the demand side, including tastes for price and reliability and the effect of household observable characteristics on tastes. We then use the Berry (1994) contraction mapping

$$\delta'_{jvt} \leftarrow \delta_{jvt} + \log\left(s_{jvt}\right) - \log\left(\Pr\left(j \mid z_{it}, x_{vt}, \hat{\beta}, \delta_{vt}\right)\right) \tag{14}$$

to infer the quality of each source-market-wave in the African surveys. Here $z_{it}$ are the household characteristics, $x_{vt}$ the source characteristics and $s_{jvt}$ the market shares in the African survey. This yields unique $\delta_{jvt}$ such that predicted market shares in the model equal observed shares in the data. We can then solve (6) for the unobserved qualities $\xi_{jvt}$ of each source in each market and survey.

We favor this approach to value the particular supply conditions in each market and to gauge the household surplus from electrification under local conditions. It would not be realistic to assume that the source quality distribution in Africa is the same as in Bihar. Our approach is also helpful in investigating the role of preferences vs policy-determined grid characteristics in explaining the market-shares we observe in Africa. However, re-estimating the quality parameters means that we cannot test model fit against out-of-sample market shares, as we are refitting source quality in each survey and so will fit market shares by construction.

### 6.2 Results

We apply the model to calculate household surplus from electrification (11) by country and source. Figure 5 summarizes the results. Panel A shows surplus at the supply side in the data; panel B lowers the price of the grid in each country to the Bihar median grid price; panel C additionally extends the grid to all villages. Within each panel, the three bars for each country show the household surplus from electrification (hollow bar at left), the surplus if only the grid were present (brown bar at center), and the surplus if only solar were present (orange bar at right). These values are showin in INR against the left axis. The countries are ordered from left to right by per capita income, measured in panel A against the right axis.

There are three main findings from these cross-country counterfactuals. First, our setting in rural Bihar is an outlier as compared to African countries of comparable income. In Figure 5, panel A, we find that the surplus from electrification is roughly increasing in household income from poor countries at left (Niger, Malawi) to richer countries at right (Kenya, Nigeria). However, the surplus from electrification in Bihar stands well above that for the rural population in all other countries considered. It is roughly twice as high as per household surplus in Kenya and Nigeria, despite the latter two countries being richer than Bihar.

Second, solar matters more for the surplus from electrification in Africa than in India. In relatively poor African countries, including Ethiopia, Uganda and Mali, solar generates more surplus than the grid, consistent with the high market shares of solar in those places. In richer countries like Tanzania and Kenya, solar generates about half as much surplus. Only in Nigeria, the richest country in the sample, is the grid as relatively dominant as in Bihar.

Third, and most remarkably, the differences in the value of electrification across these settings are accounted for to a good extent by only two factors: energy subsidies and the extent of the grid. In panel B we cut grid prices to the median price in Bihar. This change is significant because many African countries do not have electricity subsidies like India. The surplus from the grid in Tanzania, Kenya and Nigeria rise to meet or exceed that in Bihar. The surplus in Uganda and Mali increase markedly, but not to Bihar's level. In panel C we additionally make the grid universal in each country. The surplus in Mali now also exceeds that in Bihar. There is a striking uniformity in that in each of Mali, Bihar, Tanzania, Kenya and Nigeria we now see a Bihar-like pattern of higher household surplus (around INR 2000 per household-year), most of which comes from the grid, with a residual role for off-grid solar power.

The results suggest that some of the stark differences in the value of electricity across countries can be rationalized by simple differences in energy policy on the supply side. Solar power's progress has made it a viable option for households in many parts of the world. Solar generates surplus from electrification when the grid is not present or not subsidized. As households grow richer, they prefer the grid for its higher quality (e.g., higher loads), but the point at which they move from solar on to the grid depends on the supply environment. In Bihar, the state has made a "big push" to extend the grid and also to subsidize electricity supply. Under these policies the grid is dominant and will only grow more so. For households that live in a state that cannot afford or does not choose to subsidize energy, solar will remain an importance source of electricity until they reach markedly higher levels of income.

# 7 Conclusion

Electricity markets on the global electrification frontier are undergoing radical changes, driven by innovation in solar power and a traditional "big push" for grid electricity. We model the demand for *all sources* of electricity and apply the model to break down the gains to the poor from a large expansion of energy access.

Our approach uses revealed preference measures of demand, estimated with medium-run experimental variation, to value historic changes in energy supply, which would not have been feasible directly to manipulate with an experiment. As a point of comparison, the increase in electrification rates within our four-year study period is roughly twice as large as the increase in farm electrification rates studied by Kitchens and Fishback (2015), who consider a decade of gains from the beginning of the Rural Electrification Administration in the United States.

There are four main findings from our analysis of the value of electrification on the global frontier. First, electricity is very much a differentiated product with significant heterogeneity in household choices and fast changing market shares and product characteristics. Innovation is rapidly expanding the electricity choices for households on the electrification frontier. By simultaneously measuring demand from different sources we are able to estimate the value of each of these in this fast changing environment.

Second, in our setting of Bihar, household surplus from electrification increased by five times during the study period. Both an expanding grid and new off-grid choices add value for households. This is an important finding because it demonstrates how expanding the electricity choice set for households confers value. Existing technologies will change their characteristics and new energy technologies will come onto the market and so creating frameworks and collecting data that allow us to measure household surplus from each source is important.

Third, we find that the characteristics of different sources has an important bearing on how much each contributes to household surplus. In Bihar, the fact that large public investments have led to significant grid extension and subsidized connections and tariffs, is partly responsible for the grid accounting for the majority of household surplus gains. Prices of different technologies (e.g. grid versus off-grid solar) as well as technological characteristics (e.g. load) will affect consumer choice and hence the value they derive from different electricity sources. These, in turn, will be affected both by public policies and the innovation and diffusion rates of different technologies.

Fourth, we find that the breathtakingly rapid expansion of the electrification frontier driven by the twin forces of grid expansion and take-up of off-grid solar is a phenomena not limited to Bihar. By piecing together data sets across African countries and Indian states we can demonstrate that this is a global phenomena. Analysis of this data through the lens of our demand model shows that, as in Bihar, both the grid and off-grid solar increase household surplus thus underlining how the expansion of choice on electricity sources has increased the value of electrification. We find that, in Africa, household surplus gains from electrification come more from off-grid solar relative to the grid when compared to Bihar. This may be due to the grid being less subsidized and, indeed, when we price grid supply in African countries at Bihar prices household surplus from this source increases considerably.

Our analysis considers household surplus and not household or social welfare. Governments may favor energy subsidies for the poor for two main reasons. First, a preference for redistribution or a view that access to modern energy is a right. The large loss that Bihar is willing to incur to increase electrification rates is one measure of the value that the state places on energy access *per se*. A drawback of redistribution through energy subsidies is that subsidies, ironically, can undercut the quality of energy supply (Burgess et al., 2020; Dzansi et al., 2019). A pullback in supply, in turn, may slow economic growth, for example the growth of manufacturing firms (Allcott, Collard-Wexler and O'Connell, 2016). The second factor that favors subsidies, for the grid in particular, is external returns, for example a big push for electrification might generate spillovers in consumption or increases in productivity for firms, beyond the direct value of electricity to consumers. The experiences of early electrifiers such as South Korea would point in this direction, however, evidence on external returns is uncertain and more research in this area is needed (Lipscomb, Mobarak and Barham, 2013; Kline and Moretti, 2014). Governments in pursuit of such returns may reasonably push for universal electrification, even if it may seem *too early* in the process of development, as measured by household willingness-to-pay alone. Bihar's big push has come at a very low level of income, relative to historical precedent (Lee, Miguel and Wolfram, 2020*a*).

Part of the attraction of renewables derives from how clean they are. They lie at the core of country's aim to reduce emissions and particulate pollution. Though we have modelled the choice between off-grid and grid choices we have been silent on what powers the grid. One way to square the need for economic growth with the need to reduce the externalities from growth is to use more renewables like solar in the generation of grid electricity. This is particularly important if the grid is able to deliver more load to the firms that underpin industrialization and structural change of the economy.

Taking everything together we see that off-grid solar can help to drive electrification in particular when the grid is not present, of low quality or unsubsidised. It is in this sense that this innovation is important for the world. But when the quality and extent of the grid get better then reliance on off-grid seems to fall away. So there not a sense in which it will leap-frog the grid. However, in places where governments are unwilling or unable to pour money into improving the extent and quality of the grid it is likely to remain a key and growing source of electricity as innovation drives solar prices down.

The expansion of the electricity choice set that we study has brought a large increase in household surplus from electrification in a short time. Off-grid solar is not leaping over the grid, as mobile telephony made the landline obsolete before landline networks were ever completed. However, it is a worthy stopgap for the needs of the rural poor and has accelerated growth in energy access. The pace of change we observe in Bihar is unusual, but the same forces driving it are at work in large parts of South Asia and Africa. Soon enough the global electrification frontier may disappear.

## References


**Aklin, Michaël, Patrick Bayer, SP Harish, and Johannes Urpelainen.** 2018. "Economics of household technology adoption in developing countries: Evidence from solar technology adoption in rural India." *Energy Economics*, 72: 35–46.

**Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias.** 2012. "Targeting the poor: evidence from a field experiment in Indonesia." *American Economic Review*, 102(4): 1206–1240.

**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.

**Attanasio, Orazio P, Costas Meghir, and Ana Santiago.** 2012. "Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresa." *The Review of Economic Studies*, 79(1): 37–66.

**Attanasio, Orazio, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina.** 2020. "Estimating the production function for human capital: results from a randomized controlled trial in Colombia." *American Economic Review*, 110(1): 48–85.

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

**Berry, James, Gregory Fischer, and Raymond Guiteras.** 2020. "Eliciting and utilizing willingness-to-pay: evidence from field trials in northern Ghana." *Journal of Political Economy*, 128(4): 1436–1473.

**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.

**Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman.** 2010. "What's advertising content worth? Evidence from a consumer credit marketing field experiment." *The Quarterly Journal of Economics*, 125(1): 263–306.

**Bihar State Government.** 2015. “Bihar Report Card.” Bihar State Government.

**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.

**Bureau of the Census.** 1975. “Historical Statistics of the United States: Colonial Times to 1970.” Bureau of the Census 1.

**Burgess, Robin, Michael Greenstone, Nicholas Ryan, and Anant Sudarshan.** 2020. “The Consequences of Treating Electricity as a Right.” *Journal of Economic Perspectives*, 34(1): 145–169.

**Burlig, Fiona, and Louis Preonas.** 2016. “Out of the Darkness and Into the Light? Development Effects of Rural Electrification.” Mimeo, University of Chicago.

**Business Today.** 2017. “CM Nitish Kumar orders free power connection for all in Bihar.”

**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, Michael Greenstone, Rohini Pande, and Nicholas Ryan.** 2018. “The value of regulatory discretion: estimates from environmental inspections in India.” *Econometrica*, 86(6): 2123–2160.

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

**Dupas, Pascaline.** 2014. “Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment.” *Econometrica*, 82(1): 197–228.

**Dzansi, James, Stephen Puller, Brittany Street, and Belinda Yebuah-Dwamena.** 2019. *The Vicious Cycle of Blackouts and Revenue Collection in Developing Economies: Evidence from Ghana.* Mimeo, Texas A&M University.

**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.

**Fowlie, Meredith, Yashraj Khaitan, Catherine Wolfram, and Derek Wolfson.** 2019. “Solar Microgrids and Remote Energy Access: How Weak Incentives Can Undermine Smart Technology.” *Economics of Energy & Environmental Policy*, 8(1).

**Galiani, Sebastian, Alvin Murphy, and Juan Pantano.** 2015. “Estimating neighborhood choice models: Lessons from a housing assistance experiment.” *American Economic Review*, 105(11): 3385–3415.

**Goolsbee, Austan, and Amil Petrin.** 2004. “The consumer gains from direct broadcast satellites and the competition with cable TV.” *Econometrica*, 72(2): 351–381.

**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.** 2020. “Demand for Off-grid Solar Electricity: Experimental Evidence from Rwanda.” *Journal of the Association of Environmental and Resource Economists*, 7(3).

**Haushofer, Johannes, and Jeremy Shapiro.** 2016. "The short-term impact of unconditional cash transfers to the poor: experimental evidence from Kenya." *The Quarterly Journal of Economics*, 131(4): 1973–2042.

**Hausman, Jerry A.** 1996. "Valuation of new goods under perfect and imperfect competition." In *The economics of new goods*. 207–248. University of Chicago Press.

**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."

**Karlan, Dean, and Jonathan Zinman.** 2018. "Price and control elasticities of demand for savings." *Journal of Development Economics*, 130: 145–159.

**Karlan, Dean, and Jonathan Zinman.** 2019. "Long-run price elasticities of demand for credit: evidence from a countrywide field experiment in Mexico." *The Review of Economic Studies*, 86(4): 1704–1746.

**Kitchens, Carl, and Price Fishback.** 2015. "Flip the switch: the impact of the rural electrification administration 1935–1940." *The Journal of Economic History*, 75(4): 1161–1195.

**Kline, Patrick, and Enrico Moretti.** 2014. "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority." *The Quarterly Journal of Economics*, 129(1): 275–331.

**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*.

**Lee, Kenneth, Edward Miguel, and Catherine Wolfram.** 2016. "Appliance ownership and aspirations among electric grid and home solar households in rural Kenya." *The American Economic Review*, 106(5): 89–94.

**Lee, Kenneth, Edward Miguel, and Catherine Wolfram.** 2020*a*. "Does Household Electrification Supercharge Economic Development?" *Journal of Economic Perspectives*, 34(1): 122–44.

**Lee, Kenneth, Edward Miguel, and Catherine Wolfram.** 2020*b*. "Experimental Evidence on the Economics of Rural Electrification." *Journal of Political Economy*, 128(4).

**Lee, Kenneth, Eric Brewer, Carson Christiano, Francis Meyo, Edward Miguel, Matthew Podolsky, Javier Rosa, and Catherine Wolfram.** 2014. "Barriers to Electrification for "Under Grid" Households in Rural Kenya." National Bureau of Economic Research.

**Lewis, Joshua, and Edson Severnini.** 2019. "Short-and long-run impacts of rural electrification: evidence from the historical rollout of the US power grid." *Journal of Development Economics*, 102412.

**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.** 2001. "Measuring market power in the ready-to-eat cereal industry." *Econometrica*, 69(2): 307–342.

**Peletz, Rachel, Alicea Cock-Esteb, Dorothea Ysenburg, Salim Haji, Ranjiv Khush, and Pascaline Dupas.** 2017. "Supply and demand for improved sanitation: results from randomized pricing experiments in rural Tanzania." *Environmental science & technology*, 51(12): 7138–7147.

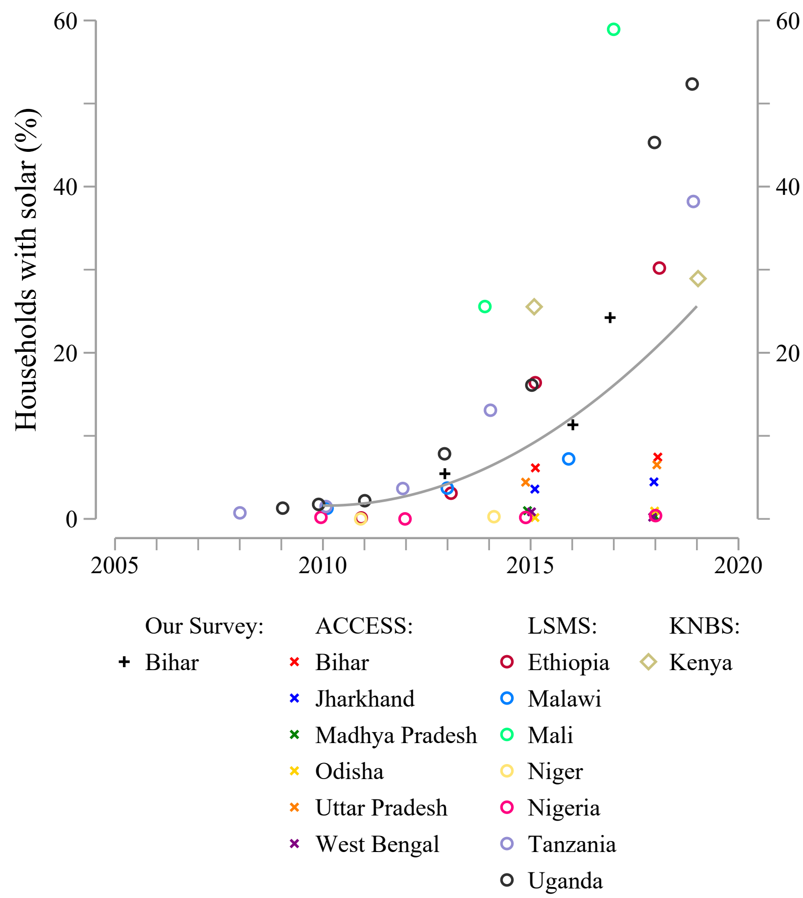
**Petrin, Amil.** 2002. "Quantifying the benefits of new products: The case of the minivan." *Journal of political Economy*, 110(4): 705–729.

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

**Todd, Petra E, and Kenneth I Wolpin.** 2006. "Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility." *American economic review*, 96(5): 1384–1417.

# 8 Figures

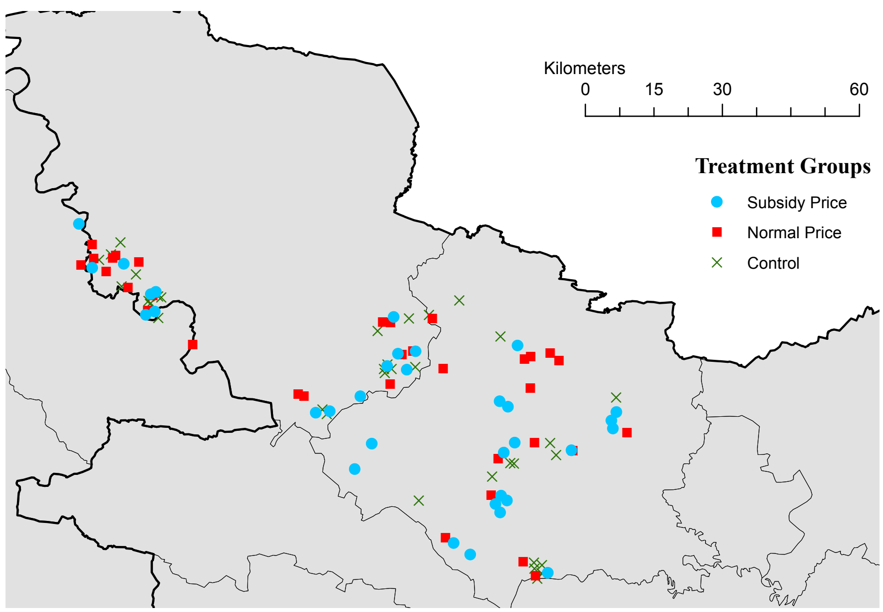
Figure 1: Household solar adoption in select developing countries




The figure shows the growth of solar power in rural areas of select Indian states and African countries, collectively representing a population of roughly 1 billion people. The figure uses four sources of data. First, the "+" markers presents our survey data from several districts in Bihar. Second the "×" markers represent data from the Access to Clean Cooking Energy and Electricity Survey of States (ACCESS) survey, conducted by the Council on Energy, Environment, and Water (CEEW). Third, the circle markers present data from the Living Standards Measurement Study (LSMS) surveys conducted by the World Bank. Fourth, the diamond markers present data from the Kenya Intergrated Household Budget Survey (2015) and Kenya Continuous Household Survey Programme (2019), both conducted by the Kenya National Bureau of Statistics (KNBS). All surveys are representative of the rural population in the state or country covered. Solar market shares are calculated as the proportion of sample households who own a solar panel. The grey line is a quadratic function fit to the underlying share data beginning in 2010.

Figure 2: Maps of study area

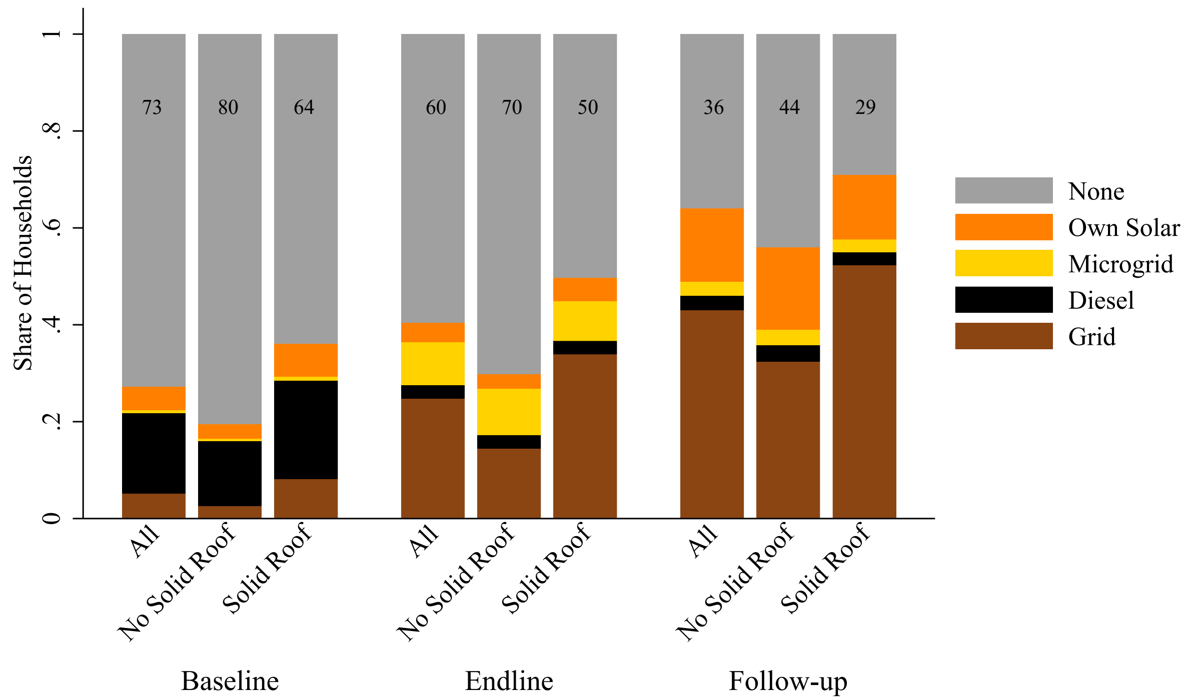
A. Study districts within the state of Bihar, India




B. Sample villages within study districts

The figure shows the study area. Panel A highlights the two districts of West Champaran and East Champaran, in the northwest corner of Bihar, which contain the study villages. 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 3: Household Electricity Sources Over Time

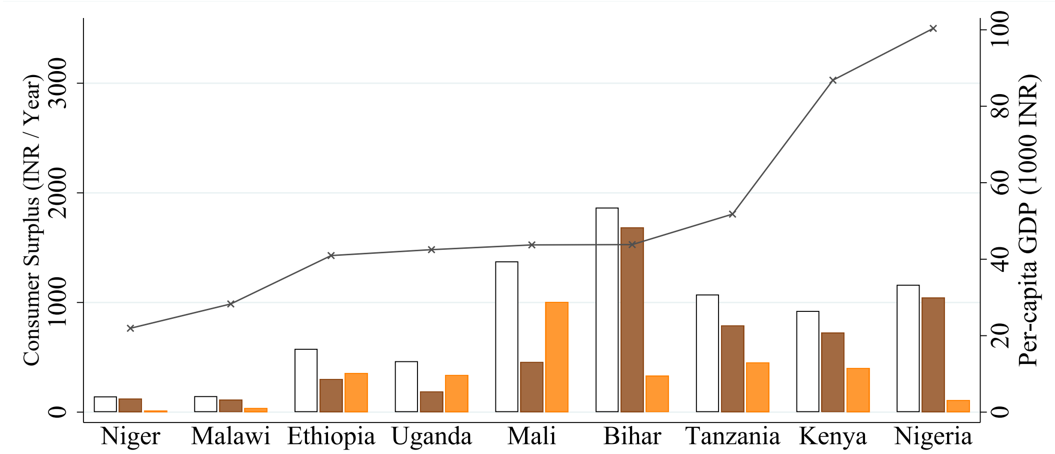



The figure shows the market shares of different sources of electricity over time. Each stacked bar gives the share of households, from bottom to top, that use grid electricity, diesel generators, solar microgrids, own solar systems or no electricity. These market shares are calculated with respect to the total sample of households, without regard for whether a source is available in a village or not; in a village where the grid is not present, for example, the grid necessarily has a zero share. There are three clusters of bars, for shares in the baseline (starting November 2013), endline (starting May 2016) and follow-up (starting May 2017) survey waves. We use a dummy variable for whether a household has a solid roof as a proxy for household assets. Within each cluster of bars, the three bars from left to right give the market shares amongst all households, households that do not have a solid roof, and households that do have a solid roof, respectively.

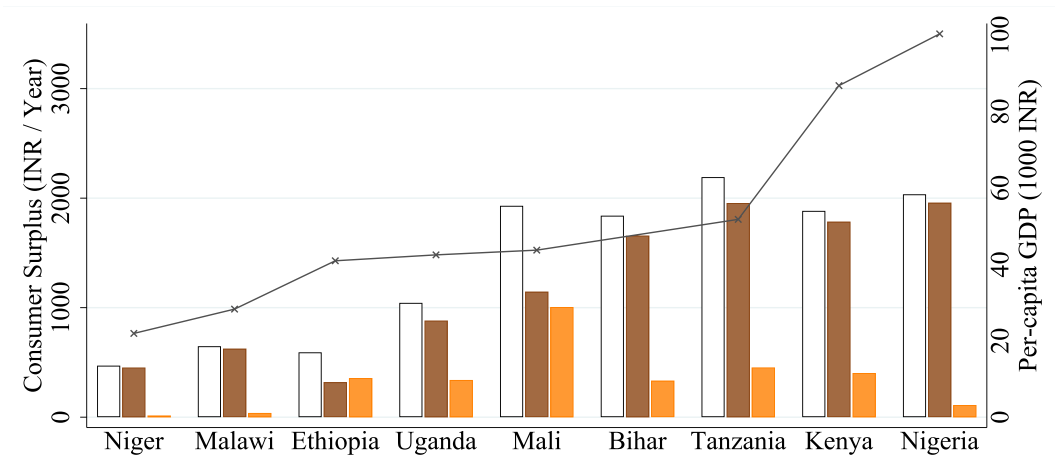
Figure 4: Consumer Surplus from microgrids versus all sources from electricity


The figures compares estimates of consumer surplus from microgrids to the surplus from all electricity sources. The bars on the left give estimates of the surplus from microgrids using the reduced-form IV demand estimates of Table 3. The bars on the right give surplus estimates from our full structural demand model, presented in Table 4 column 3 and Table 5 columns 6 through 10. The "Microgrids only" bars present the change in surplus, in the full demand model, if microgrids are removed from the market. The "All sources" bars present the consumer surplus from all sources of electricity. We present two possible scenarios at which to evaluate surplus. Hollow bars present surplus evaluated at endline, using the observed prices for microgrids. Filled bars present surplus evaluated at endline, but assuming microgrids are available everywhere at a counterfactual price of 100 INR.

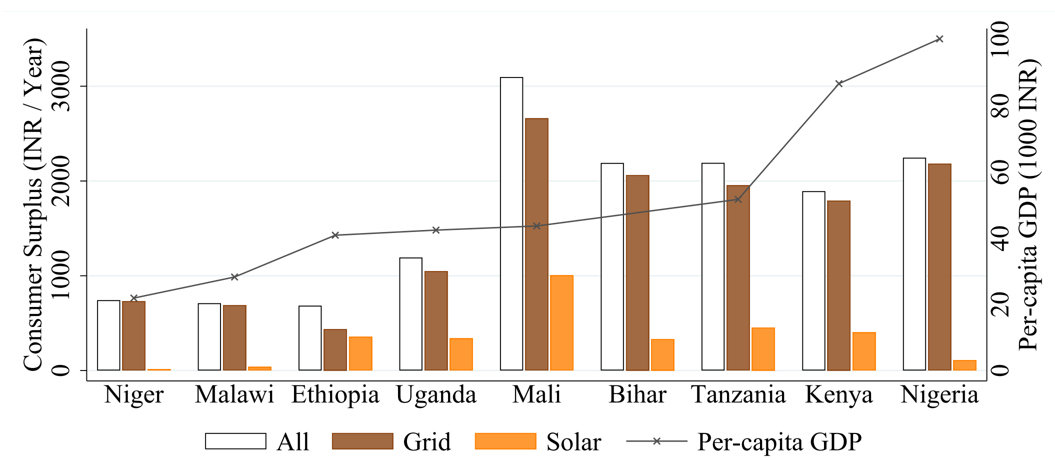
Figure 5: Consumer surplus from electricity on the global electrification frontier, by source




A. Consumer surplus estimates at data




B. Grid supply priced at Bihar median grid price




C. Grid supply priced at Bihar median + universal grid

The figure compares estimates of consumer surplus from electrification versus consumer surplus when only grid or only solar is available. The estimates are derived from out full structural demand model extended to select African countries. The "X" line shows per capita GDP in a give year/state in thousands INR, adjusted using PPP. Changes in consumer surplus are presented for Niger in 2014, Malawi in 2016, Mali in 2017, Ethiopia in 2018, Uganda in 2019, Tanzania in 2019, Nigeria in 2018, and Kenya in 2019. Estimates for Indian states are based on 2018 wave of the ACCESS survey. Surplus for Bihar is based on our experimental sample.

# 9 Tables

Table 1: Summary of Electricity Sources

| | Baseline | | | | | Endline | | | | | Follow-up | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Grid (1) | Diesel (2) | Own solar (3) | Micro-grid (4) | None (5) | Grid (6) | Diesel (7) | Own solar (8) | Micro-grid (9) | None (10) | Grid (11) | Diesel (12) | Own solar (13) | Micro-grid (14) | None (15) |
| *Panel A. Source characteristics* | | | | | | | | | | | | | | | |
| Price (INR per month) | 72 | 99 | 80 | 200 | - | 60 | 88 | 91 | 164 | - | 59 | 89 | 72 | 170 | - |
| Load (watts) | 322 | 134 | 247 | 31 | - | 145 | 22 | 39 | 31 | - | 147 | 40 | 13 | 31 | - |
| *Hours of supply* | | | | | | | | | | | | | | | |
| Total | 10.9 | 3.4 | 7.4 | 5.3 | - | 11.0 | 3.1 | 5.6 | 5.6 | - | 13.6 | 3.1 | 5.6 | 5.6 | - |
| Peak (5 - 10 pm) | 2.0 | 3.4 | 4.7 | 4.3 | - | 2.1 | 3.1 | 4.9 | 5.0 | - | 2.8 | 3.1 | 4.9 | 5.0 | - |
| Off-peak | 8.6 | 0.0 | 2.7 | 1.0 | - | 8.8 | 0.0 | 0.7 | 0.6 | - | 10.4 | 0.0 | 0.7 | 0.6 | - |
| Source in village (%) | 29 | 57 | 100 | 0 | - | 53 | 18 | 100 | 66 | - | 72 | 13 | 100 | 66 | - |
| *Panel B. Household appliance ownership* | | | | | | | | | | | | | | | |
| Light bulb (%) | 84 | 93 | 72 | 55 | 2 | 100 | 100 | 99 | 66 | 1 | - | - | - | - | - |
| Mobile phone (%) | 87 | 89 | 97 | 90 | 74 | 95 | 95 | 97 | 92 | 86 | - | - | - | - | - |
| Fan (%) | 22 | 2 | 1 | 0 | 0 | 34 | 4 | 9 | 3 | 1 | - | - | - | - | - |
| Television (%) | 15 | 3 | 10 | 15 | 1 | 11 | 1 | 4 | 2 | 0 | - | - | - | - | - |
| Radio (%) | 11 | 11 | 14 | 10 | 7 | 6 | 9 | 4 | 5 | 4 | - | - | - | - | - |
| Pump (%) | 4 | 1 | 3 | 0 | 2 | 4 | 5 | 7 | 6 | 2 | - | - | - | - | - |
| Iron (%) | 4 | 1 | 2 | 0 | 0 | 3 | 0 | 1 | 0 | 0 | - | - | - | - | - |

The table summarizes the characteristics of electricity sources available in our sample. The overarching column headers show each electricity source in each survey wave: baseline (starting November 2013), endline (starting May 2016) and follow-up (starting May 2017). The individual columns then indicate each electricity source. Panel A shows source attributes weighted by sample size at the village level. Price shown is the average monthly price for each electricity source; for grid, the price takes theft into account by multiplying reported payment by the percentage of households that actually pay. Load is imputed based on what appliances the households say they have plugged in. Hours of supply refers to hours per day of electricity supply; for grid, supply comes from administrative data and for the non-grid sources, supply comes from the respective household survey. The final row in Panel A shows the percent of villages where the given source is available. Panel B shows the share of households that own the most popular appliances. Appliance ownership at the follow-up survey is not available, as we did not collect these variables during this thin round of survey.

Table 2: Household Characteristics and Experimental Balance

| | Control (1) | Normal (2) | Subsidy (3) | N - C (4) | S - C (5) | *F*-Test (6) |
|---|---|---|---|---|---|---|
| | *Panel A. Demographics* | | | | | |
| Education of household head (1-8) | 2.41 | 2.67 | 2.58 | 0.26* | 0.17 | 1.48 |
| | [2.03] | [2.14] | [2.09] | (0.15) | (0.15) | (0.23) |
| 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) |
| | *Panel B. Wealth proxies* | | | | | |
| Household income (INR ’000s/month) | 7.46 | 7.32 | 7.28 | -0.14 | -0.18 | 0.068 |
| | [6.88] | [6.86] | [7.03] | (0.56) | (0.50) | (0.93) |
| 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) |
| Solid house (=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) |
| Owns ag. land (=1) | 0.67 | 0.69 | 0.67 | 0.015 | 0.0022 | 0.045 |
| | [0.47] | [0.46] | [0.47] | (0.056) | (0.053) | (0.96) |
| Solid roof (=1) | 0.42 | 0.46 | 0.51 | 0.042 | 0.095** | 3.08* |
| | *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) |
| Uses grid (=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) |
| Uses diesel (=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) |
| Uses own solar (=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) |
| 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) |
| Observations | 1052 | 983 | 1001 | | | |

The table reports the balance of covariates in our baseline survey across treatment 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 treatment 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. 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. The rightmost 3 columns have standard errors clustered at the village-level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3: Solar Microgrid Demand

| | ITT Estimates | | | IV Estimates | |
|---|---|---|---|---|---|
| *Survey wave:* | Baseline | Endline | Follow-up | Endline | Endline |
| *Dependent variable:* | Share | Share | Share | Share | log(Share) |
| | (1) | (2) | (3) | (4) | (5) |
| Treatment: Subsidized price | -0.001 | 0.193$^{***}$ | 0.081$^{***}$ | | |
| | (0.005) | (0.049) | (0.027) | | |
| Treatment - normal (R. 160/200) | 0.009 | 0.060$^{**}$ | 0.020$^{*}$ | | |
| | (0.010) | (0.028) | (0.012) | | |
| Price (INR '00s) | | | | -0.129$^{**}$ | |
| | | | | (0.052) | |
| log(Price) | | | | | -0.997$^{***}$ |
| | | | | | (0.386) |
| Constant | 0.006 | 0.023$^{***}$ | 0.002 | 0.347$^{***}$ | -2.079$^{***}$ |
| | (0.004) | (0.005) | (0.002) | (0.091) | (0.189) |
| Observations | 100 | 100 | 100 | 66 | 66 |
| First-stage $F$-Stat | | | | 676 | 1107 |

The table shows estimates of microgrid demand. The dependent variable in the first 3 columns is the village-level market share of microgrid solar. The independent variables are the subsidized price arm (microgrids offered at INR 100) and a normal price arm (microgrids offered at the prevailing price of INR 200, later cut to INR 160 in some villages). The control arm (microgrids not offered) is omitted. Each column measures market share at one of the three survey waves. Columns 4 and 5 show instrumental variables estimates of the demand curve using linear and log-log specifications, respectively. We instrument for price using a dummy for the subsidized treatment arm. Standard errors are shown in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table 4: Demand for Electricity: Estimates of Linear Stage

| | OLS | Price IV | Price & Hours IV | | |
|---|---|---|---|---|---|
| *Instruments* | | RCT | RCT | BLP | Hausman |
| | (1) | (2) | (3) | (4) | (5) |
| | | *Panel A. Linear Stage Estimates* | | | |
| Price (INR 100) | -2.05*** | -5.05*** | -4.97** | 10.5 | 15.2 |
| | (0.29) | (1.96) | (1.95) | (18.2) | (38.0) |
| Hours of peak supply | 3.71** | 3.54** | 5.15** | 5.85** | 6.05* |
| | (1.80) | (1.77) | (2.59) | (2.65) | (3.45) |
| Hours of off-peak supply | -1.10*** | -1.08*** | -1.54*** | -1.67*** | -1.71** |
| | (0.41) | (0.40) | (0.58) | (0.59) | (0.71) |
| $\xi_{tj}$ mean effects | Yes | Yes | Yes | Yes | Yes |
| | | *Panel B. First Stage Estimates (Price)* | | | |
| Treatment normal price × Endline (=1) | | 0.064** | 0.064** | | |
| | | (0.029) | (0.028) | | |
| Treatment subsidy price × Endline (=1) | | -0.16*** | -0.16*** | | |
| | | (0.021) | (0.021) | | |
| Hours of peak supply | | -0.050 | | | |
| | | (0.049) | | | |
| Hours of off-peak supply | | 0.0081 | | | |
| | | (0.013) | | | |
| Mean peak hours in nearby villages | | | -0.032 | -0.040 | -0.038 |
| | | | (0.045) | (0.043) | (0.044) |
| Mean off-peak hours in nearby villages | | | 0.0040 | 0.0057 | 0.0051 |
| | | | (0.0094) | (0.0090) | (0.0091) |
| Total supply of competing products | | | | -0.0017 | |
| | | | | (0.0060) | |
| Load of competing products | | | | 0.0079 | |
| | | | | (0.011) | |
| Mean price in nearby villages | | | | | 0.038 |
| | | | | | (0.080) |
| $\xi_{tj}$ mean effects | | Yes | Yes | Yes | Yes |
| Observations | 999 | 999 | 999 | 945 | 989 |
| First Stage F-Stat | | 42.1 | 21.1 | 0.4 | 0.5 |

The table presents estimates of the second, linear stage of our demand system (equation 12). The dependent variable is mean indirect utility at the market-by-survey wave level, estimated in the non-linear first stage. Peak hours refers to electricity supply during the evening, from 5 to 10 pm, and off-peak to other hours of the day. The columns estimate the same equation either by ordinary least squares (column 1) or instrumental variables (columns 2 to 5). Each column uses a different set of instruments. In column 2, we use the experimental treatment assignments interacted with a dummy for the endline survey as instruments (equation 13). In column 3, we additionally instrument for hours of supply, on-peak and off-peak, using the predicted hours of supply based on supply in nearby villages. In columns 4 and 5 we replace the experimental instruments with instruments from the industrial organization literature (Berry, Levinsohn and Pakes, 1995; Nevo, 2001; Hausman, 1996). Column 4 uses the average characteristics of the other products available in a given village as instruments (Berry, Levinsohn and Pakes, 1995). The characteristics we use are hours of supply and load. Column 5 uses the average price of each product in the nearest three villages as instrument for its price in a given village (Nevo, 2001; Hausman, 1996). All regressions control for wave-by-source fixed effects. The final row of the table reports the first-stage $F$-statistic from the price equation. Standard errors are clustered at the village-level and shown in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table 5: Electricity Source Choice Probabilities by Household Characteristics - Median Households

| | Simple Model | | | | | Full Model | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Grid (1) | Diesel (2) | Own Solar (3) | Micro-grid (4) | None (5) | Grid (6) | Diesel (7) | Own Solar (8) | Micro-grid (9) | None (10) |
| *Panel A. Marginal effects of household characteristics* | | | | | | | | | | |
| Number of adults | 0.259 | 0.015 | -0.022 | -0.020 | -0.233 | 0.228 | 0.012 | -0.017 | -0.019 | -0.205 |
| | (0.022) | (0.006) | (0.005) | (0.008) | (0.016) | (0.042) | (0.007) | (0.008) | (0.014) | (0.031) |
| Solid roof (=1) | 0.226 | 0.013 | -0.001 | -0.020 | -0.218 | 0.174 | 0.008 | 0.006 | -0.022 | -0.165 |
| | (0.017) | (0.007) | (0.008) | (0.012) | (0.016) | (0.029) | (0.007) | (0.008) | (0.016) | (0.024) |
| Owns ag. land (=1) | | | | | | 0.092 | 0.001 | -0.014 | 0.003 | -0.082 |
| | | | | | | (0.031) | (0.009) | (0.010) | (0.015) | (0.024) |
| Education of household head (1-8) | | | | | | 0.086 | 0.009 | -0.013 | -0.010 | -0.072 |
| | | | | | | (0.031) | (0.005) | (0.006) | (0.012) | (0.022) |
| Household income | | | | | | 0.029 | 0.002 | 0.003 | 0.024 | -0.059 |
| | | | | | | (0.027) | (0.005) | (0.006) | (0.014) | (0.020) |
| *Panel B. Random coefficient parameter* | | | | | | | | | | |
| Price Dispersion Parameter | 3.30 | | | | | 3.09 | | | | |
| | (0.167) | | | | | (0.173) | | | | |
| Observations | 8822 | | | | | 8822 | | | | |
| Objective Value | .1283 | | | | | .0956 | | | | |
| Wald Test Statistic ($p$-value) | | | | | | 601.71 (0.000) | | | | |

The table shows the effects of household characteristics on the probability of a household choosing a given electricity source. The table reports the results of two models. A simple model, reported in columns 1 through 5, includes as covariates the number of adults in the household and a dummy variable for whether the household has a solid roof. Our full model, reported in columns 6 through 10, includes three additional observable proxies for household demand: household income, whether the household owns agricultural land, and years of education of the household head. Both models also include a triangularly-distributed random coefficient on price, and the estimated dispersion parameter of the price distribution (corresponding to half the range of the distribution) is reported in Panel B. The effects of household characteristics are nonlinear. The table therefore reports "marginal" effects evaluated for a "median" household facing the end-line availabilities, qualities, and prices of each good. The profile of a median household is defined as a household of three adults living in a house with a solid roof, owning agricultural land, and income at the endline median. See Appendix Table C4 for the characteristics of a median household. The marginal effects are not truly marginal; for binary variables, we report the effect on choice probability of changing the value from one to zero, and for continuous variables the effect of a one standard deviation increase in that variable. We also report the minimized objective value for both models. The objective function is different between the two models. First, the objective functions use different weighting matrices derived from the 2-step "optimal" GMM estimator. Second, the full model objective function contains an additional moment for each additional parameter estimated. Therefore, the magnitude of the objective functions reported is not directly comparable. Finally, we report a Wald test statistic, distributed chi-squared with 12 degrees of freedom, from a test of the restriction that the coefficients on the covariates added in the full model are jointly zero.

Table 6: Price Elasticities of Electricity Source Demand

| | Elasticity of share for source: | | | | |
|---|---|---|---|---|---|
| with respect to price of source: | Grid (1) | Diesel (2) | Own solar (3) | Micro-grid (4) | None (5) |
| *Panel A. Elasticities at endline prices* | | | | | |
| Grid | -0.20 | 0.05 | 0.05 | 0.05 | 0.06 |
| Diesel | 0.01 | -1.68 | 0.09 | 0.08 | 0.06 |
| Own solar | 0.02 | 0.23 | -2.14 | 0.30 | 0.16 |
| Microgrid | 0.05 | 0.37 | 0.79 | -1.95 | 0.21 |
| *Panel B. Elasticities at break-even prices for grid* | | | | | |
| Grid | -0.82 | 0.18 | 0.14 | 0.15 | 0.10 |
| Diesel | 0.02 | -1.73 | 0.08 | 0.08 | 0.06 |
| Own solar | 0.03 | 0.21 | -2.17 | 0.31 | 0.17 |
| Microgrid | 0.07 | 0.36 | 0.80 | -1.99 | 0.22 |

The table presents aggregate own- and cross-price elasticities of demand by electricity source. In panel A, the elasticities are calculated using a 10% increase in each source's price from its mean endline price. Panel B is similar but fixes the price of grid at the break-even price for the utility (calculated as 233 INR = 60 KwH per month mean consumption $\times 3.88$ INR per KwH average cost of procurement). Both panels use IV estimates of mean utility from source characteristics based on the RCT. The elasticities are calculated for the market share of each column source with respect to the price of each row source.

Table 7: The Value of Electrification under Counterfactual Policies

| | Market shares | | | | | Surplus (INR per hh per year) | | |
|---|---|---|---|---|---|---|---|---|
| | Grid | Diesel | Own solar | Microgrid | All | Consumer | Producer | Total |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Panel A. Model market shares and surplus by survey wave* | | | | | | | | |
| 1. Model at baseline | 6 | 17 | 7 | 1 | 31 | 380 | -108 | 272 |
| 2. Model at endline | 24 | 3 | 7 | 10 | 43 | 1052 | -492 | 560 |
| 3. Model at follow-up | 40 | 3 | 17 | 5 | 65 | 1866 | -842 | 1024 |
| *Panel B. Disruption due to the improvement of solar power, relative to baseline* | | | | | | | | |
| 1. Model at baseline, no solar | 6 | 18 | 0 | 0 | 24 | 341 | -109 | 232 |
| 2. Model at baseline + Improved solar | 5 | 13 | 23 | 8 | 50 | 562 | -104 | 459 |
| 3. Model at baseline + Improved solar + Grid priced at cost | 2 | 15 | 24 | 8 | 49 | 494 | -0 | 494 |
| *Panel C. Disruption due to the improvement of grid electricity, relative to baseline* | | | | | | | | |
| 1. Model at baseline, no grid | 0 | 19 | 8 | 1 | 28 | 243 | | |
| 2. Model at baseline + Improved grid | 37 | 9 | 4 | 1 | 51 | 1588 | -766 | 822 |
| 3. Model at baseline + Improved grid + Grid priced at cost | 21 | 12 | 5 | 1 | 40 | 978 | 0 | 978 |
| *Panel D. Disruption due to improvements in grid and solar* | | | | | | | | |
| 1. Model at baseline + Improved grid and solar | 36 | 7 | 15 | 5 | 63 | 1699 | -750 | 950 |
| 2. Model at baseline + Improved grid and solar + Price grid at cost | 21 | 10 | 19 | 5 | 54 | 1107 | 0 | 1107 |
| *Panel E. Future growth in electrification via supply and demand shifts, relative to follow-up* | | | | | | | | |
| 1. Model at follow-up | 40 | 3 | 17 | 5 | 65 | 1866 | -842 | 1024 |
| 2. ... + 50% solar cost reduction | 38 | 2 | 27 | 15 | 81 | 2050 | -781 | 1269 |
| 3. ... + Grid in all villages | 47 | 1 | 22 | 12 | 83 | 2334 | -982 | 1352 |
| 4. ... + Increase in peak grid hours | 70 | 1 | 13 | 6 | 90 | 3895 | -1463 | 2432 |
| 5. ... + All households at least median income | 92 | 0 | 5 | 1 | 98 | 5856 | -1908 | 3948 |
| 6. ... + Price grid at cost | 77 | 1 | 11 | 4 | 94 | 4049 | 0 | 4049 |

The table presents market shares and surplus under counterfactual changes in the electricity market. The counterfactual scenarios are laid out in Section 5 of the text and the detailed assumptions behind the counterfactuals are in Appendix Table D1. All counterfactuals are calculated using the full demand model estimates of Table 5, columns 6 through 10. For each counterfactual, columns 1 to 4 give the market shares of each source, column 5 gives the electrification share, and columns 6 through 8 give consumer, producer and total surplus. Consumer surplus is the amount in INR per household per year that households would be willing to pay for a given choice set, relative to having only the outside option of no electricity. The amounts of both consumer and producer surplus are averaged over the entire sample of consumers, regardless of their choice. Producer surplus is the variable profit of the state utility that provides grid electricity. Levels rows are unindented, whereas changes rows (where the numbers displayed are differences in two counterfactual scenarios) are indented.

# A Appendix: Data

This Appendix describes our data collection and the construction of instrumental variables for hours of electricity supply. Our primary data for Bihar comes from four sources: a panel survey, microgrid administrative data, state utility administrative data, and a survey of diesel generator operators. In additional to these primary data, we utilize secondary household-level survey data covering selected states in India as well as Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, Uganda, and Kenya.

Survey data from additional Indian states comes from the Access to Clean Cooking Energy and Electricity - Survey of States (ACCESS). This survey covers over 9000 households in India across Madhya Pradesh, Uttar Pradesh, Bihar, Jharkhand, West Bengal, and Odisha. Households are asked questions regarding energy access and a vector of household characteristics is also collected.

The data on seven African countries (Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda) is sourced from the Living Standards Measurement Study - Integrated Surveys of Agriculture (LSMS). The data on Kenya is sourced from two surveys by the Kenya National Bureau of Statistics: the Kenya Integrated Household Budget Survey in (KIHBS) in 2015 and the Kenya Continuous Household Survey Programme (KCHSP) in 2019.

## 1.1 Bihar experimental data

We describe our panel survey in Section 2 2.2. We also draw on three other original data sources, which are described below.

*Microgrid administrative data.* The second source of data is an administrative dataset on microgrid customers from HPS. We partnered with HPS to roll-out solar microgrids experimentally in the sample villages (see Section 3). The dataset includes enrollment, pricing and customer payments from January 2014 to January 2016, which we match with our household surveys. This matching allows us to estimate demand in administrative payments data, to complement our survey-based estimates.

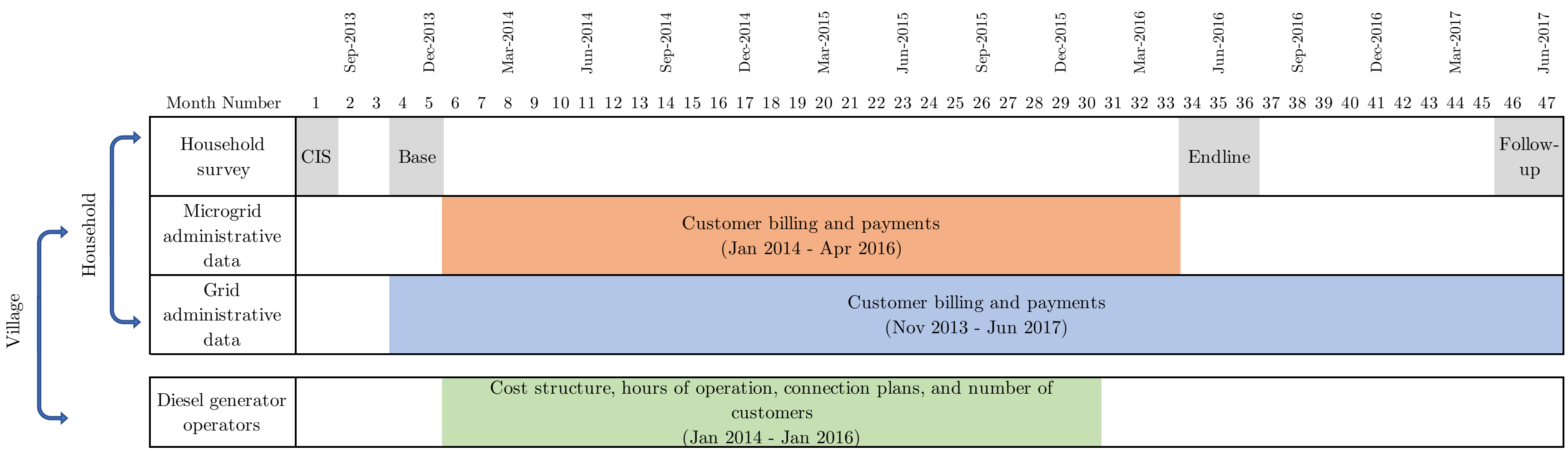
*State utility administrative data.* We use three datasets pertaining to grid electricity: (i) a consumer database for all formal customers, (ii) a billing and collections dataset containing bills and customer payments, and (iii) village-level hours of supply, recorded from administrative log-books. The data sources (i) and (ii) are matched at the customer level to our survey respondent households. Many households using the grid in the survey are not matched to the administrative database, as there are high rates of informal connections, i.e. electricity theft, in Bihar. We can measure informal connections by designating households informal if they could not provide a customer ID from their electricity bill, or the ID provided did not match the utility's billing database.

*Survey of diesel generator operators.* Our final source of data is a survey of diesel generator operators. Entrepreneurs set-up diesel generators and connect customers within non-electrified villages, providing electricity to fifty or more households at a time. We surveyed these operators to collect data on operating costs, hours of operation, pricing and customers served from January 2014 to 2016.

These sources of data allow us to see, on the demand side, a rich set of household characteristics and the sources and uses of electricity. On the supply side, we have data on all the competing sources in the marketplace, in some cases from both administrative sources and our household and operator surveys.

**Sampling and timeline .—**Figure A1 shows the timeline for the implementation of the experiment and the timing of the data collection. The microgrid experiment ran for roughly 2.5 years but the data collection spanning the experiment covered roughly 4 years in total.

Figure A1: Data Collection Timeline




The figure shows the collection of data during the study from several sources: a household survey, microgrid administrative data, grid administrative data and a diesel operator survey. See Section 2 2.2 for a detailed description of the data sources. In August 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 sampling frame for our household sample. The experiment ran from January 2014, after the conclusion of the baseline survey, through the endline survey in mid-2016.

**Construction of hours of supply for grid .—**The household survey provides most of the characteristics of sources that we use in our demand model. An exception is the grid hours of supply, which we obtained from administrative logbooks maintained by the North Bihar Power Distribution Company Limited. The logbooks record the hours when the grid is switched on and off at the level of the feeder, the lowest level of the distribution network at which the company exercises control over power supply. We aggregated this data from the hourly level to compute average daily hours of electricity supply to each feeder, both on-peak (from 5 to 10 pm) and off-peak. We then mapped our 100 sample villages to their respective supply feeders.

Some villages were missing data around the time of our endline survey. If a village was missing data during the endline survey, we imputed hours of supply with the hours of supply data for that same village within a window running from 6 months before to 6 months after the survey. If the village had no data in that window, we imputed hours of supply data based on the hours of supply for the three nearest villages for which we had data, using a random forest model. The model additionally included as covariates latitude and longitude, division fixed effects and their interactions. The root mean squared error of our prediction, for villages where data is available, is 1.9 hours.

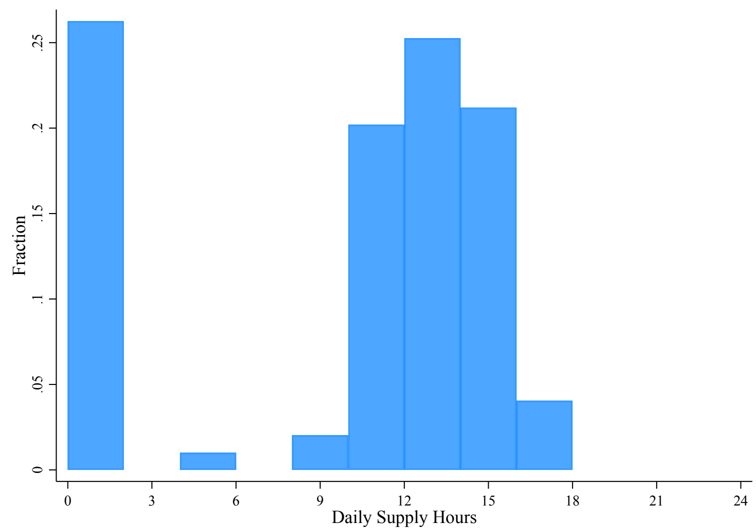
**Construction of instrument for hours of supply .—**The experiment provides instruments for price but not for other product characteristics, which in principle may also be endogenous to demand: for example, a high-demand village may be given more supply by the distribution company. Our preferred specification for the second-stage linear IV estimation therefore instruments for price, peak, and off-peak hours of electricity supply.

The instruments for peak and off-peak hours of supply are the predicted peak and off-peak hours of supply for a given village based upon hours of supply to nearby villages, as described in Appendix Section above. We expect hours of supply to nearby villages to be correlated since they are served by the same feeders or by separate feeders from the same substation, which would experience correlated supply shocks such as for rationing decisions.

For non-grid sources we set predicted hours of supply based on their technological characteristics. We set off-peak hours for diesel and microgrid solar to be zero, and assume that all supply is on peak. For own solar, we set peak and off-peak supply to be constant and equal to the global mean of each variable. In this way, there is no variation in predicted supply for off-grid sources and so the variation to identify the coefficients on supply hours come solely from variation in predicted supply for grid electricity.

**Summary statistics .**—Appendix Figure A2 shows the distribution of daily hours of electricity supply on the grid and Figure A3 the distributions of supply hours during off-peak and on-peak times. Appendix Table A1 shows the market shares of electricity sources at endline accounting for the possibility of ownership of multiple sources.

Figure A2: Daily Hours of Supply on the Grid



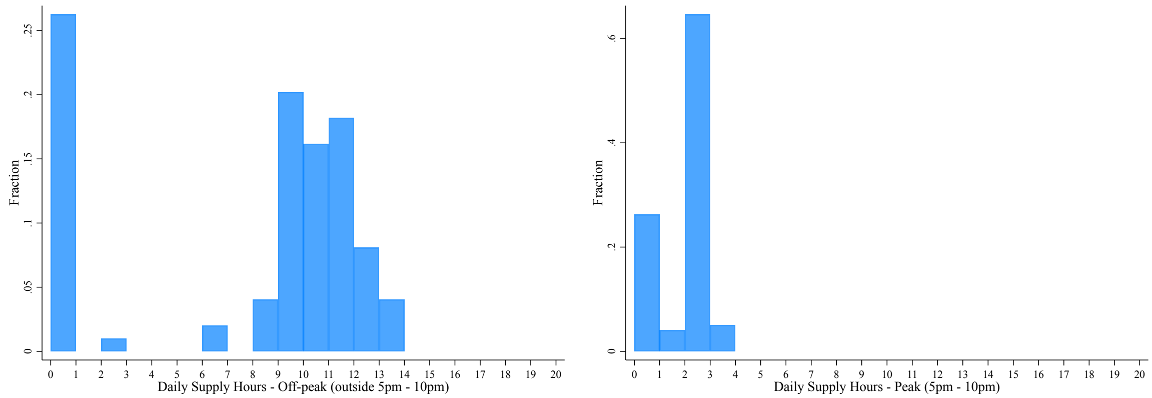

This figure shows the distribution of the daily average hours of grid electricity supply across villages in our sample at the endline survey.

Table A1: Electricity Source Ownership at Endline

| | Frequency (1) | Percentage (2) | Cumulative Percentage (3) |
|---|---|---|---|
| Grid | 681 | 22.43 | 22.43 |
| Diesel | 81 | 2.67 | 25.10 |
| Own solar | 148 | 4.87 | 29.97 |
| Microgrid | 141 | 4.64 | 34.62 |
| Grid & Own solar | 28 | 0.92 | 35.54 |
| Grid & Microgrid | 14 | 0.46 | 36.00 |
| None | 1824 | 60.08 | 96.08 |
| No data | 119 | 3.92 | 100.00 |
| Total | 3036 | 100.00 | 100.00 |

This table shows the household level take-up rate for different electricity sources, accounting for joint ownership, at the endline survey.

Figure A3: Hours of grid supply off-peak and on-peak




A. Off-peak supply B. Peak supply

The figure shows the distribution of grid hours of supply. The data come from administrative logbooks of hourly supply to sample villages. Panel A shows the distribution of hours of supply during the off-peak period and Panel B during the peak period of 5 to 10 pm. The maximal possible hours of supply in the peak period is therefore 5 hours and during the complementary off-peak period 19 hours.

## 1.2 ACCESS

The data on selected Indian states was sourced from the Access to Clean Cooking Energy and Electricity - Survey of States (ACCESS), conducted by the Council on Energy, Environment and Water among the rural population in six states: Madhya Pradesh, Uttar Pradesh, Bihar, Jharkhand, West Bengal, and Odisha. ACCESS uses a three-stage stratified sampling design. First, districts are sampled from each administrative division within each state. Second, villages are divided into “small" and “large" village groups and seven villages are sampled from each group. Finally, twelve households are sampled from each village. This procedure leads to a sample that is representative of the six states in aggregate. State-specific estimates are somewhat sensitive to the districts sampled within each state. Currently two waves of ACCESS are available: 2015 and 2018.

**Market identification.—**Our research designs requires a market definition. For identification purposes, it is preferred that each market is about thirty households or larger. In the experimental sample, markets are defined at the village level. In the ACCESS data, we define markets at the district level because villages do not have sufficiently many households to accurately capture source availability. Each state has anywhere between three and eighteen districts, and each district is comprised of 167-168 households. We will refer to these districts as village equivalents. Table A2 contains summary statistics for the village equivalents in our data.

Table A2: Summary of village equivalents in the ACCESS data

| | Year<br>(1) | Households<br>(2) | Villages<br>(3) | Min<br>(4) | Median<br>(5) | Max<br>(6) | Mean<br>(7) |
|---|---|---|---|---|---|---|---|
| Bihar | 2015 | 1511 | 9 | 167 | 168 | 168 | 167.89 |
| Bihar | 2018 | 1512 | 9 | 168 | 168 | 168 | 168.00 |
| Jharkhand | 2015 | 840 | 5 | 168 | 168 | 168 | 168.00 |
| Jharkhand | 2018 | 840 | 5 | 168 | 168 | 168 | 168.00 |
| Madhya Pradesh | 2015 | 1680 | 10 | 168 | 168 | 168 | 168.00 |
| Madhya Pradesh | 2018 | 1680 | 10 | 168 | 168 | 168 | 168.00 |
| Odisha | 2015 | 504 | 3 | 168 | 168 | 168 | 168.00 |
| Odisha | 2018 | 1008 | 6 | 168 | 168 | 168 | 168.00 |
| Uttar Pradesh | 2015 | 3023 | 18 | 167 | 168 | 168 | 167.94 |
| Uttar Pradesh | 2018 | 3024 | 18 | 168 | 168 | 168 | 168.00 |
| West Bengal | 2015 | 1008 | 6 | 168 | 168 | 168 | 168.00 |
| West Bengal | 2018 | 1008 | 6 | 168 | 168 | 168 | 168.00 |

The table presents summary statistics for the number of households in a given village equivalent for each state of India in each year. In the ACCESS data, villages are too small to be used directly, so districts are used as village equivalents.

**Variable construction and imputation.—Household characteristics**: With the exception of solid roof and income, ACCESS data has all household characteristics included in our demand model. To infer whether a household has a solid roof variable, we regressed the solid roof indicator in our experimental data on an indicator for pucca. We then used these coefficients to predict the probability of a solid roof in the ACCESS dataset. Note that these predicted solid roof variables are probabilistic and not discrete. Income was constructed as Consumption + Savings - Borrowings. Table A3 displays summary statistics for each household characteristic.

Table A3: Summary of household characteristics in ACCESS data

| | | Household | | | | House Characteristics | | | | Assets | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Year (1) | Members (2) | Adults (3) | Income (4) | Education (5) | Pucca (6) | Rooms (7) | Land (8) | Roof (9) | Fan (10) | Mo-bile (11) | TV (12) | Fridge (13) |
| Bihar | 2015 | 7.43 | 4.57 | 5,618 | 5.04 | 0.34 | 2.84 | 0.58 | 0.52 | 0.20 | 0.85 | 0.12 | |
| Bihar | 2018 | 7.02 | 4.27 | 6,992 | 4.91 | 0.41 | 2.87 | 0.61 | 0.57 | 0.69 | 0.95 | 0.24 | |
| Jharkhand | 2015 | 6.49 | 4.08 | 5,459 | 4.97 | 0.24 | 3.04 | 0.77 | 0.46 | 0.35 | 0.76 | 0.27 | |
| Jharkhand | 2018 | 5.97 | 3.81 | 5,174 | 4.80 | 0.24 | 2.98 | 0.76 | 0.45 | 0.60 | 0.84 | 0.34 | |
| Madhya Pradesh | 2015 | 6.15 | 3.96 | 4,298 | 5.06 | 0.21 | 2.53 | 0.69 | 0.44 | 0.50 | 0.80 | 0.43 | |
| Madhya Pradesh | 2018 | 5.87 | 3.94 | 4,289 | 4.65 | 0.26 | 2.59 | 0.70 | 0.47 | 0.74 | 0.87 | 0.47 | |
| Odisha | 2015 | 5.91 | 4.05 | 4,157 | 4.63 | 0.19 | 2.58 | 0.52 | 0.42 | 0.52 | 0.63 | 0.34 | |
| Odisha | 2018 | 4.86 | 3.39 | 4,982 | 4.45 | 0.25 | 2.66 | 0.69 | 0.46 | 0.71 | 0.82 | 0.47 | |
| Uttar Pradesh | 2015 | 7.51 | 4.62 | 4,373 | 5.14 | 0.59 | 2.95 | 0.78 | 0.68 | 0.44 | 0.90 | 0.30 | |
| Uttar Pradesh | 2018 | 6.90 | 4.38 | 6,590 | 5.06 | 0.66 | 2.92 | 0.77 | 0.73 | 0.65 | 0.97 | 0.40 | |
| West Bengal | 2015 | 5.08 | 3.57 | 4,785 | 4.61 | 0.22 | 2.20 | 0.69 | 0.44 | 0.88 | 0.64 | 0.51 | |
| West Bengal | 2018 | 4.94 | 3.73 | 6,324 | 4.69 | 0.40 | 2.53 | 0.47 | 0.56 | 0.90 | 0.84 | 0.63 | |

The table presents means of the variables after within and cross-country imputations. Solid roof in the Indian states is predicted from pucca. For the Indian states, income is constructed as Consumption + Savings - Borrowings. Education is converted to the scale used in the experimental data from years of education.

**Supply side characteristics**: Source characteristics were aggregated at the village equivalent level in the ACCESS dataset following the same procedure that was performed in the experimental dataset.

For source availability, we assume that all households have access to solar. We infer that grid or diesel is available to a household if any household in their village equivalent uses grid or diesel respectively. Source availability is summarized in Table A4.

Table A4: Market shares and source availability in ACCESS data

| | | Choice | | | | Availability | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Year (1) | Grid (2) | Solar (3) | Diesel (4) | None (5) | Grid (6) | Solar (7) | Diesel (8) | N (9) |
| Bihar | 2015 | 0.41 | 0.04 | 0.11 | 0.44 | 0.71 | 1.00 | 0.94 | 1,51 |
| Bihar | 2018 | 0.88 | 0.02 | 0.00 | 0.10 | 0.97 | 1.00 | 0.44 | 1,51 |
| Jharkhand | 2015 | 0.64 | 0.02 | 0.00 | 0.33 | 0.88 | 1.00 | 0.25 | 840 |
| Jharkhand | 2018 | 0.83 | 0.03 | 0.00 | 0.14 | 0.93 | 1.00 | 0.25 | 840 |
| Madhya Pradesh | 2015 | 0.86 | 0.00 | 0.00 | 0.14 | 0.92 | 1.00 | 0.00 | 1,68 |
| Madhya Pradesh | 2018 | 0.92 | 0.00 | 0.00 | 0.08 | 0.97 | 1.00 | 0.00 | 1,68 |
| Odisha | 2015 | 0.70 | 0.00 | 0.00 | 0.30 | 0.90 | 1.00 | 0.00 | 504 |
| Odisha | 2018 | 0.86 | 0.00 | 0.00 | 0.14 | 0.98 | 1.00 | 0.00 | 1,00 |
| Uttar Pradesh | 2015 | 0.57 | 0.02 | 0.01 | 0.40 | 0.88 | 1.00 | 0.34 | 3,02 |
| Uttar Pradesh | 2018 | 0.75 | 0.04 | 0.00 | 0.21 | 0.95 | 1.00 | 0.00 | 3,02 |
| West Bengal | 2015 | 0.93 | 0.00 | 0.00 | 0.07 | 1.00 | 1.00 | 0.17 | 1,00 |
| West Bengal | 2018 | 0.96 | 0.00 | 0.00 | 0.04 | 1.00 | 1.00 | 0.00 | 1,00 |

The table presents household-level means of source choice and availability. Market shares are constructed from data on source choice and asset ownership. Where multiple sources are chosen, grid is assumed to be the primary choice. Availability for grid and solar is inferred using village equivalents.

Grid price is set equal to average monthly grid expenditure. Diesel and solar price is calculated as the monthly amortization of the purchase price of a diesel generator or solar panel respectively using a 7-year amortization window and a 20% interest rate. This is analogous to the treatment of solar prices in the experimental dataset. Due to a low number of observations for solar price in West Bengal, Odisha, and Madhya Pradesh, solar prices in these Indian states were replaced with medians of solar prices in the neighboring states (Bihar for West Bengal, Jharkhand for Odisha, and Uttar Pradesh for Madhya Pradesh). For states that were missing source prices for a single wave, the source price for the missing wave was set equal to the mean source price of the non-missing wave. For states that were missing source prices for both waves, prices were set equal to India's mean source price. The hours of supply variable was available for each source directly in the ACCESS dataset. Table A5 summarizes supply side characteristics for the

ACCESS data.

Table A5: Summary of supply side characteristics in ACCESS data

| | | Price (Rs. PPP) | | | Black Out (days) | | | Offpeak Hours | | | Peak Hours | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Year | Grid | Solar | Diesel | Grid | Solar | Diesel | Grid | Solar | Diesel | Grid | Solar |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Bihar | 2015 | 84.36 | 128.06 | 114.92 | 4.03 | 1.75 | 2.77 | 7.47 | 2.84 | 2.66 | 2.52 | 1.79 |
| Bihar | 2018 | 87.01 | 188.75 | 100.28 | 1.52 | 0.58 | 2.41 | 10.64 | 9.17 | 6.68 | 3.85 | 3.81 |
| Jharkhand | 2015 | 91.25 | 78.47 | 78.85 | 3.80 | 3.09 | 3.00 | 6.54 | 2.90 | 4.00 | 1.85 | 2.44 |
| Jharkhand | 2018 | 103.15 | 72.16 | 150.00 | 3.46 | 2.65 | 1.50 | 6.74 | 4.61 | 2.50 | 2.50 | 2.95 |
| Madhya Pradesh | 2015 | 297.10 | 177.65 | 193.53 | 2.48 | 2.53 | | 9.05 | 6.96 | 5.30 | 4.13 | 3.06 |
| Madhya Pradesh | 2018 | 307.30 | 222.06 | 110.89 | 0.37 | 0.71 | | 13.81 | 11.74 | 5.79 | 4.55 | 3.98 |
| Odisha | 2015 | 161.95 | 66.62 | 193.53 | 1.71 | 15.00 | | 13.29 | 2.00 | 5.30 | 4.29 | 1.00 |
| Odisha | 2018 | 146.19 | 59.96 | 110.89 | 1.50 | 4.51 | | 13.97 | 13.38 | 5.79 | 4.64 | 4.80 |
| Uttar Pradesh | 2015 | 195.99 | 271.05 | 180.14 | 3.23 | 2.98 | 2.87 | 6.17 | 5.41 | 1.98 | 2.07 | 2.74 |
| Uttar Pradesh | 2018 | 173.32 | 277.84 | 110.89 | 1.25 | 1.15 | | 9.62 | 7.38 | 5.79 | 3.16 | 2.48 |
| West Bengal | 2015 | 172.59 | 133.24 | 389.80 | 0.88 | 0.94 | 0.00 | 15.93 | 12.62 | 18.00 | 4.79 | 4.09 |
| West Bengal | 2018 | 215.30 | 133.24 | 110.89 | 1.50 | 1.50 | | 15.67 | 17.50 | 5.79 | 4.83 | 4.00 |

The table presents means of the variables after within and cross-country imputations. Grid prices are equal to average monthly grid expenditure. Diesel and solar prices are calculated as the monthly amortization of the purchase price of a diesel generator or solar panel, respectively, using a 7-year amortization at 20% interest rate. Blackout frequency is presented in terms of days in a month. Solar prices in West Bengal were replaced with median solar prices in Bihar, in Odisha - median prices in Jharkhand, and in Madhya Pradesh - median prices in Uttar Pradesh, due to low data availability.

**Market shares**: We constructed market shares for grid, solar, and diesel from household-level electricity source choice and asset ownership variables. In the ACCESS data, it is possible to see if households utilize multiple sources of electricity. Where bundles are observed, we impose the same priority order as in the experimental dataset, in which households are assumed to have chosen the grid if it is part of their chosen bundle. In the ACCESS data, the only large observed bundles are grid and solar (1-11% of households). We summarize market shares in Table A4.

## 1.3 LSMS

The data on seven African Countries (Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda) is sourced from the Living Standards Measurement Study - Integrated Surveys of Agriculture (LSMS), a study conducted by the World Bank. The specific sampling design of the surveys vary, but all produce representative statistics for the rural population at the national level. We selected the first and last wave for each country available as of June 2022. We included all LSMS countries except for Burkina Faso, which only had one wave available. We checked surveys for the inclusion of variables on solar ownership in intervening waves. This variable was available for Ethiopia 2011, 2013, 2015, and 2018; Malawi 2010, 2013, and 2016; Mali 2014 and 2017; Niger 2011 and 2014; Nigeria 2010, 2012, 2015, and 2018; Tanzania 2008, 2010, 2012, 2014, and 2019; and Uganda 2009, 2010, 2011, 2013, 2015, 2018, and 2019.

**Market identification.—**Markets in the experimental data are defined as villages. Because villages are not available in the LSMS data, it is necessary to define markets using a village equivalent. The LSMS survey uses a two-stage sampling scheme. First, the country is divided into small enumeration areas (EAs), typically defined as the smallest administrative unit available in a country. EAs are sampled. Then, households are sampled from each EA. For some countries, the LSMS survey provides the EA to which households belong and the average coordinates of the households that were sampled from that EA. EAs are too small to be used as villages themselves, so they must be grouped together into village equivalents. The LSMS surveys for Nigeria, Niger, Mali, Malawi, and Ethiopia provide coordinates for the EAs for both waves.

We employ a constrained K-mean clustering algorithm (Bradley, Bennett, and Demiriz, 2000) to group EAs into village equivalents. This algorithm takes as an input the minimum number of EAs per cluster. In order to accurately assess source availability, we aim to include no fewer than 30 households per cluster. To translate our inference constraint into the constraint used by the algorithm, it was necessary to drop EAs with very few households. This allowed us to guarantee that if all clusters contained at least some minimum number of enumeration areas, those clusters would also contain at least some minimum number of households.

In Uganda and Tanzania, the coordinates of enumeration areas were not available, so villages are defined by geospatially clustering districts and regions, respectively. We employ a similar constrained K-means clustering algorithm as for the LSMS countries providing EA coordinates, and use centroids for each administrative unit for clustering purposes. Table A6 contains summary statistics for the village-equivalents derived using this method.

Table A6: Summary of village equivalents in the LSMS and KNBS data

| | Year<br>(1) | Households<br>(2) | Villages<br>(3) | Min<br>(4) | Median<br>(5) | Max<br>(6) | Mean<br>(7) |
|---|---|---|---|---|---|---|---|
| Ethiopia | 2011 | 3466 | 68 | 45 | 48 | 156 | 50.97 |
| Ethiopia | 2018 | 3029 | 66 | 40 | 43 | 82 | 45.89 |
| Kenya | 2015 | 13087 | 45 | 210 | 294 | 393 | 290.82 |
| Kenya | 2019 | 12598 | 45 | 132 | 280 | 416 | 279.96 |
| Malawi | 2010 | 10038 | 198 | 47 | 48 | 112 | 50.70 |
| Malawi | 2016 | 10113 | 200 | 46 | 48 | 113 | 50.56 |
| Mali | 2014 | 1786 | 47 | 32 | 36 | 66 | 38.00 |
| Mali | 2017 | 6841 | 176 | 31 | 36 | 107 | 38.87 |
| Niger | 2011 | 2430 | 44 | 40 | 54 | 125 | 55.23 |
| Niger | 2014 | 2261 | 44 | 38 | 50 | 115 | 51.39 |
| Nigeria | 2010 | 3351 | 64 | 47 | 50 | 110 | 52.36 |
| Nigeria | 2018 | 3376 | 66 | 42 | 49 | 100 | 51.15 |
| Tanzania | 2008 | 2039 | 25 | 32 | 72 | 152 | 81.56 |
| Tanzania | 2019 | 671 | 12 | 26 | 49 | 109 | 55.92 |
| Uganda | 2009 | 2086 | 28 | 38 | 70 | 143 | 74.50 |
| Uganda | 2019 | 2233 | 29 | 35 | 75 | 138 | 77.00 |

Villages are defined artificially in African countries. In Nigeria, Niger, Mali, Malawi, and Ethiopia villages are defined by geospatially clustering enumeration areas. In Uganda and Tanzania, villages are defined by geospatially clustering districts and regions, respectively. For Kenya, villages are defined as counties.

**Variable construction and imputation.— Household characteristics**: All household characteristics included in our demand model other than income were available directly in the LSMS data. Income was set equal to monthly consumption. Summary statistics for households characteristics in the LSMS data are displayed in Table A7.

Table A7: Summary of household characteristics in LSMS and KNBS data

| | | Household | | | | House Characteristics | | | | Assets | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Year | Members | Adults | Income | Education | Pucca | Rooms | Land | Roof | Fan | Mobile | TV | Fridg |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) |
| Ethiopia | 2011 | 4.33 | 2.19 | 4,590 | 4.35 | 0.18 | 1.68 | 0.81 | 0.44 | | 0.24 | 0.02 | 0.01 |
| Ethiopia | 2018 | 5.17 | 2.27 | 8,137 | 4.44 | 0.11 | 1.90 | 0.94 | 0.61 | | | 0.02 | 0.00 |
| Kenya | 2015 | 4.52 | 2.15 | 7,699 | 4.75 | 0.27 | 1.52 | 0.62 | 0.84 | | 0.85 | 0.16 | |
| Kenya | 2019 | 4.45 | 2.24 | 8,162 | 5.29 | 0.31 | 3.38 | 0.62 | 0.88 | | | | |
| Malawi | 2010 | 4.60 | 2.10 | 2,875 | 5.77 | 0.23 | 2.47 | 0.87 | 0.26 | 0.01 | 0.29 | 0.04 | 0.01 |
| Malawi | 2016 | 4.30 | 2.02 | 4,587 | 4.75 | 0.20 | 2.38 | 0.83 | 0.39 | 0.01 | 0.40 | 0.04 | 0.01 |
| Mali | 2014 | 10.97 | 4.78 | 18,129 | 4.27 | 0.18 | 5.49 | 0.77 | 0.38 | 0.05 | 0.76 | 0.18 | 0.03 |
| Mali | 2017 | 11.05 | 4.65 | 18,129 | 4.32 | 0.15 | 5.82 | 0.80 | 0.47 | 0.08 | 0.79 | 0.28 | 0.03 |
| Niger | 2011 | 6.58 | 2.66 | 5,906 | 4.50 | 0.19 | 2.66 | 0.93 | 0.11 | 0.00 | 0.38 | 0.02 | 0.01 |
| Niger | 2014 | 6.65 | 2.66 | 7,575 | 4.46 | 0.21 | 2.79 | 0.85 | 0.17 | 0.01 | 0.56 | 0.02 | 0.01 |
| Nigeria | 2010 | 5.72 | 2.65 | 5,398 | 5.10 | 0.43 | 3.85 | 0.64 | 0.92 | 0.25 | | 0.25 | 0.09 |
| Nigeria | 2018 | 6.00 | 2.67 | 5,377 | 5.57 | 0.45 | 3.72 | 0.79 | 0.88 | 0.30 | 0.70 | 0.33 | 0.10 |
| Tanzania | 2008 | 5.44 | 2.49 | 4,119 | 5.74 | 0.60 | 2.70 | 0.98 | 0.48 | | 0.28 | 0.03 | 0.01 |
| Tanzania | 2019 | 5.04 | 2.45 | 7,490 | 5.79 | 0.78 | 2.87 | 0.77 | 0.81 | | 0.79 | 0.13 | 0.02 |
| Uganda | 2009 | 5.98 | 2.39 | 2,264 | 5.24 | 0.56 | 2.88 | 0.80 | 0.63 | | 0.45 | 0.05 | |
| Uganda | 2019 | 5.25 | 2.33 | 2,266 | 5.26 | 0.64 | 2.05 | 0.78 | 0.72 | | 0.75 | 0.12 | |

The table presents means of the variables after within and cross-country imputations. For LSMS countries, income is equated with consumption. Income in Kenya 2015 was set equal to monthly consumption. For each county in Kenya in 2015, we regressed consumption on earnings. For each county in Kenya 2019, we predicted monthly consumption using the results of the corresponding regression. Income in Kenya 2019 was set equal to predicted monthly consumption. Income is converted to rupees using World Bank's PPP conversion factor. Education is converted to the scale used in the experimental data from years of education.

**Supply side characteristics**: Source characteristics were aggregated at the village equivalent level in the LSMS data following the same procedure that was performed in the experimental dataset. For source availability, we assume that solar is available to all households. We infer that grid or diesel is available to households if any household in their village equivalent has access to grid or diesel respectively. In places where large administrative units (or clusters of administrative units) are used, source availability is likely overestimated. Table A8 summarizes source availability in the LSMS data.

Table A8: Market shares and source availability in LSMS and KNBS data

| | | Choice | | | | Availability | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Year (1) | Grid (2) | Solar (3) | Diesel (4) | None (5) | Grid (6) | Solar (7) | Diesel (8) | N (9) |
| Ethiopia | 2011 | 0.07 | 0.00 | 0.00 | 0.93 | 0.65 | 1.00 | 0.06 | 3,466 |
| Ethiopia | 2018 | 0.09 | 0.30 | 0.00 | 0.60 | 0.50 | 1.00 | 0.01 | 3,029 |
| Kenya | 2015 | 0.17 | 0.23 | 0.01 | 0.59 | 0.97 | 1.00 | 0.67 | 13,087 |
| Kenya | 2019 | 0.27 | 0.29 | 0.00 | 0.44 | 0.99 | 1.00 | 0.07 | 12,598 |
| Malawi | 2010 | 0.02 | 0.01 | 0.01 | 0.96 | 0.73 | 1.00 | 0.22 | 10,038 |
| Malawi | 2016 | 0.03 | 0.07 | 0.00 | 0.90 | 0.82 | 1.00 | 0.19 | 10,113 |
| Mali | 2014 | 0.09 | 0.26 | 0.07 | 0.58 | 0.75 | 1.00 | 0.64 | 1,786 |
| Mali | 2017 | 0.10 | 0.58 | 0.06 | 0.26 | 0.43 | 1.00 | 0.39 | 6,841 |
| Niger | 2011 | 0.03 | 0.00 | 0.02 | 0.95 | 0.32 | 1.00 | 0.64 | 2,430 |
| Niger | 2014 | 0.05 | 0.00 | 0.01 | 0.94 | 0.43 | 1.00 | 0.54 | 2,261 |
| Nigeria | 2010 | 0.35 | 0.00 | 0.07 | 0.58 | 0.97 | 1.00 | 0.97 | 3,351 |
| Nigeria | 2018 | 0.36 | 0.00 | 0.11 | 0.53 | 0.87 | 1.00 | 0.97 | 3,376 |
| Tanzania | 2008 | 0.02 | 0.01 | 0.00 | 0.97 | 0.63 | 1.00 | 0.21 | 2,039 |
| Tanzania | 2019 | 0.13 | 0.38 | 0.00 | 0.48 | 1.00 | 1.00 | 0.00 | 671 |
| Uganda | 2009 | 0.05 | 0.01 | 0.00 | 0.93 | 0.87 | 1.00 | 0.58 | 2,086 |
| Uganda | 2019 | 0.08 | 0.50 | 0.00 | 0.41 | 0.85 | 1.00 | 0.14 | 2,233 |

The table presents household-level means of source choice and availability. Market shares are constructed from data on source choice and asset ownership. Where multiple sources are chosen, grid is assumed to be the primary choice. Availability for grid and solar is inferred using village equivalents.

Grid price is available for both waves of all LSMS countries other than Ethiopia and Tanzania for which it is only available for the later wave and Mali for which it is only available for the earlier wave. If grid price is only available for one wave, but not the other, grid price for the missing wave is set equal to the median grid price of the non-missing wave.

Solar and diesel price is calculated as the monthly amortization of the purchase price of a solar panel or diesel generator respectively using a 7-year amortization window and a 20% interest rate. This is analogous to the treatment of solar prices in the experimental dataset. Solar prices are generally not available for LSMS countries, with the exception of Uganda and Malawi for which it is available for both waves. To infer prices for the other countries, we calculate solar price for each wave in which it is available in the ACCESS and LSMS data, and regress solar price on year. We then use this regression model to predict solar prices in the remaining LSMS countries. For countries that were missing source prices for a single wave, the source price for the missing wave was set equal to the mean source price of the non-missing wave. For countries that were missing source prices for both waves, prices were set equal to the global mean source price.

Hours of supply are not available for most LSMS countries across sources. However, some LSMS survey waves had data on grid blackout frequency. To predict hours of supply using blackouts, we regressed blackout frequency on hours of supply with country-by-year fixed effects using the subset of survey waves that had both variables[32]. We used the coefficient on blackout frequency from this regression to predict hours of supply as a function of blackouts, with the median hours of source availability for all countries and years as the intercept term.

After this procedure, 9 of 14 LSMS waves had predicted hours of grid supply, but most waves were missing hours of solar and diesel supply. We constructed the remaining hours of supply variables for these survey waves as the global median from the other surveys. ACCESS and one of the LSMS survey waves disaggregated hours of supply into peak and off-peak hours. We used the median ratio between peak and off-peak hours in these surveys to disaggregate hours of supply in the other survey-waves. Table A9 summarizes supply side characteristics in the LSMS data.

[32] For grid, countries with both variables for at least one wave were: Ethiopia, India, and Nigeria. For solar, India. For diesel, India and Nigeria.

Table A9: Summary of supply side characteristics in the LSMS and KNBS data

| | | Price (Rs. PPP) | | | Black Out (days) | | | Offpeak Hours | | | Peak Hours | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Year | Grid | Solar | Diesel | Grid | Solar | Diesel | Grid | Solar | Diesel | Grid | Solar | Diesel |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) |
| Ethiopia | 2011 | 49.55 | 170.73 | 443.28 | 10.32 | 0.00 | 0.00 | 9.42 | 7.52 | 3.90 | 3.36 | 2.55 | 2.10 |
| Ethiopia | 2018 | 49.55 | 121.18 | 443.28 | 17.12 | | 12.86 | 7.13 | 7.53 | 3.91 | 2.55 | 2.55 | 2.11 |
| Kenya | 2015 | 259.66 | 142.41 | 363.32 | | | | 9.05 | 7.53 | 3.93 | 3.23 | 2.55 | 2.11 |
| Kenya | 2019 | 259.66 | 114.10 | 363.32 | | | | 9.05 | 7.53 | 3.93 | 3.23 | 2.55 | 2.11 |
| Malawi | 2010 | 342.12 | 73.17 | 87.54 | 16.72 | | | 9.18 | 7.53 | 3.93 | 3.28 | 2.55 | 2.11 |
| Malawi | 2016 | 352.32 | 30.97 | 100.79 | 18.62 | | | 9.17 | 7.53 | 3.93 | 3.27 | 2.55 | 2.11 |
| Mali | 2014 | 422.91 | 149.49 | 339.02 | 10.38 | | | 9.22 | 7.53 | 3.93 | 3.29 | 2.55 | 2.11 |
| Mali | 2017 | 422.91 | 128.26 | 339.02 | | | | 9.05 | 7.53 | 3.93 | 3.23 | 2.55 | 2.11 |
| Niger | 2011 | 234.28 | 170.73 | 39.46 | 20.31 | | | 9.10 | 7.53 | 3.93 | 3.25 | 2.55 | 2.11 |
| Niger | 2014 | 251.03 | 149.49 | 21.49 | 16.68 | | | 9.18 | 7.53 | 3.93 | 3.28 | 2.55 | 2.11 |
| Nigeria | 2010 | 149.87 | 48.11 | 225.66 | 23.59 | | 22.55 | 3.73 | 7.53 | 3.23 | 1.33 | 2.55 | 1.74 |
| Nigeria | 2018 | 218.99 | 121.18 | 226.13 | 19.36 | 25.71 | | 2.76 | 7.51 | 3.47 | 2.44 | 2.55 | 2.51 |
| Tanzania | 2008 | 183.91 | 191.97 | 805.45 | | | | 9.05 | 7.53 | 3.93 | 3.23 | 2.55 | 2.11 |
| Tanzania | 2019 | 183.91 | 114.10 | 805.45 | | | | 9.05 | 7.53 | 3.93 | 3.23 | 2.55 | 2.11 |
| Uganda | 2009 | 598.71 | 257.70 | 647.76 | | | | 13.45 | 7.53 | 3.93 | 4.80 | 2.55 | 2.11 |
| Uganda | 2019 | 352.37 | 56.39 | 983.88 | | | | 14.04 | 7.53 | 3.93 | 5.01 | 2.55 | 2.11 |

The table presents means of the variables after within and cross-country imputations. Grid price is set equal to average monthly grid expenditure. Diesel and solar prices are calculated as the monthly amortization of the purchase price of a diesel generator or solar panel, respectively, using a 7-year amortization at 20% interest rate. For countries where solar prices are not available, solar prices are adjusted for linear decline in solar prices. Where hours of supply were not available, they were predicted using blackout frequency. Blackout frequency is presented as days of blackout in a month.

**Market shares**: As was done in the ACCESS data, we constructed market shares for grid, solar, and diesel from household-level electricity source choice and asset ownership variables. In the LSMS data, it is not generally possible to see if households utilize multiple sources of electricity. Where bundles are observed, we impose the same priority order as in the experimental dataset, in which households are assumed to have chosen the grid if it is part of their chosen bundle. In practice, the only large observed bundles in the LSMS data are grid and diesel in Nigeria (10-13 % of households). Table A8 summarizes the calculated market shares in the LSMS data.

### 1.4 KNBS

The data on Kenya is sourced from two surveys by Kenya National Bureau of Statistics: Kenya Integrated Household Budget Survey in (KIHBS) in 2015 and Kenya Continuous Household Survey Programme (KCHSP) in 2019. Although the 2020 KCHSP survey is available, it is currently missing some key variables, so it has been excluded. Both surveys cover all counties and are designed to be nationally representative.

**Market identification.—**Villages are not provided in these datasets, so we needed to define village-equivalents. Because counties were the only administrative unit supplied and enumeration-area coordinates were not, we used counties directly as village equivalents. Summary statistics for these village equivalents are provided in Table A6.

**Variable construction and imputation.—Household characteristics**: With the exception of income, all household characteristics included in our demand model were available in the KNBS surveys. Income in Kenya 2015 was set equal to monthly consumption. For each county in Kenya in 2015, we regressed consumption on earnings. For each county in Kenya 2019, we predicted monthly consumption using the results of the corresponding regression. Income in Kenya 2019 was set equal to predicted monthly consumption. Summary statistics for these variables are available in Table A7.

**Supply side characteristics**: Source characteristics were aggregated at the village equivalent level in the KNBS data following the same procedure that was performed in the experimental dataset. For source availability, we assume that solar is available to all households. We infer that grid or diesel is available to households if any household in their village equivalent has access to grid or diesel respectively. Counties are large, so source availability may be overestimated in Kenya. Table A8 summarizes source availability in the KNBS data.

Neither grid, solar, nor diesel price is available for Kenya, so we use the global median to impute grid and diesel price. We fill predict solar price using a regression of solar price on year for the countries for which solar prices are available. We are missing blackouts for Kenya, so we use the global median to infer hours of source availability. We compute peak and off-peak hours using the same method as implemented in the LSMS data. Table A9 summarizes supply side characteristics in the KNBS data.

**Market shares**: As was done in the LSMS data, we constructed market shares for grid, solar, and diesel from household-level electricity source choice and asset ownership variables. In the KNBS data, it is possible to see if households utilize multiple sources of electricity. Where bundles are observed, we impose the same priority order as in the experimental dataset, in which households are assumed to have chosen the grid if it is part of their chosen bundle. In practice, there are very few households utilizing multiple sources. Table A8 summarizes calculated market shares for the KNBS data.

# B Appendix: Model

## 2.1 Estimation moments in detail

This section derives the moments used in estimation.

**Market share moments.—** The first set of moments are based upon the electricity source market shares in each village (market) and time period. The model predicts a source market share of

$$\Pr\left(j' \mid \mathbf{z_{it}}, \mathbf{x_{vt}}, \boldsymbol{\beta}, \boldsymbol{\delta}_{vt}\right).$$

This prediction relies on an integral over the distribution of $\boldsymbol{\nu}_i$. We approximate this integral with simulation. Let $\mathscr{I}_{vt}$ be the surveyed households in village $v$ at time $t$. We draw $S$ tuples of independent standard normal draws for $\boldsymbol{\nu}_i$ once for a household and hold them fixed across simulations. These draws are used to form predicted market shares as

$$\mathbb{E}[s_{vtj} \mid \mathbf{z_t}, \mathbf{x_{vt}}, \boldsymbol{\beta}, \boldsymbol{\delta}_{vt}] = \frac{1}{N_{vt}} \sum_{i \in \mathscr{I}_{vt}} \frac{1}{S} \sum_{s} \Pr\left(j' \mid \mathbf{z_{it}}, \mathbf{x_{vt}}, \boldsymbol{\beta}, \boldsymbol{\delta}_{vt}, \boldsymbol{\nu}_{is}\right)$$

The moment for market shares at the village-time-product level is then

$$G^1_{vtj}(\boldsymbol{\beta}) = \tilde{s}_{vtj} - \mathbb{E}[s_{vtj} \mid \mathbf{z_t}, \mathbf{x_{vt}}, \boldsymbol{\beta}, \boldsymbol{\delta}_{vt}]$$

where $\tilde{s}_{vtj}$ are the observed market shares in the data, slightly corrected for sampling error.[33] We write $\boldsymbol{\delta}(\boldsymbol{\beta})$ because we concentrate the $\boldsymbol{\delta}$ out of estimation using the contraction mapping of **?**. The concentration out of unobservable village-time-product fixed effects is necessary to reduce the dimensionality of the estimates. The fixed effects will also allow us to gain insight, within the demand model, into how unobserved characteristics of electricity sources are changing over time.

Given the contraction mapping to find $\boldsymbol{\delta}$ the predicted market shares will match actual market shares exactly for all candidate values of $\boldsymbol{\delta}(\boldsymbol{\beta})$. Therefore $G^1$ is not used as a moment in the GMM estimator below, since it will be satisfied by construction for all candidate $\boldsymbol{\beta}$ parameters.

[33] We use a Laplace correction to adjust market shares if a source is available but not purchased by any household in our survey sample. This correction is needed because the model will always predict a strictly positive, though small, share for a given source, while exact zero shares are observed in finite samples. For a sample of size $n$, this correction replaces observed market shares $s_j$ with $\tilde{s}_j = (ns_j + 1)/(n + J + 1)$, which has the effect of giving small, positive shares to any source with a precise zero share, while slightly deflating the shares of other sources. Since we observe availability on the supply side for the grid, microgrid and diesel, separately from whether any household in our sample used a given source, we do not apply this correction if a source was not available in a village. Instead, we remove that choice from the choice set for that village.

**Covariances of household characteristics and chosen source characteristics.—**The second set of moments is based on covariances between household characteristics and the characteristics of the electricity source they chose. Let $x^{i,1}_{vtk}$ represent the product characteristic $k$ in village $v$ and time $t$ of household $i$'s chosen source $j$. We can form this chosen characteristic for $k = 1, \ldots, K_2$ as

$$x^{i}_{vtk} = \sum_j \mathbf{1}\{y_{it} = j\} x_{jvtk}$$

where $\mathbf{1}\{y_i = j\}$ denotes that $i$ chose source $j$. Let $\boldsymbol{z_i} = (z_{i1} \ldots z_{iT})$ be an $R \times T$ matrix of household characteristics and $\boldsymbol{x^i_{vk}} = (x^i_{v1k} \ldots x^i_{vTk})'$ be a $T \times 1$ vector stacking the source characteristics $k$ chosen by $i$ in each time period.

We form moments at the household level by summing across time periods. The moments for household characteristics $r = 1, \ldots, R$ interacted with product characteristic $k$ are:

$$\underbrace{G_{ik}\left(\tilde{\boldsymbol{\beta}}, \boldsymbol{\delta}_v\right)}_{R\times 1} = z_i \underbrace{\left(\boldsymbol{x}^i_{vk} - \mathbb{E}\left[\boldsymbol{x}^i_{vk} \mid \mathbf{z_i}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\delta}_v\right]\right)}_{T\times 1}$$

This moment gives the product of the household characteristic and the deviation of the source characteristic, for the household's chosen source, from their expected source characteristic in the model.

Expected source characteristics for each household are formed in the model using the simulation draws for each household as

$$\mathbb{E}\left[\boldsymbol{x}^i_{vtk} \mid \mathbf{z_{it}}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\delta}_v\right] \approx \frac{1}{S}\sum_s \sum_j x_{vtjk} \Pr\left(j' \mid \mathbf{z_{it}}, \mathbf{x_{vt}}, \boldsymbol{\beta}, \boldsymbol{\delta}_{vt}, \boldsymbol{\nu}_{is}\right).$$

These expectations are formed for each time period $t$ and source characteristic $k$. The expected source characteristics for each household in each time period are then stacked to plug-in to the moment.

There are $r = 1, \ldots, R$ household characteristics and $k = 1, \ldots, K$ source characteristics. We form a vector of moment conditions

$$\underbrace{G_i^2\left(\tilde{\boldsymbol{\beta}},\boldsymbol{\delta}_v\right)}_{RK_2\times 1} = \begin{bmatrix} G_{i1}^2\left(\tilde{\boldsymbol{\beta}},\boldsymbol{\delta}_v\right) \\ G_{i2}^2\left(\tilde{\boldsymbol{\beta}},\boldsymbol{\delta}_v\right) \\ \vdots \\ G_{iK}^2\left(\tilde{\boldsymbol{\beta}},\boldsymbol{\delta}_v\right) \end{bmatrix} = \begin{bmatrix} G_{i,k=1,r=1}^2 \\ G_{i,k=1,r=2}^2 \\ \vdots \\ G_{i,k=2,r=1}^2 \\ G_{i,k=2,r=2}^2 \\ \vdots \\ G_{i,k=K_2,r=R}^2 \end{bmatrix}.$$

Households are arrayed horizontally and the moment interactions are arrayed vertically, so that the matrix is $(R\times K_2)\times N$.

**Transition matrix between products over time.—**A key question in discrete choice is what variation in the data identifies the coefficients on unobserved tastes. The availability of data on alternate (Berry, Levinsohn and Pakes, 2004) or repeated choices can provide information on unobserved tastes.

To capture this kind of variation we form moments based on the transition matrix between sources across periods. The transition matrix consists of conditional probabilities for the choice of each $j_{t+1}$ given the household chose $j_t$ in period $t$. Let $\boldsymbol{a}_{j_t,j_{t+1}}$ give the matrix of conditional probabilities such that the row entries across each current technology $j_t$ sum to one.

To allow that we may wish to aggregate transition probabilities across groups of goods, let $\ell = 1,\ldots,L$ denote a group of goods. For example, $\ell$ may denote grid electricity, off-grid electricity, and no electricity as different groups of possible choices. Then the transition matrix between groups is

$$\boldsymbol{a}_{\ell_t,\ell_{t+1}} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1L} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{L1} & & & a_{LL} \end{bmatrix}.$$

The transition probabilities in a row sum to one so the last column is linearly dependent on the others. There are two such transition matrices in our data, for $t = 1$ and $t = 2$.

The moment at the household level is the difference between actual household transitions and the transition probabilities predicted for that household.

$$G^3_{it,\ell,\ell'}\left(\tilde{\boldsymbol{\beta}},\boldsymbol{\delta}_v\right) \quad = \quad \mathbf{1}\{\ell_{it}=\ell,\ell_{i,t+1}=\ell'\} - \Pr\left(\ell_{it}=\ell,\ell_{i,t+1}=\ell' \mid \mathbf{z_{it}},\mathbf{x_{vt}},\boldsymbol{\beta},\boldsymbol{\delta}_{vt}\right)$$

where the first term is an indicator function for the household choosing the pair $(\ell_t=\ell,\ell_{t+1}=\ell')$ in the data and the second term is the predicted probability of observing the same transition for that household in the model. The moment vector is formed by stacking the moments for the full set of linearly independent transitions

$$\underbrace{G^3_i\left(\tilde{\boldsymbol{\beta}},\boldsymbol{\delta}_v\right)}_{L(L-1)(T-1)\times 1} = \begin{bmatrix} G^3_{i,t=1,\ell=1,\ell'=1} \\ G^3_{i,t=1,\ell=2,\ell'=1} \\ \vdots \\ G^3_{i,t=1,\ell=L,\ell'=1} \\ \vdots \\ G^3_{i,t=1,\ell=1,\ell'=L-1} \\ G^3_{i,t=1,\ell=2,\ell'=L-1} \\ \vdots \\ G^3_{i,t=1,\ell=L,\ell'=L-1} \\ G^3_{i,t=2,\ell=1,\ell'=1} \\ G^3_{i,t=2,\ell=2,\ell'=1} \\ \vdots \\ G^3_{i,t=2,\ell=L,\ell'=1} \\ \vdots \\ G^3_{i,t=2,\ell=1,\ell'=L-1} \\ G^3_{i,t=2,\ell=2,\ell'=L-1} \\ \vdots \\ G^3_{i,t=2,\ell=L,\ell'=L-1} \end{bmatrix}.$$

The elements of the moment vector are thus drawn from the columns of the transition matrix, omitting the last column. The vector therefore contains $L(L-1)(T-1)=3(3-1)(3-1)=12$ moments, in our example where there are $L=3$ product groups and $T=3$ time periods. We horizontally concatenate the $G^3_i$ vectors to from $G^3$ of dimension $[L(L-1)(T-1)]\times N$ for all households.

The model predictions for transitions are drawn from household-level joint probabilities over source choices in both periods. As above, we approximate the predicted joint probabilities through simulation

$$Pr(\ell_t, \ell'_{t+1}|\mathbf{z}, \mathbf{x}, \beta, j_t) \quad \approx \quad \frac{1}{S}\sum_s \Pr\left(\ell_t \mid \mathbf{z_{it}}, \mathbf{x_{vt}}, \boldsymbol{\beta}, \boldsymbol{\delta}_{vt}, \boldsymbol{\nu}_{is}\right).\Pr\left(\ell'_{t+1} \mid \mathbf{z_{i,t+1}}, \mathbf{x_{v,t+1}}, \boldsymbol{\beta}, \boldsymbol{\delta}_{v,t+1}, \boldsymbol{\nu}_{is}\right).$$

The fact that draws of unobserved tastes ***nu*** are persistent over time within households will induce dependence between household choices over time periods, even conditional on household and product observable characteristics. We expect this variation helps to identify the variance of unobserved taste shocks. In a similar way, Berry, Levinsohn and Pakes (2004) report that information on the second choice vehicle a household would have chosen is critical to their getting precise estimates of random coefficients.

## 2.2 Estimation procedure and objective

First we estimate $\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{\beta}^o, \boldsymbol{\beta}^u\right)$ along with the mean indirect utilities $\boldsymbol{\delta}_{vt}$ using GMM.

1. Begin with an intial guess $\tilde{\boldsymbol{\beta}}_0$ of the parameters. We can start with the values $\boldsymbol{\beta}^o$ equal to the coefficient estimates from a nested logit model with the same specification and $\boldsymbol{\beta}^u = 0$.

2. Minimize the GMM objective function. Let it be called Obj$(\tilde{\boldsymbol{\beta}}, W)$.

   (a) Use $G^1$ and the BLP contraction mapping to solve for $\delta_{vtj}$ such that predicted market shares equal observed shares. **?** shows that

   $$\delta_{jvt} \leftarrow \delta_{jvt} + \log\left(s_{jvt}\right) - \log\left(\Pr\left(j \mid \mathbf{z_{it}}, \mathbf{x_{vt}}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\delta}_{vt}\right)\right) \tag{15}$$

   is a contraction mapping. For any $\tilde{\boldsymbol{\beta}}$, iterating until convergence recovers the unique $\delta_{jvt}$ such that the predicted market shares of product $j$ in village $v$ at time $t$ exactly equal observed shares. Let $\boldsymbol{\delta}(\tilde{\boldsymbol{\beta}})$ be the $(V \times T \times J) \times 1$ vector of such shares.

   (b) Stack the household level moments as

   $$G_i(\tilde{\boldsymbol{\beta}}) = G_i\left(\tilde{\boldsymbol{\beta}}, \boldsymbol{\delta}_v(\tilde{\boldsymbol{\beta}})\right) = \begin{bmatrix} G_i^2\left(\tilde{\boldsymbol{\beta}}, \boldsymbol{\delta}_v(\tilde{\boldsymbol{\beta}})\right) \\ G_i^3\left(\tilde{\boldsymbol{\beta}}, \boldsymbol{\delta}_v(\tilde{\boldsymbol{\beta}})\right) \end{bmatrix}.$$

   of row dimension $M = RK_2 + L(L-1)(T-1) = 7 \cdot 4 + 3 \cdot 2 \cdot 2 = 40$ and column dimension $N$. Let $G(\tilde{\boldsymbol{\beta}})$ be the $M \times 1$ vector with the row means

   $$G(\tilde{\boldsymbol{\beta}}) = \frac{1}{N}\sum_i G_i(\tilde{\boldsymbol{\beta}}).$$

(c) Form the objective as

$$\hat{Q}\left(\tilde{\boldsymbol{\beta}},\boldsymbol{\delta},W\right) = G(\tilde{\boldsymbol{\beta}})'WG(\tilde{\boldsymbol{\beta}} \tag{16}$$

With the initial conformable identity weighting matrix $W = \mathscr{I}_{M\times M}$.

(d) Recover a first estimate of $\tilde{\boldsymbol{\beta}}$

$$\hat{\tilde{\boldsymbol{\beta}}}_1 = \arg\min_{\tilde{\boldsymbol{\beta}}} \mathrm{Obj}\left(\tilde{\boldsymbol{\beta}},\mathscr{I}_{M\times M}\right).$$

3. Find the optimal two-step estimate $\hat{\tilde{\boldsymbol{\beta}}}_2$.

(a) Form a new weighting matrix. Calculate the covariance of the moments

$$\hat{\Omega}_{M\times M} = \frac{1}{N}G\left(\hat{\tilde{\boldsymbol{\beta}}}_1\right)G\left(\hat{\tilde{\boldsymbol{\beta}}}_1\right)'.$$

Form $W_{(2)} = \hat{\Omega}^{-1}$.

(b) Repeat the above minimization with $\mathrm{Obj}(\tilde{\boldsymbol{\beta}},W_{(2)})$ as the objective.

The above minimization will return parameter estimates $\widehat{\tilde{\boldsymbol{\beta}}} = (\hat{\boldsymbol{\beta}}^{\boldsymbol{u}},\hat{\boldsymbol{\beta}}^{\boldsymbol{o}})$ and $\widehat{\boldsymbol{\delta}}\left(\widehat{\tilde{\boldsymbol{\beta}}}\right)$.

# C Appendix: Additional Results

This section presents additional results on demand. Subsection 3.1 reconciles market shares in the raw data, with Laplace correction, and as predicted by our structural model. Subsection 3.2 presents estimates of the first stage from the estimation of the second, linear part of our structural demand model. Subsection 3.4 gives the profiles of households, which are used to calculate marginal effects in the demand model, and shows the heterogeneity of the estimated marginal effects by household profile. Subsection 3.5 provides additional estimates to check the robustness of the structural demand estimates to alternative nest structures in the nested logit model.

## 3.1 Market shares: model versus data

Table C1 presents the fit of market shares in the model to the data by survey wave and electricity source. In principle the model can fit the data exactly, since village-source-wave specific mean indirect utility terms are free parameters. The fit is very close, but not exact, for two reasons. First, the raw data contain zero market shares for some sources that were available in a given village and wave. For example, we take own solar to be universally available and yet there are some villages where no household said they use own solar. These zeros are not surprising in a sample of 30-odd households, but in the model, all sources must have positive shares, though they can be arbitrarily small. To force the data to have positive shares, we implement a Laplace correction (see footnote 33), which raises market shares slightly for sources with low take-up (Table C1, panels A through C, row 2 versus row 1). Second, we classify availability for some sources based on our supply-side data on village-source-level availability, rather than the survey data on household reports. This classification allows us to observe when a source is not offered (as opposed to not bought), and therefore remove the choice from the choice set instead of modeling it as available but not selected. However, in a small number of cases households report buying sources that we do not believe were offered in their village and survey wave, which we attribute to survey misreports. Again, these differences in classification have a very small affect on market shares (Table C1, panels A through C, row 3 versus row 2).

## 3.2 First stage estimates for structural demand model

Table C2 presents the first-stage from the linear, instrumental variables estimates of the second part of the structural demand model. The endogenous variables are either price, peak hours of supply, or off-

Table C1: Structural Model Fit versus Data

| | Market shares | | | | |
|---|---|---|---|---|---|
| | Grid (1) | Diesel (2) | Own solar (3) | Microgrid (4) | All (5) |
| *Panel A. Baseline* | | | | | |
| Raw data | 5 | 17 | 5 | 1 | 27 |
| Data with Laplace correction | 6 | 17 | 7 | 1 | 31 |
| Model | 6 | 17 | 7 | 0 | 30 |
| *Panel B. Endline* | | | | | |
| Raw data | 25 | 3 | 4 | 9 | 40 |
| Data with Laplace correction | 24 | 3 | 7 | 10 | 43 |
| Model | 24 | 3 | 7 | 9 | 43 |
| *Panel C. Follow-up* | | | | | |
| Raw data | 43 | 3 | 15 | 3 | 64 |
| Data with Laplace correction | 40 | 3 | 17 | 5 | 65 |
| Model | 41 | 3 | 17 | 4 | 65 |

The table presents market shares in the electricity market, and juxtaposes data vs our model's fit. Data with Laplace correction adjusts each product's market share to ensure that no product has a zero share. Small differences between data with Laplace correction and model for a given wave can exist due to the use of market-level source availability in the model. Data with Laplace correction uses actual household-level availability, and there can be inconsistencies between household-level and market-level availability in the data due to a very small number of households in the control villages saying that they used microgrid solar.

peak hours of supply. In columns 1 through 4 the instruments for price are the interactions between the experimental treatment assignments and the endline survey waves. Column 1 gives the first stage for price when instrumenting only for price. Column 2 gives the first stage for price when instrumenting for price, peak hours of supply, and off-peak hours of supply. Columns 3 and 4 give the respective first stage estimates for peak and off-peak hours of supply. Columns 5 and 6 give the first stage estimates of the price equation, when instrumenting for price and both hours measures, and replacing the experimental instruments with instrumental variables constructed along the lines of Berry, Levinsohn and Pakes (1995) and Hausman (1996). We have two sets of alternative instruments for source-village-wave prices. First, the average hours of supply and load from the other products in the same village, which should affect source mark-ups and prices under oligopolistic competition (Berry, Levinsohn and Pakes, 1995). Second, the average price for a given source in the nearest three villages where that source is available, which will covary with source price due to common supply shocks (Hausman, 1996; Nevo, 2001).

Table C2: First-Stage of Linear Estimation of Demand for Electricity

| | Price IV | Price & hours IV | | | Price & hours IV BLP | Price & hours IV Hausman |
|---|---|---|---|---|---|---|
| | Price (1) | Price (2a) | Peak hours (2b) | Off-peak hours (2c) | Price (3) | Price (4) |
| Treatment normal price X Endline | 0.064** | 0.064** | 0.0050 | -0.0046 | | |
| | (0.029) | (0.028) | (0.0051) | (0.030) | | |
| Treatment subsidy price X Endline | -0.16*** | -0.16*** | 0.0075 | 0.014 | | |
| | (0.021) | (0.021) | (0.0063) | (0.031) | | |
| Hours of peak supply | -0.050 | | | | | |
| | (0.049) | | | | | |
| Hours of off-peak supply | 0.0081 | | | | | |
| | (0.013) | | | | | |
| Peak hours instrument | | -0.032 | 0.94*** | 0.19 | -0.040 | -0.038 |
| | | (0.045) | (0.063) | (0.15) | (0.043) | (0.044) |
| Off-peak hours instrument | | 0.0040 | 0.032** | 0.88*** | 0.0057 | 0.0051 |
| | | (0.0094) | (0.013) | (0.030) | (0.0090) | (0.0091) |
| Total supply of competing products | | | | | -0.0017 | |
| | | | | | (0.0060) | |
| Load of competing products | | | | | 0.0079 | |
| | | | | | (0.011) | |
| Avg price in nearby villages | | | | | | 0.038 |
| | | | | | | (0.080) |
| $\xi_{tj}$ mean effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 999 | 999 | 999 | 999 | 945 | 989 |
| First-stage $F$-Stat | 42.1 | 21.1 | 524.1 | 1057.2 | 0.4 | 0.5 |
| Control mean | 0.95 | 0.95 | 4.09 | 3.07 | 0.95 | 0.95 |

This table presents the first-stage of the IV estimates provided in column 2 through 5 of Table 4. Each outcome variable is an endogenous variable that we instrument for in the IV estimations. The second cluster of columns correspond to our preferred IV specification, which uses our experiment to instrument for price, peak, and off-peak hours of supply. Details on instrument construction for hours of supply can be found in Appendix A, Subsection . Standard errors are clustered at the village-level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### 3.3 Elasticities of substitution

Table **??** presents aggregate own- and cross-price elasticities by source, using the structural demand model. The elasticities are calculated using the full model of demand, using the specification for mean indirect utility in Table 4, column 3 and the specification for household heterogeneity in Table 5, columns 6 through 10. Each entry in the table gives the elasticity of the market share for the column source with respect to the price of the row source.

Table C3: Price Elasticities of Electricity Source Demand by Instrument Set

| | **RCT IV** | | | | | **BLP IV** | | | | | **Hausman IV** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Elasticity of share of source:* | Grid (1) | Diesel (2) | Own solar (3) | Micro-grid (4) | None (5) | Grid (6) | Diesel (7) | Own solar (8) | Micro-grid (9) | None (10) | Grid (11) | Diesel (12) | Own solar (13) | Micro-grid (14) | None (15) |
| *w.r.t price of* | | | | | | | | | | | | | | | |
| *Panel A. Endline Prices* | | | | | | | | | | | | | | | |
| Grid | -0.20 | 0.05 | 0.05 | 0.05 | 0.06 | 0.43 | -0.14 | -0.17 | -0.19 | -0.11 | 0.63 | -0.20 | -0.24 | -0.26 | -0.17 |
| Diesel | 0.01 | -1.68 | 0.09 | 0.08 | 0.06 | -0.05 | 5.86 | -0.37 | -0.34 | -0.18 | -0.09 | 8.78 | -0.53 | -0.48 | -0.27 |
| Own solar | 0.02 | 0.23 | -2.14 | 0.30 | 0.16 | -0.11 | -1.41 | 10.64 | -1.65 | -0.76 | -0.18 | -2.17 | 16.51 | -2.47 | -1.19 |
| Microgrid | 0.05 | 0.37 | 0.79 | -1.95 | 0.21 | -0.30 | -2.17 | -3.24 | 9.00 | -0.94 | -0.47 | -3.12 | -4.35 | 13.32 | -1.43 |
| *Panel B. Break-even Grid Price* | | | | | | | | | | | | | | | |
| Grid | -0.82 | 0.18 | 0.14 | 0.15 | 0.10 | 0.47 | -0.16 | -0.32 | -0.41 | -0.40 | 0.33 | -0.03 | -0.20 | -0.22 | -0.40 |
| Diesel | 0.02 | -1.73 | 0.08 | 0.08 | 0.06 | -0.01 | 4.62 | -0.38 | -0.27 | -0.13 | -0.00 | 6.44 | -0.56 | -0.42 | -0.21 |
| Own solar | 0.03 | 0.21 | -2.17 | 0.31 | 0.17 | -0.03 | -1.55 | 10.48 | -1.73 | -0.72 | -0.02 | -2.37 | 15.81 | -2.69 | -1.17 |
| Microgrid | 0.07 | 0.36 | 0.80 | -1.99 | 0.22 | -0.06 | -1.34 | -3.06 | 8.22 | -0.77 | -0.04 | -1.97 | -4.16 | 11.62 | -1.17 |

The table presents aggregate own- and cross-price elasticities of demand by electricity source. In panel A, the elasticities are calculated using a 10% increase in each source's price from its mean endline price. Panel B is similar but fixes the price of grid at the break-even price for the utility (calculated as 233 INR = 60 KwH per month mean consumption $\times$3.88 INR per KwH average cost of procurement). Columns 1 through 5 use IV estimates of mean utility from source characteristics based on the RCT corresponding to Column 3 in Table 4. Columns 6 through 10 use IV estimates of mean utility that rely on average characteristics of other products available in a given market as instruments (Berry, Levinsohn and Pakes (1995)) corresponding to Column 4 in Table 4. Columns 11 through 15 use IV estimates that rely on the average price of each product in the nearest three villages as instrument for its price in a given village (Nevo (2001); Hausman (1996)) corresponding to Column 5 in Table 4. The elasticities are calculated for the market share of each column source with respect to the price of each row source.

### 3.4 Marginal effects for alternative household profiles

Table 5 presents the marginal effects of household characteristics on electricity choice probabilities for a "median" household. Section 5 shows the results of counterfactuals where we increase the income and wealth of households from "poor" to "median" and "rich" levels. This subsection defines these household profiles and shows marginal effects for alternative household profiles to complement the estimates in the main text. BM>NR: Just heads up, marginal effects are evaluated at endline, whine the counterfactuals are estimated at the follow-up.

Appendix Table C4 shows the characteristics of households that are used to create the three profiles of household covariates. The number of adults (column 1) is integer valued, house characteristics (2 and 3) are indicator variables, the number of rooms is integer valued (4), agricultural land ownership is an indicator variable (5), literacy is integer valued (6) and income is continuous. Each row gives the values that these variables take on for each of the three household profiles we use to calculate marginal effects and to run counterfactuals.

The levels of these variables were chosen in order to roughly place a household, on an univariate basis, at the 20th, 50th and 80th percentile of the income or wealth distributions. Table C6 shows detailed summary statistics for the household covariates that enter our demand model in order to place the household profiles in context.

To calculate the marginal effects of these covariates on choice probabilities, we change their values by either one unit, for dummy variables, or one standard deviation, for integer valued and continuous variables. Table C5 shows the changes that this entails for each household covariate that enters the profiles. For the binary variables (Pukka, roof, land), this approach necessarily means that we cannot calculate the discrete effect for these variables when they are already equal to one in a given profile. For example, we cannot calculate the impact of having a roof for a median household, as a median household already has a roof. We therefore omit these entries from the corresponding tables of marginal effects.

Table C6: Summary Statistics of Household Characteristics

| | Mean | SD | Min | Q1 | Median | Q3 | Max | Obs |
|---|---|---|---|---|---|---|---|---|
| Adults in the household | 3.79 | 1.88 | 1 | 2 | 3 | 5 | 15 | 2917 |
| Indicator for solid roof | .53 | .5 | 0 | 0 | 1 | 1 | 1 | 2917 |
| Indicator for agricultural land | .6 | .49 | 0 | 0 | 1 | 1 | 1 | 2917 |
| Education of household head (1–8) | 2.44 | 2.01 | 1 | 1 | 1 | 4 | 8 | 2917 |
| Monthly household income (INR) | 768.85 | 625.93 | 0 | 425 | 600 | 875 | 6500 | 2917 |

The table summarizes each household covariates used in out structural estimation. Each observation is for a household at the endline.

Table C4: Profile Details

| | Adults (1) | Roof (2) | Land (3) | Education (4) | Income (INR) (5) |
|---|---|---|---|---|---|
| Poor | 3 | 0 | 0 | 1 | 5000 |
| Median | 3 | 1 | 1 | 1 | 6000 |
| Rich | 4 | 1 | 1 | 3 | 7500 |

This table details the characteristics for a poor, median, and rich household. Each profile was constructed by independently taking a fixed percentile of each column attribute at the endline. The fixed percentiles corresponding to poor, median, and rich are 33, 50, and 66, respectively. For example, a poor household has three adults, which corresponds to the 33rd percentile of households in our sample at the endline.

Table C5: Definition of Household Characteristics and Magnitude of Marginal Change

| Characteristic | Definition | Marginal Change (Poor) |
|---|---|---|
| Adults | Adults in the household | 1 SD (1.88 persons) |
| Roof | Indicator for solid roof | 0 to 1 |
| Land | Indicator for agricultural land | 0 to 1 |
| Education | Education of household head (1-8) | 1 SD (2 levels) |
| Income | Monthly household income | 1 SD (INR 6259) |

The table defines the household characteristics used in our choice model and shows the magnitude of the change in each covariate for a poor household, as used in the marginal impact analysis of household covariates on choice probabilities (Table **??**). Base profiles for a representative poor, median, and rich household can be found in Table C4. Education 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.

Tables C7 and C8 show the estimated discrete effects for a poor and rich household, respectively, to be compared to Table 5 in the main text. The main finding is that, at all levels of household income, the discrete effects of increasing income or wealth proxies is to increase the demand for grid electricity and decrease, or barely alter, the demand for other sources of electricity. The discrete effects of household characteristics on choice probabilities are slightly smaller for rich than for poor households on some measures (e.g., the effect of income on grid choice), though these differences are small and not generally statistically significant. This relative lack of attenuation may reflect that even rich households, in our sample, have far from complete take-up of any electricity source.

Table C7: Electricity Source Choice Probabilities by Household Characteristics - Poor Households

| | Simple Model | | | | | Full Model | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Grid | Diesel | Own Solar | Micro-grid | None | Grid | Diesel | Own Solar | Micro-grid | None |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| *Panel A. Marginal effects of household characteristics* | | | | | | | | | | |
| Number of adults | 0.121 | 0.028 | -0.013 | 0.003 | -0.138 | 0.051 | 0.020 | -0.003 | 0.014 | -0.082 |
| | (0.028) | (0.006) | (0.005) | (0.010) | (0.019) | (0.042) | (0.006) | (0.008) | (0.011) | (0.033) |
| Solid roof (=1) | 0.226 | 0.013 | -0.001 | -0.020 | -0.218 | 0.092 | 0.010 | 0.014 | -0.001 | -0.115 |
| | (0.017) | (0.007) | (0.008) | (0.012) | (0.016) | (0.020) | (0.007) | (0.010) | (0.012) | (0.018) |
| Owns ag. land (=1) | | | | | | 0.011 | 0.004 | -0.006 | 0.023 | -0.032 |
| | | | | | | (0.014) | (0.006) | (0.008) | (0.013) | (0.016) |
| Education of household head (1-8) | | | | | | 0.006 | 0.009 | -0.008 | 0.008 | -0.015 |
| | | | | | | (0.006) | (0.004) | (0.005) | (0.007) | (0.009) |
| Household income | | | | | | 0.002 | 0.002 | 0.005 | 0.036 | -0.045 |
| | | | | | | (0.007) | (0.003) | (0.007) | (0.011) | (0.011) |

The table shows the effects of household characteristics on the probability of a household choosing a given electricity source. The table reports the results of two models. A simple model, reported in columns 1 through 5, includes as covariates the number of adults in the household and a dummy variable for whether the household has a solid roof. Our full model, reported in columns 6 through 10, includes three additional observable proxies for household demand: household income, whether the household owns agricultural land, and years of education of the household head. Both models also include a triangularly-distributed random coefficient on price, and the estimated dispersion parameter of the price distribution (corresponding to half the range of the distribution) is reported in Panel B. The effects of household characteristics are nonlinear. The table therefore reports "marginal" effects evaluated for a "poor" household, lacking the binary indicators of wealth and with an income at the 33rd percentile, facing the end-line availabilities, qualities, and prices of each good. The profile of a poor household is defined as a household of three adults living in a house without a solid roof and lacking agricultural land ownership. See Appendix Table C4 for the characteristics of a poor household. The marginal effects are not truly marginal; for binary variables, we report the effect on choice probability of changing the value from zero to one, and for continuous variables the effect of a one standard deviation increase in that variable. We also report the minimized objective value for both models. The objective function is different between the two models. First, the objective functions use different weighting matrices derived from the 2-step "optimal" GMM estimator. Second, the full model objective function contains an additional moment for each additional parameter estimated. Therefore, the magnitude of the objective functions reported is not directly comparable. Finally, we report a Wald test statistic, distributed chi-squared with 12 degrees of freedom, from a test of the restriction that the coefficients on the covariates added in the full model are jointly zero.

Table C8: Electricity Source Choice Probabilities by Household Characteristics - Rich Households

| | Simple Model | | | | | Full Model | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Grid | Diesel | Own Solar | Micro-grid | None | Grid | Diesel | Own Solar | Micro-grid | None |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| *Panel A. Marginal effects of household characteristics* | | | | | | | | | | |
| Number of adults | 0.163 | 0.012 | -0.012 | -0.011 | -0.153 | 0.119 | 0.010 | -0.008 | -0.013 | -0.108 |
| | (0.019) | (0.006) | (0.004) | (0.009) | (0.015) | (0.027) | (0.007) | (0.005) | (0.014) | (0.019) |
| Solid roof (=1) | 0.185 | 0.003 | 0.002 | -0.020 | -0.171 | 0.138 | 0.001 | 0.005 | -0.024 | -0.120 |
| | (0.021) | (0.009) | (0.008) | (0.012) | (0.018) | (0.035) | (0.010) | (0.009) | (0.012) | (0.025) |
| Owns ag. land (=1) | | | | | | 0.090 | -0.005 | -0.010 | 0.000 | -0.075 |
| | | | | | | (0.043) | (0.009) | (0.008) | (0.016) | (0.032) |
| Education of household head (1-8) | | | | | | 0.071 | 0.001 | -0.010 | -0.005 | -0.058 |
| | | | | | | (0.025) | (0.006) | (0.005) | (0.009) | (0.019) |
| Household income | | | | | | 0.034 | -0.002 | 0.000 | 0.017 | -0.050 |
| | | | | | | (0.025) | (0.008) | (0.006) | (0.011) | (0.018) |

The table shows the effects of household characteristics on the probability of a household choosing a given electricity source. The table reports the results of two models. A simple model, reported in columns 1 through 5, includes as covariates the number of adults in the household and a dummy variable for whether the household has a solid roof. Our full model, reported in columns 6 through 10, includes three additional observable proxies for household demand: household income, whether the household owns agricultural land, and years of education of the household head. Both models also include a triangularly-distributed random coefficient on price, and the estimated dispersion parameter of the price distribution (corresponding to half the range of the distribution) is reported in Panel B. The effects of household characteristics are nonlinear. The table therefore reports "marginal" effects evaluated for a "rich" household facing the end-line availabilities, qualities, and prices of each good. The profile of a rich household is defined as a household of four adults living in a house with a solid roof, owning agricultural land, and income at the 66th percentile. See Appendix Table C4 for the characteristics of a median household. The marginal effects are not truly marginal; for binary variables, we report the effect on choice probability of changing the value from one to zero, and for continuous variables the effect of a one standard deviation increase in that variable. We also report the minimized objective value for both models. The objective function is different between the two models. First, the objective functions use different weighting matrices derived from the 2-step "optimal" GMM estimator. Second, the full model objective function contains an additional moment for each additional parameter estimated. Therefore, the magnitude of the objective functions reported is not directly comparable. Finally, we report a Wald test statistic, distributed chi-squared with 12 degrees of freedom, from a test of the restriction that the coefficients on the covariates added in the full model are jointly zero.

### 3.5 Robustness of demand estimates

Table C9 shows estimates of the intent-to-treat effects of the experimental treatment assignments on microgrid demand using administrative data on microgrid payments. These estimates are analogous to the Table 3 estimates in the main text but use administrative data on payments rather than survey data on source usage as the measure of demand. The estimated market share in subsidized price villages is very similar across both data sources, while the estimated market share in normal price villages is higher in the survey data than in the payments data. Payments for microgrids may differ from survey reports due to measurement error or because households still use microgrids, for a time, even after they have stopped paying the monthly price. We understood from our field work that the pace at which HPS repossessed systems for non-payment was slow.

Table C9: Solar Microgrid Demand by Village Treatment Arm

| | Administrative | | |
|---|---|---|---|
| | Baseline | Endline | Paid ever |
| | (1) | (2) | (3) |
| Treatment: Subsidized price | 0.033 | $0.179^{***}$ | $0.271^{***}$ |
| | (0.025) | (0.052) | (0.066) |
| Treatment: Normal price | 0.003 | 0.013 | 0.022 |
| | (0.002) | (0.010) | (0.034) |
| Constant | 0.000 | 0.005 | 0.030 |
| | (0.000) | (0.005) | (0.029) |
| Observations | 100 | 100 | 100 |

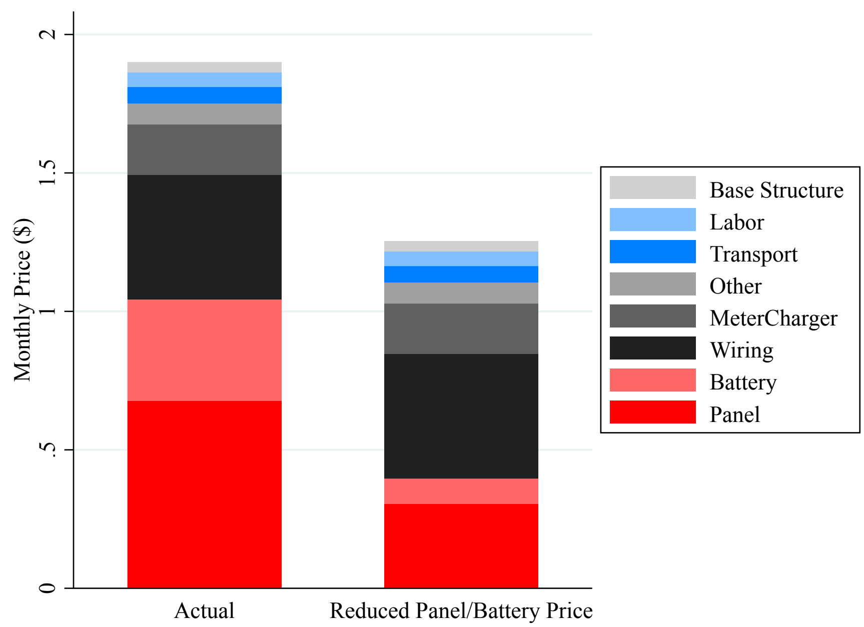
The table shows estimates of microgrid demand by treatment status. The dependent variable is the village-level market share of microgrid solar from HPS administrative payments data, which measures whether households have paid for the source recently. There are three treatment arms: a subsidized price arm (microgrids offered at INR 100), a normal price arm (microgrids offered at the prevailing price of INR 200, later cut to INR 160 in some villages), and a control arm (microgrids not offered). Each column measures market share for a specific time frame: the household paid in the first month after baseline; the household paid in the three months leading up to the endline; the household ever paid. Standard errors are shown in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

**BM:** We could add RCL with heterogeneity by price here. The table would replace estimates for Alternate Nests specification for NL.

# D Appendix: Counterfactual Scenarios

This section gives additional details on our counterfactual scenarios. Figure D1 shows the breakdown of costs for solar microgrids, which we use to forecast the effects of declines in the prices of solar photovoltaic panels and batteries on solar systems. Because the capital cost of PV panels and batteries make up a significant, but incomplete, share of total cost, this breakdown is necessary to calculate the effect that any proportional decline in capital costs will have on the total costs of solar systems.

Figure D1: Microgrid Solar Cost Structure: Current and Predicted




In this figure, we show the cost components of a microgrid, to provide transparency on how we derived the solar prices in our counterfactual scenarios involving a fall in solar prices. We only take into account price changes for solar photovoltaics and batteries, which are clearly correlated with R&D. We assume a 55% reduction in the cost of solar PV, which is in line with the National Renewable Energy Laboratory's projections for 2022. For batteries, we assume a cost reduction of 75%, in accordance with the US Department of Energy's 2022 goal. These two changes translate into a 30% reduction in the overall price of a microgrid. We use the same proportional change in price for own solar in our counterfactuals.

For reference, Table D1 enumerates the assumptions in our counterfactual scenarios from Table **??**. Column 1 gives the name of each scenario and columns 2 through 4 detail assumptions made in the row scenario regarding source availability, supply hours, pricing and subsidies, and any additional details.

Table D1: Counterfactual Analysis: Assumptions

| Scenario | Source availability | Source hours (peak) | Other notes |
|---|---|---|---|
| Improved solar | Follow-up | Follow-up | Solar technologies at their market characteristics as of the follow-up survey, with consumer characteristics held constant at baseline levels |
| Improved grid | Follow-up | Follow-up | Grid technology at its market characteristics as of the follow-up survey, with consumer characteristics held constant at baseline levels |
| 50% Solar Cost Reduction | Follow-up | Follow-up | Reduction in microgrid price from INR 170 to INR 85, proportional (50%) reduction in own solar price |
| Grid in all villages | Follow-up for diesel and solar, grid everywhere | Follow-up | |
| Increase peak grid hours | Follow-up | Two additional peak hours for grid (capped at 5 hours), follow-up peak hours for all other sources | |
| All households at least $X$ | Follow-up | Follow-up | Each household covariate is at least as large as it is under profile $X$ where $X \subset \{\text{Median}, \text{Rich}\}$. Profile *Rich* corresponds to the 66th percentile (details on each profile can be found in Appendix Table **??**) |
| Price grid at cost | Follow-up | Follow-up | Increase grid price from INR 60 to INR 233 to break even |