The Global Electrification Frontier and Climate Change*

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

March 12, 2025

Abstract

Off-grid solar promises a low-cost and carbon-free path to electrification. But will poor households choose off-grid power? We run a multi-year pricing experiment in rural India to estimate demand over all electricity sources, including off-grid solar, diesel generators and the grid. We find that off-grid solar is an important stop-gap, but households value grid electrification 6.6 times more. The grid, however, barely increases global surplus, because grid carbon damages nearly offset households' gains. We apply our model to data from Africa and find a similarly strong preference for the grid among households there, underscoring the external validity of our results.

*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, Marsya Candradewi, Catherine Che, Aliya Jabbar, Ananya Kotia, Bogdan Mukhametkaliev, Rakesh Pandey, Aditya Petwal, Victor Quintas-Martinez, Johanna Rayl, Rashi Sabherwal, and Jared Stolove 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].

The global electrification frontier is the collection of places in the world where households are getting electricity for the first time. The 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. Today, billions of people, mainly in South Asia and sub-Saharan Africa, still lack electricity or consume paltry amounts. Most of the power they do use comes from burning fossil fuels; therefore, expanding the grid and increasing supply along the frontier will increase carbon emissions and climate change damages.[1]

Off-grid solar is often held up as a solution to the dual challenges of electrification and climate, because it can expand electricity access *and* decarbonize its supply in one stroke. Former UN Secretary General Ban Ki-moon said "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," a belief that is widely held.[2] After steep cost declines in the last decade, off-grid solar is showing signs of realizing its potential. Figure 1 reports off-grid solar's market share for rural populations in a group of African countries and Indian states. Off-grid solar has grown from nothing, as recently as 2010, to now powering 30% to 60% of rural homes in parts of sub-Saharan Africa. Notably, solar take-up has also risen in rural India, but not to the same heights.

This paper studies how households choose between the grid and off-grid alternatives like solar. We estimate demand over *all sources* of electricity using experimental variation in prices that lasted for two-and-a-half years and accounting for supply-side variation in the availability of sources and their carbon intensity. While off-grid solar is carbon-free, it is unknown how much rural households themselves value this new technology. On the frontier, electricity is not homogenous, and off-grid solar, the grid and other sources differ widely in price, load, intermittency, reliability and other characteristics. Moreover, the trade-offs between these sources are changing rapidly, due to technological progress and income growth over time.

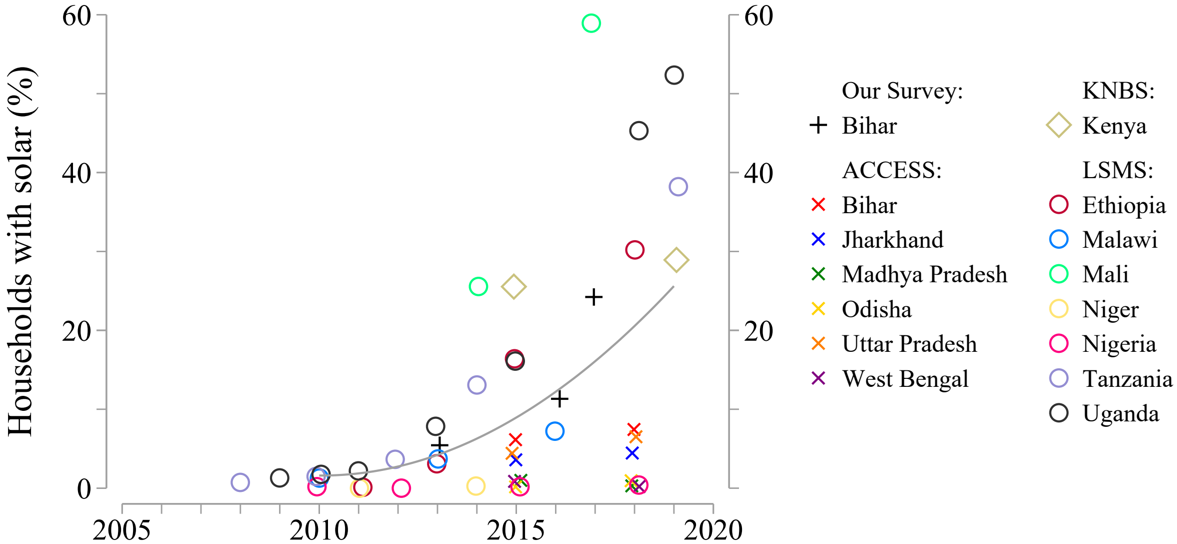
Previous research has estimated the demand for one source of power at a time, with an unspecified outside good that varies across contexts.[3] This approach cannot answer urgent policy

[1] Fossil fuels account for about 70% of electricity supply in Asia and 75% in Africa (Ember, 2024).

[2] "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 particular. In 2023, the International Energy Agency stated "Off-grid connections play a central role in bringing about universal access by the end of this decade" (IEA, 2023). The flagship 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. Notwithstanding this wave of investment, there were more people without electricity in Africa in 2023 than in 2010 (IEA, 2023).

[3] Our study contributes to a small literature on the demand for electrification. In contemporaneous experiments,

Figure 1: Household solar adoption in select developing countries and Indian states



The figure shows the growth of solar power in rural areas of select Indian states and African countries. The data collectively represent a population of roughly 1 billion people. The figure uses four sources of data. First, the "+" markers present 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 Integrated 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.

questions like whether households consider off-grid solar an adequate substitute for the grid or what are the emissions caused by grid extension, because it does not explicitly account for how households substitute between different technologies. A large related literature has measured the impacts of access to electricity on economic outcomes and well-being, but has not estimated household demand, and so cannot speak to household adoption of and willingness-to-pay for alternative policies and technologies.[4]

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. A couple of papers have hypothesized that solar and grid electricity are imperfect substitutes for the rural poor. Fowlie et al. (2019) argue 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 survey in Kenya showing that grid users own more high-load appliances than solar users. In the United States, where all households are on the grid, recent research on electricity demand has focused on improving estimates of the intensive margin elasticity of power demand with respect to price (Ito, 2014; Deryugina, MacKay and Reif, 2020).

[4]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.

The study is conducted in Bihar, an Indian state of 123 million people and a typical outpost on the global electrification frontier. Between 2000 and 2016, India dominated world electrification, contributing over 80% of the total gain in the number of households connected to the grid (International Energy Agency, 2017). We capture the transformation of the electricity sector in Bihar with an original, three-wave household panel survey spanning 2013 to 2017. The survey reveals a competitive electricity market in which each of diesel generators, solar panels and the grid are the main source of electricity in different villages at different points in time. The overall electrification rate increased by an astounding 40 percentage points during the four years covered by our sample, due to both rapid grid extension and household adoption of off-grid solar. Overall, our sample in rural Bihar is drawn from an informative time and place to learn about the global electrification frontier and its connections to climate change.

We estimate demand over *all* sources of electricity in rural Bihar using a discrete-choice demand model and variation in off-grid solar prices from a randomized experiment. In the model, 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. The model allows for rich heterogeneity in household characteristics and source characteristics in 100 village-level electricity markets over a four-year period. We identify the price elasticity of demand using experimental variation in the price and availability of off-grid solar microgrids supplied by Husk Power Systems, an international company that partnered in this research. Our demand estimates are rare for combining a structural demand system with randomized variation from an experiment.[5] The experimental price variation lasted two-and-a-half years, which mitigates external validity concerns that arise from the short-run variation in many experiments. The model estimates allow for quantification of the household surplus from electrification, attribution to different sources, and counterfactual examination of how changes in technology and policy affect household choices and

[5]We join a methodological movement in the development literature that combines structural models with experimental variation to aid interpretation and increase the external validity of experimental results (Todd and Wolpin, 2006; Attanasio, Meghir and Santiago, 2012; Duflo, Hanna and Ryan, 2012; Bryan, Chowdhury and Mobarak, 2014; Duflo et al., 2018). Experimental estimates of demand have been an area of growth in development economics and are used both to test theories of behavior and to consider optimal policy (Berry, Fischer and Guiteras, 2020; Peletz et al., 2017; Dupas, 2014; Karlan and Zinman, 2019). Though many products have close substitutes, few prior studies that experimentally estimate demand explicitly model substitution between competing products. 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. With respect to the industrial organization literature, the use of an experiment to estimate a discrete choice demand model removes the need to rely on traditional assumptions, about conduct or demand, to form instrumental variables (Berry, Levinsohn and Pakes, 1995; Hausman, 1996).

the greenhouse gas emissions that cause climate change.

The analysis leads to three main findings. First, households value off-grid solar, but the fossil-powered grid dominates once it reaches a village. Off-grid solar's value comes from serving as a stop-gap source of electricity, when the grid is not present, or for the most price-sensitive households, rather than from leap-frogging the grid. For example, the own-price elasticities of demand for off-grid sources (i.e., household solar, solar microgrids, diesel generators) are large, at around -2, indicating that households do not gain much surplus from these sources and readily substitute away from them in response to even modest price increases, while demand for the grid is relatively inelastic. The grid's greater contribution to household surplus is due to both demand factors (e.g., the grid allows households to run appliances like fans and televisions that have higher loads than off-grid sources can support) and supply factors (e.g., when available, grid electricity supply is heavily subsidized).

Second, increases in electrification during the sample affect household, producer and social surplus very differently. Using the model, we estimate that the observed increase in electrification, due to falling solar prices and grid expansion, raised household surplus from electrification by about 5 times, with most of this gain due to the grid alone. Households value the improvements in the grid 6.6 times more than the advances in off-grid solar within our sample period. However, the state loses money on grid expansion, because grid electricity is heavily subsidized. Thus, the gain in *Bihar*'s social surplus (= household + producer) from the increase in overall electrification is half as large as the gain in household surplus. The gain in *global* social surplus (= household + producer + climate damages) is even smaller, at only 20% of the gain in household surplus, after accounting for the increase in climate damages using the US EPA's estimated social cost of carbon (Environmental Protection Agency, 2023).[6]

If the grid is so dominant in Bihar, why is off-grid solar surging in Africa (Figure 1)? The large gap in solar adoption between rural India and rural sub-Saharan Africa motivates us to explore the Bihar findings' external validity. We assemble household data, representative of a rural population of 500 million people, for eight African countries that are also on the global electrification frontier. We then apply a re-fitted version of our demand model to these data to reconcile differences in household choices across countries.

[6]The *aggregate* climate impacts of grid extension in Bihar remain small, because electricity consumption is very low, even among those connected to the grid. As a point of comparison, per capita electricity sector CO_2 emissions are about 5 metric tons in the US (calculated using US electricity power emission and population data from U.S. Energy Information Administration (2024)), nearly 90 times baseline per capita emissions in this sample of rural Biharis.

Our third finding is that the substantially higher rates of solar adoption in sub-Saharan Africa, compared to Bihar, are explained by policy-induced differences in supply, not differences in household demand (i.e., preferences). The high take-up of off-grid solar in rural Africa is due to the grid's limited extent and high prices there, relative to Bihar. We find that the grid would dominate choices and surplus in these eight sub-Saharan African countries if households there faced the Bihar grid, with its wide availability and heavily subsidized energy prices. More broadly, this out-of-sample application shows how our model can concurrently explain both skyrocketing solar shares in Africa and the low willingness-to-pay to connect to grid electricity found, for example, in some Kenyan studies (Lee et al., 2014; Lee, Miguel and Wolfram, 2020*b*).

Our results have important implications for policy as it seeks to balance the goals of electrification and climate change mitigation that are often in conflict, on the global frontier, because the grid runs mainly on fossil fuels. We find that poor households strongly prefer the grid to the off-grid solar present in our sample. Although off-grid solar will continue to play an important role for poor households, it appears to be a bridge to the grid, rather than a means to leap past it. Therefore, subsidizing utility-scale low-carbon electricity is likely to be more effective at reducing CO_2 emissions than would be bundling the same subsidies with an inferior off-grid product. Grid-connected low-carbon generation also yields more durable decarbonization, since our model projects that households will overwhelmingly switch to the grid as they grow richer and their demand for electricity rises.

The remainder of the paper goes as follows. Section 1 describes our data sources and the rapidly changing electricity landscape in Bihar. Section 2, as a benchmark, estimates demand for a single good, solar microgrids, without accounting for substitution opportunities. Section 3 lays out and estimates our preferred model of demand for all electricity sources. Section 4 uses the model estimates to calculate the surplus gains from electrification, the decline in solar prices, the expansion of the grid, and a series of counterfactual policies. Next, Section 5 explores the external validity of the Bihar results by applying the demand model to data from eight African countries. Section 6 concludes.

1 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.

1.1 Context

Bihar, with a population of 123 million in 2021 (Census of India Population Projections), is one of India's poorest states. In 2012, the year before our baseline survey, the electrification rate in Bihar was only 25%, about one-third of the all-India rate of 79% and below the rate of 37% in sub-Saharan Africa. 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 drawing 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. The supply of electricity is limited due to both a low electrification rate and to the rationing of power supply (Jain et al., 2018).

Our study was well-timed to capture two forces changing the electricity market. The first force is a decline in the price of solar power. Innovation has brought down the price of solar panels over several decades, but only in the last decade has it reached a level low enough to make off-grid solar an affordable alternative to the grid for poor households, as seen in the rapid uptake of solar in India and especially 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 changing electrification in India specifically was a "big push" for the grid 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.[7] 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.[8] In Bihar, the state government made electricity access a priority (Kumar, 2019).

[7]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 that 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, 2022).

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

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). The government not only invested in grid extension but also held camps to sign-up households and subsidized connections, including offering free connections to all households designated as Below the Poverty Line (BPL). These state investments allowed the grid to reach progressively poorer households. 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.

1.2 Study Sample and Data

Our study sample consists of 100 villages in two districts in Bihar (see Appendix Figure A1 for a map). We worked in these villages in partnership with a solar microgrid provider called Husk Power Systems (HPS) (see Section 2.1). These villages were sampled from a broader population with poor access to electricity at baseline. Poor access was based on 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 required that villages be close enough to existing service areas for the company to reach 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 to characterize both the demand and supply sides of the market over a nearly four-year period from four sources. First, our main source of data, on the demand side, is a household-level panel survey on the sources and uses of electricity. Second, on the supply side, we obtained household-level administrative data on customer enrollment and payments for solar microgrid connections from HPS. Third, on the supply side, we collected village-level survey data from the operators of common diesel generators, an off-grid source of electricity. Fourth, on the supply side, we gained access to 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 with 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, the sample is nearly representative of the population as a whole.[9] 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 A2 shows the timing of survey rounds). The two thick rounds used nearly the same survey instrument and covered demographics, the sources and uses of electricity, and welfare relevant outcomes likely to be influenced by electricity use. The follow-up round took place one year after the endline of the experiment. This round was not part of our original plan but, in light of the massive changes we observed on the supply side, was added to update household electricity sources and choices. All told, the three surveys capture a three-and-a-half-year period of extraordinary dynamism.

1.3 Characteristics of electricity sources

In developed countries, electricity is the archetype of a homogeneous good: power is available from the grid 24×7 and can run all kinds of appliances. In many developing countries, electricity sources are differentiated products and the characteristics of sources are rapidly changing. 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: diesel generators, microgrid solar, and own solar systems, all provided in private markets. 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. A *solar microgrid* is a system composed of a solar panel and batteries that serves a small group of six to nine households.[10] An *own solar* system is a panel and battery bought and

[9] 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.

[10] 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 keypad 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 outlet, typically used for mobile phone charging.

operated by a single household. The outside option is not to have electricity from any of these sources which means relying on kerosene, candles or flashlights for lighting.

Table 1 gives summary statistics on the characteristics of electricity sources at baseline (columns 1 to 5), endline (6 to 10), and follow-up (11 to 15). Sources differ on a number of dimensions including energy services (load), 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 (total, peak and off-peak), and the share of villages in which a source is present. We highlight four findings that characterize household trade-offs.

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.

Table 1: Summary of Electricity Source Characteristics

	Baseline (2013)					Endline (2016)					Follow-up (2017)				
	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	87	91	164	-	59	89	72	170	-
Load (watts)	322	134	247	31	-	145	22	39	31	-	147	66	13	31	-
Hours of supply															
Total	10.9	3.4	7.8	3.0	-	11.0	3.1	5.5	6.4	-	13.6	3.1	5.6	6.4	-
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.

Second, the pricing in the market is fairly tightly clustered. We define the price of each source to be the monthly price of maintaining service, which is how the price is actually charged for diesel, micro-grid and many grid customers.[11] 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.[12] The highest-priced product at baseline, above this tight cluster, is microgrids, with a price of INR 200 per month. Our experiment subsidized the price of this product (Section 2.1).

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 A3 and A4 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 radically changed over our nearly four years of 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 follow-up. The availability of diesel *fell* from 57% (column 2) to 12% of villages (column 12) in the same span, because diesel generators, run by private entrepreneurs, exited as they lost market share 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 there is no power when one wants

[11] Households pay upfront 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, minimum monthly payments and infrequent meter readings imply that many poor consumers are *de facto* billed at a flat monthly rate.

[12] 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 price would make the grid among the most costly sources. We calculate the break-even price for grid electricity payments to cover variable supply costs to be INR 233.

to use the fan. 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 affect households' choices between electricity sources.

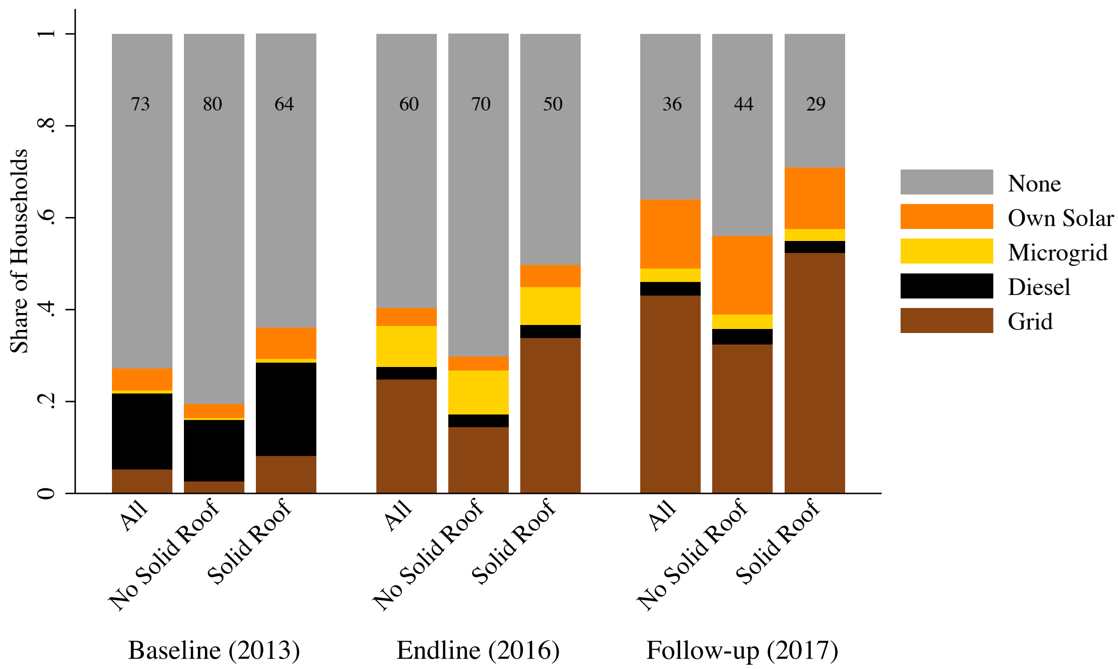
1.4 Market shares of electricity sources in Bihar

The two disruptions of solar innovation and grid expansion transformed electricity access during our study. Figure 2 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).

During the study period, all sources play an important role in electrification. Three features of Figure 2 show the transformation of Bihar's electricity sector in this period. First, there was 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, the net gain in electrification conceals a churning of market shares away from diesel and towards solar power and especially grid electricity. 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, when subsidies had been withdrawn. 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

Figure 2: Electricity Source Market Shares by Survey Wave



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 clusters of bars for shares in each respective survey wave. We use a dummy variable for whether a household has a solid roof as a proxy for household wealth. 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.

coming between the endline and follow-up rounds.

Third, there is 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.

2 Demand for Solar Microgrids

This section describes our experiment and uses the experimental variation to estimate demand for solar microgrids. Quantifying demand for microgrids, or some other source in isolation, is the common approach in the prior literature, because until recently most households could only buy electricity from a single source. We therefore estimate demand for microgrids alone as a benchmark for calculating household surplus from this product.

While we start by estimating demand for this new good, microgrids are only one of several competing electricity sources in Bihar (Section 1). Therefore, Section 3 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, reliability and quality of the grid improve.

2.1 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, added the solar microgrid product to its portfolio to reach a wider set of customers.[13] HPS was the only microgrid supplier in our sample and so we treat HPS and microgrids as synonymous.

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 where HPS did not offer microgrids (34 villages), a normal price arm where HPS offered microgrids at the prevailing price, initially INR 200 per month (33 villages), and a subsidized price arm where HPS offered microgrids at a price of INR 100 per month (33 villages). The normal price arm provides microgrid service at, or slightly above, cost and the subsidized arm at perhaps 40% below cost.[14] Within each treatment village, all households were offered the same HPS connection and pricing, regardless of whether they had previously expressed

[13]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 due to fixed costs could only serve a village if demand was sufficiently broad.

[14]We estimate the capital and installation costs of a microgrid to be INR 105 per household per month. 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.

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.[15] 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 incomes imply a 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.[16] 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.

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

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

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*
	[0.49]	[0.50]	[0.50]	(0.043)	(0.039)	(0.050)
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. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

2.2 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 survey wave t on the experimental treatment assignments:

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

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 all 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). We attribute this small but positive estimate to survey response error, since the constant in estimates using administrative data is indistinguishable from zero (Appendix Table C6). 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 when using administrative measures of household payments, rather than surveys, to measure take-up.[17]

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

[17] We have administrative data from Husk Power that contains the monthly payment history of all eligible households. Appendix Table C6 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.

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

microgrid market shares, which we investigate further in Section 3.3.

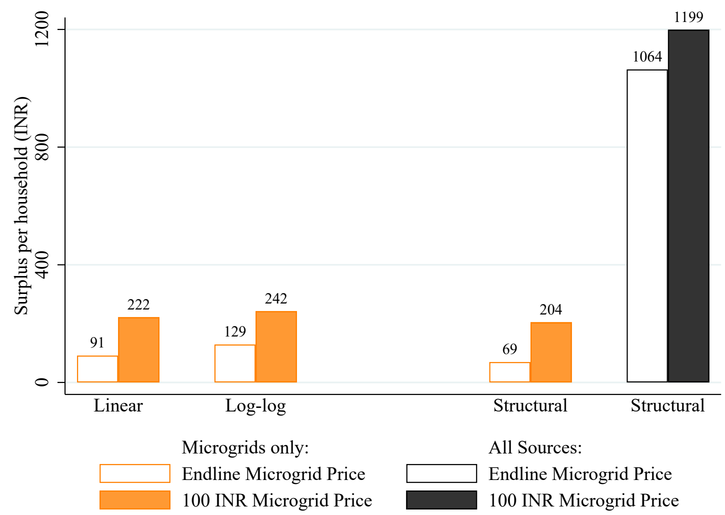
The last two columns of Table 3 give instrumental variables estimates of microgrid demand, where we instrument for the price level (or log of price) using the experimental treatment assignment. For example, the column 4 (linear) IV specification of demand consists of the two stages

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

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

A corresponding log-log specification is used in column 5. The sample for these columns is limited to the two-thirds of the 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

Figure 3: Consumer Surplus from Microgrids versus All Sources of Electricity



The figure 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. The "Microgrids only" bars show the change in surplus, in the full demand model, if microgrids are removed from the market. The "All sources" bars show the consumer surplus from all sources of electricity. We evaluate surplus at two different sets of prices. Hollow bars show surplus evaluated at endline, using the observed prices for microgrids. Filled bars show surplus evaluated at endline, but assuming microgrids are available everywhere at a counterfactual price of 100 INR.

share (standard error 0.052) for each INR 100 cut in price.

2.3 Surplus from microgrids

We use these experimental demand estimates to calculate the contribution of microgrids to household surplus as the area under the demand curve and above the price of the good. Figure 3 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 3, when we compare the results from these village-level demand specifications with those from the full household 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. The surplus numbers for the hypothetical removal of microgrids are understated in that they give the effect of removal at the time of the endline survey, when microgrids were not present in the control group, one-third of sample villages. Thus the removal of microgrids, by design, has no effect on surplus in those villages. 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 1). 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 they can easily switch to another comparable source of electricity when microgrid prices rise. On external validity, household demand for microgrids may have been drastically different in a different policy or supply environment (e.g., if the government had not made a big push for the grid or if the price of alternatives like own solar had not declined). The following sections therefore specify and estimate a demand model that covers *all* electricity sources.

3 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 mean 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.

3.1 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.[18] 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

$$u_{ijvt} = \delta_{jvt} + \mu_{ijt} + \varepsilon_{ijt}. \tag{4}$$

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

$$\delta_{jvt} = x'_{jvt}\overline{\beta}^o + p_{jvt}\overline{\beta}^p + \xi_{jvt}. \tag{5}$$

The vector x_{jvt} of source characteristics includes hours of supply on-peak (from 5 to 10 pm) and hours of supply off-peak and p_{jvt} is the price of the source. Mean unobserved source indirect utility ("quality") ξ_{jvt} is known to households but not the econometrician. For off-grid solar, we expect quality includes both unmeasured physical characteristics, such as the capacity of a battery, as well as characteristics of the service, such as the hassle to obtain a connection. For the grid, similarly, quality may include voltage fluctuations and services like village connection camps.

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

$$\mu_{ijt} = d'_{jvt}\beta^o z_{it} + p_{jvt}\beta^p \nu_i. \tag{6}$$

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 β^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 ν_i is a preference shock, unobserved by the econometrician, for the household's disutility of price that is 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

[18] At the time of our endline survey, only 1.4% of households held multiple sources. 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.

and the persistence of ν_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 2). 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 $x_{vt} = (x_{1vt}, \ldots, x_{Jvt}, d_{1vt}, \ldots, d_{Jvt}, p_{1vt}, \ldots, p_{Jvt})$. The probability that a household chooses product j' is given by

$$\begin{aligned}
\Pr\left(j' \mid z_{it}, x_{vt}, \beta, \delta_{vt}\right) &= \int_{\nu_i} \Pr\left(j' \mid z_{it}, x_{vt}, \beta, \delta_{vt}, \nu_i\right) f\left(\nu_i\right) d\nu_i && (7)\\
&= \int_{\nu_i} \frac{\exp\left(\delta_{j'vt} + \mu_{ij't}\right)}{1 + \sum_{j=1}^{J} \exp\left(\delta_{jvt} + \mu_{ijt}\right)} f\left(\nu_i\right) d\nu_i. && (8)
\end{aligned}$$

Here we make two distributional assumptions. First, in writing (8), we assume, as is common in the discrete choice literature, that the idiosyncratic utility shock ε_{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 uniform distribution with support on $[-1, 1]$. We assume the distribution is uniform, 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 β^p scales the difference between the mean coefficient on price $\overline{\beta}^p$ and the coefficient on price for a maximally price sensitive household (see 6).

Consumer surplus.— Given the coefficients in the model, we can calculate the consumer surplus for any choice set of goods, good characteristics and household characteristics. 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 (5) and (6). The consumer has price coefficient $\beta_i^p = \overline{\beta}^p + \beta^p \nu_i$. Conditional on the choice set J and a draw of ν_i, a consumer's expected surplus is

$$E[CS_{it} | \nu_i, J] = \frac{1}{\beta_i^p} \ln \sum_{j \in J} \exp(\widehat{V}_{ijvt}), \tag{9}$$

as in the logit model. We integrate over ν_i using Gauss-Legendre quadrature to approximate

$$\mathbb{E}[CS_{it}|J] = E_\nu[E[CS_{it}|J]] = \int_{\nu_i} E[CS_{it}|\nu_i, J] f(\nu_i)\, d\nu_i. \tag{10}$$

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.

3.2 Estimation

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

This two-step estimation follows a micro-BLP approach (Berry, Levinsohn and Pakes, 2004). However, the richness of our data and experiment allow several innovations relative to a typical application of mixed logit. First, with household panel data for many villages v, we can recover up to $V \times T \times J \approx 1200$ 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 with which to instrument for price. Although the experiment varies the price for only one of the sources, this variation is sufficient to identify the price coefficient. Conceptually, a household's distaste for price reflects their marginal utility of expenditure on outside goods. This parameter is the same regardless of whether a household spends an additional rupee on off-grid solar or on the grid.

Non-linear estimation of the first step.— We estimate the first step by the 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 is 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.[19] Concentrating out the δ vector in this way

[19] 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

greatly reduces the dimensionality of the non-linear search. The second and third sets of moments are used to estimate $\hat{\hat{\beta}}$. The second set of moments is 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 (5) 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 ξ_{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.

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

$$\widehat{\delta}_{jvt} = x'_{jvt}\overline{\beta}^{o} + p_{jvt}\overline{\beta}^{p} + \overline{\xi}_{jt} + \tilde{\xi}_{jvt} \tag{11}$$

$$p_{jvt} = \pi_1 T_v^{Normal}\mathbf{1}\{Endline\} + \pi_2 T_v^{Subsidized}\mathbf{1}\{Endline\} + \pi_3\widehat{Peak}_{vt} + \pi_4\widehat{OPeak}_{vt} + \alpha_{jt} + \eta_{jvt}. \tag{12}$$

Equation (11) 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 (12) gives the first-stage specification for price. The first stage constructs instruments from the experiment. Our solar microgrid experiment offers instruments that are excludable and likely to be powerful, given that the microgrid treatment changed market shares (Table 3). Equation (12) 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 α_{jt} are source-by-wave fixed effects.

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.

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 allow 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 those that are served by the same substation, will be similarly affected by the distribution companies' power supply rationing rules. The exclusion restriction is that the 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 these instruments.

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).

3.3 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[20]; 20 parameters governing household heterogeneity; 3 parameters on the average effects of source characteristics; and, a parameter for the dispersion of the random coefficient on price. We report a selection of these parameters to give a sense of how the model represents household choices. Specifically, we initially report the linear estimates of the average effects of source characteristics, from the second step. We then turn to the estimates from the non-linear first step of how household characteristics affect choice probabilities.

Linear second step estimates: Mean effect of source characteristics.—Table 4 reports estimates of the linear part of the demand model estimated using two-stage least squares. Column 1 reports results from ordinary least squares estimates as a basis for comparison. Column 2 reports

[20] Although there are four products and a hundred villages across three time periods, resulting in 1200 possible values for mean indirect utility parameters, product availability varied across villages and survey waves. We do not estimate a ξ_{jvt} parameter when a product was not offered in a village-wave.

Table 4: Demand for Electricity: Estimates of Mean Effects of Source Characteristics

	OLS	Price IV		Price & Hours IV		Alternative IVs	
Instruments		RCT	RCT	RCT	RCT	BLP	Hausman
Stage		First	Second	First	Second	Second	Second
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Price (INR 100)	-2.05***		-5.00***		-4.93**	13.1	18.8
	(0.29)		(1.93)		(1.92)	(27.7)	(57.1)
Hours of peak supply	3.70**	-0.050	3.53**		5.13**	5.95**	6.19
	(1.79)	(0.049)	(1.77)		(2.59)	(2.76)	(4.09)
Hours of off-peak supply	-1.10***	0.0081	-1.07***		-1.54***	-1.69***	-1.73**
	(0.41)	(0.013)	(0.40)		(0.57)	(0.60)	(0.82)
Treatment normal price		0.064**		0.064**			
× Endline (=1)		(0.029)		(0.029)			
Treatment subsidy price		-0.16***		-0.16***			
× Endline (=1)		(0.021)		(0.021)			
Mean peak hours				-0.032			
in nearby villages				(0.045)			
Mean off-peak hours				0.0040			
in nearby villages				(0.0094)			
F-stat for first-stage for:							
Price			43.1		21.6	0.5	0.5
Peak hours					524.1	468.8	607.3
Off-peak hours					1057.2	1209.7	1411.5
ξ_{tj} mean effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	999	999	999	999	999	945	989

The table reports estimates of the linear part of the demand model estimated using two-stage least squares. Columns (3), (5), and (7) shows estimates of the structural equation (11) and columns (2), (4), and (6) of the corresponding first stage (12). The estimates in column 1 use ordinary least squares for comparison. The other columns vary in the set of instruments used in the first stage: (3) instruments for price using experimental treatment assignments interacted with a dummy for the endline survey (equation 12); (5) additionally instruments for peak and off-peak hours of supply, using predicted hours based on supply in nearby villages; (6) replaces the experimental price instruments with average characteristics (hours of supply and load) of other products in a village (Berry, Levinsohn and Pakes, 1995); (7) replaces the experimental price instruments with the average price of each product in the nearest three villages as instrument for its price in a given village (Hausman, 1996). All specifications control for wave-by-source fixed effects. The final row of the table reports the first-stage F-statistic. First-stage estimates for hours of supply are reported in Table C2. Standard errors in parentheses are clustered at village-level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the first stage for price, using the experimental instruments, and column 3 reports instrumental variables estimates of the second stage. Columns 4 and 5 report estimates of the same pair of equations using the experimental instruments in the first stage and additionally instrumenting for peak and off-peak hours of supply. (The first stage estimates for hours are in Appendix Table C2.) Finally, columns 6 and 7 replace the experimental variables in the instrument set with alternative, non-experimental instruments for price.

We find, in the first stage estimates, that the experimental treatment assignments have a highly statistically significant effect on source prices (columns 2 and 4). Mean prices are lower at endline in villages where solar microgrids were subsidized in the experiment and higher in villages where microgrids were not subsidized, relative to control villages where they were not offered. The first-stage F-statistic for a test of the null that the instruments do not affect price ranges from 43 to 21, depending on whether we instrument only for price (column 3) or for price and hours simultaneously (column 5). Additionally, our supply instruments strongly predict hours both peak and off-peak (see Appendix Table C2, columns 2 and 3).

The experimental instrumental variables estimates show a high degree of price sensitivity. We find a coefficient of -5.00 (standard error 1.93) on price (column 3). This estimate is essentially unchanged if we additionally instrument for hours of supply (column 4). We discuss below the demand elasticities our estimates imply.

The experiment is necessary to accurately estimate household sensitivity to price. The magnitude of the coefficient on price is two-and-a-half times greater than in the OLS estimates (Table 4, column 1), which suggests that OLS is biased upward, towards zero, from some combination of endogeneity and measurement error. Alternative instrument sets commonly used in the industrial organization literature, which do not rely on the experiment, lack the power to predict price in the first stage (footer, columns 6 and 7).[21] These instrument sets therefore yield biased and imprecise estimates in this setting (first row, columns 6 and 7). The point estimates for the price coefficient are positive and the 95% confidence intervals are wide enough to include the experimental estimates, the OLS estimates and a zero coefficient on price.

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). The positive estimate for the value of peak hours is consistent with the conventional wisdom 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 also note that there is a smaller and, unexpectedly, negative coefficient for off-peak hours. We choose to proceed with the column 3 estimates, instrumenting for both price and hours, as our main specification for

[21] Neither the BLP (F-statistic 0.4) nor Hausman (F-statistic 0.5) instruments have predictive power for the endogenous price variable. 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.

counterfactual analysis.

Non-linear first step estimates: Heterogeneity in demand across households.—We report estimates for the effects of household characteristics on choice probabilities in Appendix Table C3. The table 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 choice probability of changing the value from zero to one.

The main finding of the demand estimates with respect to household characteristics is that richer households have much stronger preferences for grid electricity than all other sources. A household with a solid roof, a common wealth proxy, is 17.4 pp (standard error 0.029) more likely to choose grid electricity. 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) 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 are modestly more likely to choose microgrids (column 4). A natural interpretation of this heterogeneity is that richer households want the energy services from higher-load appliances, like a fan or television, that are mainly supported by the grid (Table 1).

Implied price elasticities.— The estimated demand model predicts how product market shares would change in response to changes in the price of each source. Table 5 reports aggregate own- and cross-price elasticities for all sources.

We find that demand for non-grid electricity sources is roughly 10 times more price elastic than demand for the grid. Panel A evaluates these elasticities at the prices observed in the endline survey. The own-price elasticities for off-grid technologies fall between -1.65 and -2.12, whereas grid electricity's is -0.20 (panel A). A plausible reason for this difference is that the grid is heavily subsidized, so that even if the grid price rises, the grid will remain the preferred source for most households that have it in their village. In panel B, we keep all non-grid prices at their endline values, but move the grid price up to the break-even price for the utility (at which households are charged the average cost of the energy they consume). We find that with the grid price set at the cost of supply, households more readily substitute to off-grid sources. The own-price elasticity of the grid increases in magnitude by a factor of four, to -0.81. The cross-price elasticities of off-

Table 5: Price Elasticities of Electricity Source Demand

	Elasticity of share of source:				
	Grid	Diesel	Own solar	Micro-grid	None
With respect to price of:	(1)	(2)	(3)	(4)	(5)
Panel A. Experimental IV estimates, at endline price					
Grid	-0.19	0.05	0.05	0.05	0.06
Diesel	0.01	-1.65	0.09	0.07	0.06
Own solar	0.02	0.23	-2.12	0.30	0.16
Microgrid	0.05	0.36	0.77	-1.93	0.20
Panel B. Experimental IV estimates, at break-even grid prices					
Grid	-0.81	0.18	0.14	0.15	0.10
Diesel	0.02	-1.70	0.08	0.08	0.06
Own solar	0.03	0.21	-2.14	0.30	0.17
Microgrid	0.07	0.36	0.78	-1.97	0.22

The table presents aggregate own- and cross-price elasticities of demand by electricity source. The elasticities are calculated for the market share of each column source with respect to the price of each row source. In panel A, the elasticities are evaluated at the mean endline price for each source. Panel B is similar but sets the price of grid at the break-even price (calculated as 233 INR = 60 kWh per month mean consumption $\times 3.88$ INR per kWh average cost of procurement). Both panels use the experimental IV estimates of mean utility from Table 4, column 5.

grid source market shares with respect to the grid price triple.[22] The price elasticities implied by alternative model estimates, not reported in the table, are much smaller in magnitude (using OLS) or have the wrong sign (using BLP instruments).

4 Source Competition 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

[22]The pattern of elasticities provides additional insights. Although grid prices are set by a natural monopolist (the government utility), these elasticities show they use pricing rules that are different from a profit-maximizing monopolist. Here grid prices are set low, on the inelastic part of the demand curve, although monopoly profits would be higher by pricing in the elastic region. On the other hand, off-grid sources can be thought of as competitively supplied and households appear quite elastic to off-grid prices, consistent with the rapid expansion of solar power in recent years as costs have fallen (Figure 1).

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.

4.1 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 3. 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 10).

The value of microgrids in the structural model is nearly the same as calculated previously with the linear model of demand. Figure 3 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 three different models is reassuring since demand is estimated with the same experimental variation in each case.

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 microgrids, is INR 1199 per year, greater than the surplus from microgrids alone by a factor of roughly six (right group of bars).

This large difference shows that the modest value of microgrids found in Section 2 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 have 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 demand estimation for all sources is critical along the global electrification frontier. Households are changing their electricity choices rapidly in response to both an expansion of the choice set and changes in the characteristics of available sources.

4.2 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 expansion of the grid, to the household surplus from electrification and to CO_2 emissions from electricity consumption.

Table 6 reports a range of counterfactual results based on the demand model estimates. Each 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 *Bihar* 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.[23] 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 break-even with variable costs (see footnote 12).

The remaining columns provide an accounting for greenhouse gas emissions and climate damages. Column 9 reports the annual mean CO2 emissions per household and column 10 converts this into the monetary value of climate damages in INR, using the US EPA's value of the social cost of carbon of $225 in 2023 (Environmental Protection Agency, 2023).[24] The *global* social surplus is calculated as the sum of columns 8 and 10 with the important caveat that the climate damages are global and that the Bihar government may not place great importance on this global externality since their state's share of climate damages is necessarily small.

We draw four main conclusions from Table 6. First, the household surplus from electrification increased by a factor of about 5 over the short 3.5-year period. At baseline, the grid had low penetration and solar was expensive and unpopular (Table 6, panel A, row A1). Diesel generators were the most popular electricity source, and households valued all electricity sources at only INR 385 per *year* (row A1, column 6). By the follow-up survey, surplus from electricity had risen to INR 1,886 (row A3, column 6). To put this gain of INR 1,501 in context, baseline household expenditures on electricity and lighting was INR 2,029 per year, and on all energy INR 6,024 per year. The increase in the surplus from electrification is therefore 74% of baseline electricity and lighting expenditures and 25% of all baseline energy expenditures.

Second, gains in household surplus come both from the grid and solar, but the gains from grid

[23] 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.

[24] To carry out this calculation, we use emissions factors for each electricity source (see Table A4 in the appendix) and multiply by household consumption from that source. For the grid, we assume 20% transmission and distribution losses because of which generated electricity exceeds end-user consumption.

Table 6: The Value of Electrification and Its Climate Damages under Counterfactual Policies

	Market Shares (%)					*Bihar* Surplus (INR per hh annually)			CO2 Emissions (per hh annually)		*Global* Social Surplus (INR per hh annually)
	Grid (1)	Diesel (2)	Own solar (3)	Microgrid (4)	All (5)	Consumer (6)	Producer (7)	Total (8)	m tons (9)	Damages (INR) (10)	Surplus - Damages (11) = (8) - (10)
Panel A. Model market shares and surplus by survey wave											
A1. Baseline	6	17	7	1	31	385	-108	277	0.17	574	-297
A2. Endline	24	3	7	10	43	1064	-492	572	0.19	659	-87
A3. Follow-up	40	3	17	5	65	1886	-842	1045	0.31	1071	-26
Panel B. Disruption due to the improvement of solar power, relative to baseline											
B1. A1 + No solar	6	18	0	0	24	345	-109	236	0.18	614	-377
B2. A1 + Improved solar	5	13	23	8	50	570	-104	467	0.14	484	-17
B3. B2 + Grid priced at cost	2	15	24	8	49	502	0	502	0.13	438	64
Panel C. Disruption due to the improvement of grid electricity, relative to baseline											
C1. A1 + No grid	0	19	8	1	28	246			0.15	506	
C2. A1 + Improved grid	37	9	4	1	51	1605	-766	839	0.33	1144	-305
C3. C2 + Grid priced at cost	21	12	5	1	40	994	0	994	0.25	846	148
Panel D. Disruption due to improvements in grid and solar, relative to baseline											
D1. A1 + Improved grid and solar	36	7	15	5	63	1718	-750	969	0.31	1071	-103
D2. D1 + Grid priced at cost	21	10	19	5	54	1125	0	1125	0.22	763	363
Panel E. Future growth in electrification via supply and demand shifts, relative to follow-up *Each row in E2-E5 adds to the previous scenario*											
E1. A3 + 50% solar cost reduction	38	2	27	15	81	2070	-782	1288	0.28	962	325
E2. E1 + Grid in all villages	47	1	22	12	83	2357	-983	1374	0.34	1190	183
E3. E2 + Increase in peak grid hours	70	0	13	6	89	3934	-1463	2470	0.50	1737	734
E4. E3 + All households at least median income	92	0	5	1	98	5915	-1909	4007	0.65	2256	1751
E5. E4 + Grid priced at cost	78	1	11	4	94	4107	0	4107	0.56	1921	2185

The table presents market shares, surplus, and climate damages under counterfactual changes in the electricity market. The counterfactual scenarios are laid out in Section 4 of the text. All counterfactuals are calculated using the full demand model estimates of Table C3, columns 6 through 10. For each counterfactual, columns 1 to 4 give the market shares of each source, column 5 gives the electrification share, columns 6 through 8 give consumer, producer, and total surplus, column 9 gives the CO2 emissions and column 10-11 give the monetary value of the CO2 emissions. 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. The emissions are the sum of grid and diesel emissions, assuming household electricity consumption of 720 kWh per year as found in the sample. The emission factor used to calculate each source's emission is in Table A4. To calculate the monetary value of climate damages, we use the Social Cost of Carbon (SCC) at 175$ per ton CO2 (2016 US$), which is the inflation-adjusted level of the average 2% discount rate EPA SCC of 190$ per ton CO2 in 2020 US$ (Environmental Protection Agency, 2023). We use the 2016 PPP-adjusted exchange rate to convert the SCC to INR. Panel A shows levels at baseline, endline, and follow-up. Panels B, C, D, and E present baseline or follow-up levels with additional specified scenarios.)

improvements are much larger. These gains are driven by dramatic improvements in the price, availability and quality of both sources over the 3.5-year study period.[25] To illustrate the gains from these shifts in supply, we consider two hypothetical pathways to electrification that improve the characteristics of only one technology at a time, holding the other constant at baseline values.

Panel B reports on the surplus and electrification rate changes due to the advances in off-grid solar, holding constant the extent and quality of the grid. We model these advances 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 (row B2 versus row B1, column 5) and the surplus from electrification by a relatively modest (given the increase in electrification) 1.7$\times$ (row B2 versus row B1, column 6). Therefore, even without state investment in the grid, off-grid solar would have doubled electrification rates. It is noteworthy that the value of solar is contingent on the state of the grid and its pricing policy (see row B3 and Panel D).[26]

Panel C reports the electrification and surplus gains due to improvements in grid extent and quality, holding the extent and quality of off-grid solar constant. We find that grid improvements on their own would have increased electrification from 31 to 51 pp and surplus by 4.2$\times$ (row C2 versus row A1). This is 81% of the total increase in the value of electrification from all sources. To put it another way, the gains from the grid alone are 6.6 times larger than the gains from improved solar alone (row C2 less row A1, as compared to row B2 less row A1.) Appendix Table C7 decomposes the gains from the grid and finds that improvements in the observable characteristics of the grid (price, peak hours) and access (grid extent and unobserved quality) are both important determinants of the increase in surplus and the grid's gain in market share.

Third, a large part of the gain in household surplus from electrification, and the advantage of the grid, is a transfer via increased producer losses (row A3, column 7). The state loses money as more households choose to connect to the heavily subsidized grid. Total surplus therefore increases by only about half of the gain in household surplus (row A3 less A1 in column 8 versus in column 6).

[25] Table 1 shows improvements in the observable characteristics of sources. Appendix Figure C1 additionally shows improvements in our model-inferred estimates of unobserved source quality over time.

[26] To illustrate the interactions between solar and the grid, Panel D values simultaneous improvements in both technologies. We estimate that improvements to solar increase consumer surplus by INR 185 per household per year when grid availability and quality are held at the baseline levels (compare row B2, column 6 to row A1, column 6). The gain from improving solar is only 60% as large (INR 114 per household per year) when carried out in the backdrop of improved grid availability and quality (compare row D1, column 6 to row C2, column 6) as compared to when the grid is left in its baseline state.

However, if we eliminate subsidies by pricing grid electricity at cost and then compare an improved grid (row C3) with improved solar (row B3), the household surplus gain from an improved grid is about twice as large as the gain from improved solar. This finding is especially striking given that the electrification *rate* from all sources is actually higher under improved solar (49% vs 40%). Solar is more prevalent, when the grid is priced at cost, but households place a higher value on grid electricity, as the improved grid allows higher-income households to connect to their most preferred source and support higher-load appliances.

Fourth, gains in electrification add much less to *global* social surplus than to *Bihari* social surplus, because households prefer the grid and the grid in Bihar is largely powered by fossil fuels (i.e., coal). The observed increase in electrification from baseline to follow-up increased annual household CO_2 emissions by 0.14 metric tons and climate damages by about INR 500 per hh per year (columns 9 and 10, row A3 less A1). The result is a net gain in *global* social surplus of only INR 271 (column 11, row A3 less A1), or 18% of the gross gain to Bihari households (recall column 6, row A3 less A1). By comparison, partial electrification, via off-grid solar alone, would slightly decrease emissions, by taking market share from the grid, and increase social surplus by INR 360, despite being much less preferred by households (column 11, row B2 less B1). This accounting, for marginal changes in electrification surplus and climate damages, must not obscure that households in rural Bihar have extremely low levels of electricity consumption and therefore emissions in all the scenarios we consider, about 1/90$^{\text{th}}$ of US levels per capita.[27] Nevertheless, this result underscores the substantial difference between *Bihar*'s surplus and *global* social surplus on the margin of rural electrification.

More broadly, this quantification makes clear why a model is necessary to understand the two great disruptions to electricity supply in Bihar. Without the model we could not estimate: (i) the total value of electrification to households; (ii) how much different electricity sources contribute to this value, either singly or in combination; or (iii) CO_2 emissions and the resulting climate damages due to electricity consumption under different counterfactuals. The demand model also allows us to step back and measure the effects of respective improvements in both the grid and off-grid power sources. Solar and the subsidized grid have similar impacts on electrification rates, but the grid creates more household surplus and CO_2 emissions.

[27] On average, households have about 6.4 people and account for 0.17 metric tons of CO_2 at the baseline, rising to 0.31 at the follow-up. As a point of comparison, *per capita* electricity sector CO_2 emissions are about 5 tons in the US (calculated using US electricity power emission and population data from U.S. Energy Information Administration (2024)), nearly 90 times baseline per capita emissions in our sample.

4.3 Counterfactual policy reforms and the future of electrification

Scenarios.—The model additionally allows us to move beyond the study period to project how demand for different electricity sources will evolve in Bihar as market conditions, government policies and household characteristics change. This subsection uses the model to develop estimates of the electrification and surplus gains from a further reduction in solar prices, additional improvements in the grid, and growth in household incomes.

Falling solar prices. One of the key disruptions in global electricity markets over the last decade 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). Here we assess the consequences if these predictions prove accurate. Specifically, we consider a 50% reduction in stand-alone solar systems and microgrids such that on a monthly basis the price of the former goes from INR 72 to INR 36 and the latter from INR 170 to INR 85.[28]

Improving grid. As of the follow-up survey, the grid was still present in only 72% of villages and supplied 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 these improvements in the grid, we construct a counterfactual where the grid is extended to all villages and peak supply hours are extended by two hours a day up to a maximum of five hours (i.e., 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 in state product of 11% from 2012 to 2018. To model how demand changes as a result of growth, we evaluate a counterfactual that increases household income and wealth proxies. In this counterfactual, we set each characteristic that enters demand to the maximum of a household's actual value and the median value of that characteristic in the sample (see Appendix Table C4 for summary statistics). This ensures all projected household characteristics remain at within-sample values. The resulting increases in income and wealth proxies are large in relative terms but yield a counterfactual population that remains poor in an absolute sense.[29]

[28]This scenario 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.

[29]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.

Results.—Table 6, panel E reports on the results of the forward-looking counterfactuals that carry out these changes. The panel considers changes relative to the state of the market at the follow-up survey, described in row A3, by cumulatively adding 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 E1), extending the grid to all villages (row E2), and increasing grid supply during peak demand hours (row E3).

Two main findings fall out of these supply side improvements. First, future improvements in solar are projected to increase electrification rates substantially, but the resulting gains in household surplus are modest. Specifically, the solar improvement increases the electrification rate in row E1, compared to row A3, by 16 percentage points (25%), but the gain in household surplus is only about 10%. It is apparent that solar improvements pull households into electrification that place a relatively low valuation on solar electricity, presumably because they are poor.

Second, improvements in the extent and quality of the grid are projected to cause vast increases in household surplus and to make the grid the dominant electricity supplier. Adding the extension of the grid to all villages to the solar efficiency improvements barely changes electrification rates, but leads to a roughly 15% increase in surplus due to households switching from solar (-8 pp market share) to the grid ($+9$ pp) (row E2 versus row E1). The big gain from an improved grid comes from ensuring that there is supply during all of the peak hours: relative to solar efficiency gains only (row E1), a universal grid increases surplus by 14% (row E2) and a universal and reliable grid by 90% (row E3). The universal, reliable grid nearly doubles the grid's market share, to 70% of the population. These future improvements in the grid are roughly $10\times$ as valuable to households as gains in solar efficiency, even when there have been large price cuts in off-grid solar. At the same time, the improvements in the grid come with substantial *percentage* increases in CO_2 emissions but still very low emissions *levels* relative to the US and other developed countries.

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 as of our follow-up survey. All of these gains occur despite these counterfactuals freezing Bihar's development on the demand side of the model.

In the next scenario, we raise all households to the median values of sample income and other observable characteristics. We find very large further gains in surplus that lead to near-universal electrification through grid dominance. Household growth increases surplus by another 50% (INR 1981, row E4 less row E3) and increases the electrification rate to fully 98%, with 92% of households on the grid and just 6% of households choosing solar. This projected grid dominance was foreshadowed by the demand model estimates, which showed that moderately richer households have much stronger preferences for grid electricity.

The competition between the grid and off-grid solar is not 'fair', as the grid is heavily subsidized. The expansion of the grid and increases in household demand together imply that the state loses INR 1909 per household per year on grid supply (row E4, column 7). Row E5 considers the consequences of the government removing its current subsidies so that the grid breaks even on a variable cost basis. Higher grid prices cause the grid's market share to decline from 92% to 78%, but the decline in the electrification rate is only 4%, because solar picks up the slack, with its market share rising by 9 pp. This finding underscores that solar is a valuable fall-back for poorer households that can provide basic electricity services when the grid is not affordable. There is, however, little suggestion that off-grid solar will leapfrog the grid for households that can afford either source. Meanwhile, this migration to the grid increases CO_2 emissions, but they again remain at low levels, undoubtedly reflecting the low income levels of households in this sample.

Discussion.—Across the counterfactual scenarios in panel E, we see that electricity policy plays a central role in determining households' choices and surplus. Off-grid solar provides basic electricity services that households use as a backstop when the state chooses to price the grid at cost, offers it in only some locations or with low reliability. Consequently, there is little evidence in this setting that off-grid solar is a viable path to increasing electrification without increasing CO_2 emissions.

It is natural to wonder whether deficiencies in grid supply explain the huge differences in off-grid solar market shares at different points on the global electrification frontier (recall Figure 1), or whether they reflect differences in households' preferences over electricity sources (i.e., demand). For example, prior work in Kenya has found that many rural households do *not* connect to the grid when it arrives and have low willingness-to-pay for grid connections (Lee et al., 2014; Lee, Miguel and Wolfram, 2020*b*), which contrasts with our finding that the grid is dominant in Bihar, and will grow more so in the near term. These differences motivate the next section's exploration of the external validity of the findings from the Bihar sample.

5 The value of electricity sources along the global frontier

This section uses our demand model estimates to value electrification outside of the Bihar sample. It is striking that off-grid solar has become dominant in parts of Africa, reaching more than half of households in some countries (Figure 1), while solar take-up remains muted across India and we find it is less valuable than the grid in Bihar. We apply our demand model to value household choices for rural households in eight African countries at different points on the global electrification frontier. The model estimates and data from these countries will help us to understand the reasons for differences in market outcomes.

5.1 Data and methodological approach

To conduct this analysis, we gathered data on household electricity choices and the supply side of the electricity sector in Africa. We then partially refit a version of the paper's demand model, using this new data, and use the adapted demand model to calculate how African households value different energy sources. This subsection describes each of these steps.

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 cover 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 sample 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. We use household consumption to replace income in the model (except in the Kenyan survey, which does cover income). On the supply side, we define markets v using enumeration areas in each survey or aggregations over nearby enumeration areas. We then infer the availability, pricing and reliability of electricity sources j available in a given market from the survey responses of households in that market using a given source (see Appendix A.2 for details). Appendix Table A3 compares the mean characteristics of households and electricity sources in our Bihar sample to those in the

African data. Household demographics and income are broadly comparable between the two samples. On the supply side, however, grid and diesel electricity are markedly more expensive for rural African households than in Bihar, while off-grid solar is somewhat cheaper

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

We, therefore, take the preference parameters $\hat{\beta}$ to be as estimated in Bihar and refit the ξ_{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

$$\delta'_{jvt} \leftarrow \delta_{jvt} + \log\left(s_{jvt}\right) - \log\left(\Pr\left(j \mid z_{it}, x_{vt}, \hat{\beta}, \delta_{vt}\right)\right) \tag{13}$$

to infer the quality of each source-market-wave that rationalizes the choices in the African surveys. Here z_{it} are the household characteristics, x_{vt} the source characteristics and s_{jvt} the market shares in each African survey. This yields unique δ_{jvt} such that predicted market shares in the model equal observed shares in the data. We can then solve (5) for the unobserved qualities ξ_{jvt} of each source in each market and survey.

We favor this approach to gauge the household surplus from electrification under local supply conditions in each country. Conceptually, we treat the $\hat{\beta}$ as fixed demand parameters and the quality terms ξ_{jvt} as supply parameters. An alternative would be to assume that the source quality distribution in Africa is the same as in Bihar, but we think this assumption is implausible. A limitation of re-estimating the quality parameters is that we cannot test model fit against out-of-sample market shares, because the refitting of source quality in each survey means that we fit market shares by construction. Our approach is, nonetheless, able to measure how the characteristics of African households and policy-determined supply factors, respectively, contribute to the household surplus from electrification.

5.2 Results

We apply the model to calculate several measures of household surplus. For each country, Figure 4 reports three bars: the surplus from electrification (hollow bar at left), as well as the surpluses if only the grid (brown bar at center) or only solar (the orange bar at right) were available. These values are calculated with (10) and shown in INR against the left axis. Each panel reports on a different scenario; A uses each location's observed supply side, B lowers the price of the grid to the Bihar median grid price, and C additionally extends the grid to all villages. The countries are ordered from left to right by per capita income, measured against the right axis.

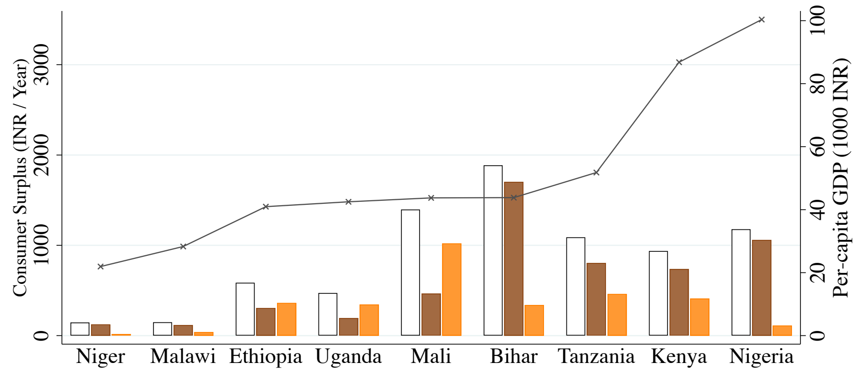
There are three main findings from these cross-country counterfactuals. First, rural Bihar is an outlier when compared to African countries of similar income. In Figure 4, panel A, we find that the surplus from electrification is roughly increasing in household income from poor countries at left (Niger, Malawi) to richer countries at right (Kenya, Nigeria). However, the surplus from electrification in Bihar greatly exceeds the surplus from electrification in the rural parts of all African countries in the sample. Most strikingly, the household surplus in Bihar is roughly *twice* that in Kenya and Nigeria, even though those countries have richer rural populations.

Second, solar matters more for the surplus from electrification in Africa than in Bihar. In relatively poor African countries, including Ethiopia, Uganda, and Mali, solar generates more surplus than the grid. In richer countries like Tanzania and Kenya, solar generates about half as much surplus as the grid. Only in Nigeria, the richest country in the sample, does the grid dominate the value of electrification as it does in Bihar.

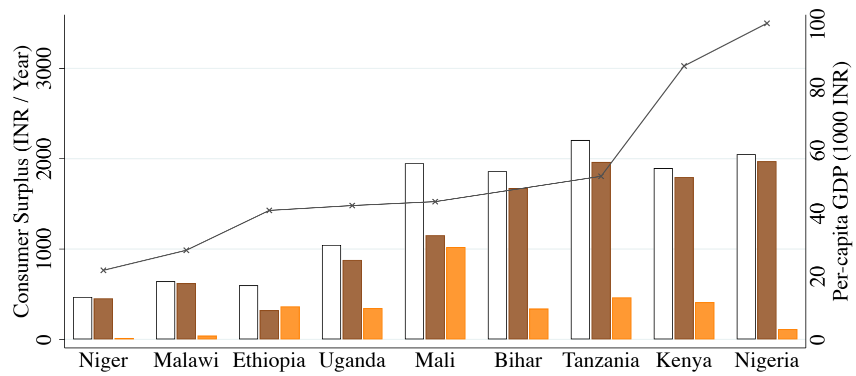
Third, remarkably, the differences in the value of electrification across countries are accounted for to a good extent by only two factors: energy subsidies and the extent of the grid. In panel B we cut grid prices everywhere to the median price in Bihar. This change is significant because many African countries do not have electricity subsidies like India. The surplus from the grid in Tanzania, Kenya and Nigeria rise to meet or slightly 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.

This stark finding, that surplus across countries would be nearly equalized if they had the same electricity policies, does not follow mechanically from our refitting of the demand model. The

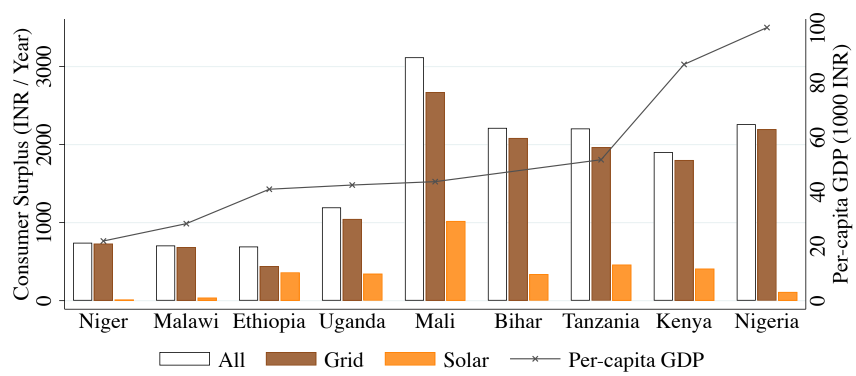
Figure 4: Consumer Surplus from the Grid and Off-grid Solar Along the Global Frontier



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 our full structural demand model extended to select African countries. The solid black line shows per capita GDP in a given 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. Surplus for Bihar is based on our experimental sample.

data for each country is distinct and heterogeneous and household valuations for electricity are determined by their characteristics in the survey data. In Niger, Malawi and Uganda, poorer rural populations imply that cheap, universal electrification would bring gains in surplus that are large, relative to baseline levels, but still modest in absolute terms.

Our findings help to understand the different paths countries are taking to universal electrification. A wide variety of explanations could be offered, from preferences to credit constraints, geography or politics (Lipscomb, Mobarak and Barham, 2013; Gertler et al., 2016; Trotter, 2016). Our results argue that the proximate causes of the different paths are only two, both on the supply side: grid extension and pricing (subsidies) explain most of the quantitative difference in the value of electrification between Mali, Tanzania, Kenya, Nigeria and Bihar. In Niger, Malawi, Ethiopia and Uganda, the value to households would remain low even after an India-style "big push" for grid electrification. The state of electrification in large parts of Africa, characterized by limited grid availability and high prices, makes off-grid solar more important than it is in India.

6 Conclusion

The global electrification frontier has radically changed. Until recently, hundreds of millions of households in developing countries had few choices for electricity. They could either make do with nothing (practically, kerosene lamps or make-shift sources, like lights attached to car batteries), or use costly, polluting power from diesel generators. Off-grid solar has brought a revolution in electrification by expanding households' choice set. At the same time, countries are moving ahead with traditional grid extension. The mode of electrification, on-grid or off-grid, determines its environmental impact, given that off-grid solar emits no carbon while the grids on the global electrification frontier remain carbon-intensive.

The paper's core contribution is to provide empirical insight into the benefits and external costs of this revolution in electrification, founded in households' own choices (i.e., demand) over all sources of electricity. An optimistic view has been that off-grid solar, with its falling costs and rising market share, will carry the day in the medium-term. We find instead that households, even at Bihar's low level of income, strongly prefer the grid to off-grid solar, and that this preference will grow even stronger as income rises.

This paper points to several areas for further research on the role of electrification in economic development. First, our analysis considers surplus and not social welfare or the downsides of energy subsidies. Governments often view access to modern energy as a right and favor subsidies as a means of redistribution (Burgess et al., 2020). A drawback of redistribution through energy is that subsidies may undercut the quality of energy supply (Dzansi et al., 2019). Second, we measure household willingness-to-pay for electrification and, apart from climage damages, no other external returns beyond the household. Some prior research suggests that a big push for electrification might generate positive spillovers in consumption or increases in productivity, beyond the direct value of electricity to consumers or firms (Lipscomb, Mobarak and Barham, 2013; Kline and Moretti, 2014).[30] Third, the economics of building renewables like solar and wind onto grids in developing countries, with distinct institutions and markets from those of the grids in the US or Europe, merit much greater study (Ryan, 2021).

From a policy perspective, it is apparent that off-grid solar is not leaping over the grid, as mobile telephony made landline networks obsolete even before they were completed. For this reason, subsidizing utility-scale low-carbon electricity generators is likely to be more effective at expanding electrification *and* limiting CO_2 emissions than bundling subsidies with off-grid electricity sources that households, as potential customers, consider inferior.

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.

Alsema, E A. 2000. "Environmental Life Cycle Assessment of Solar Home Systems." Department of Science, Technology and Society, Utrecht University NWS-E-2000-15.

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.

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

[30] In pursuit of such returns, governments may push for universal electrification, even if it may seem too early in the process of development. At the same time, Bihar's big push has come at a very low level of income, relative to historical precedent (Lee, Miguel and Wolfram, 2020*a*). On the negative side, fossil fuel intensive grids also increase *local* air pollution concentrations that undermine health, productivity and human capital.

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.

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

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

Burlig, Fiona, and Louis Preonas. 2022. "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."

Central Electricity Authority. 2021. "CO2 Baseline Database for the Indian Power Sector User Guide Version 16." Central Electricity Authority, Ministry of Power, Government of India, https://cea.nic.in/wp-content/uploads/baseline/2021/06/User_Guide_ver_16_2021-1.pdf.

Deryugina, Tatyana, Alexander MacKay, and Julian Reif. 2020. "The long-run dynamics of electricity demand: Evidence from municipal aggregation." *American Economic Journal: Applied Economics*, 12(1): 86–114.

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.

Ember. 2024. "Ember (2024); Energy Institute - Statistical Review of World Energy (2024) – with Major Processing by Our World in Data. "Share of Electricity Generated by Bioenergy – Ember and Energy Institute" [Dataset]. Ember, "Yearly Electricity Data"; Energy Institute, "Statistical Review of World Energy" [Original Data]."

Environmental Protection Agency. 2023. "EPA Report on the Social Cost of Greenhouse Gases: Estimates Incorporating Recent Scientific Advances." Environmental Protection Agency.

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).
Gertler, Paul J., Orie Shelef, Catherine D. Wolfram, and Alan Fuchs. 2016. "The Demand for Energy-Using Assets among the World's Rising Middle Classes." *The American Economic Review*, 106(6): 1366–1401.
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.
IEA. 2023. *World Energy Outlook 2023, IEA.* Paris.
International Energy Agency. 2017. "From Poverty to Prosperity, a World Energy Outlook-2017 Special Report."
Ito, Koichiro. 2014. "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing." *American Economic Review*, 104(2): 537–563.
Jain, Abhishek, Saurabh Tripathi, Sunil Mani, Sasmita Patnaik, Tauseef Shahidi, and Karthik Ganesan. 2018. "Access to Clean Cooking Energy and Electricity: Survey of States 2018." New Delhi:CEEW.
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.

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.

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

Ryan, Nicholas. 2021. "Holding Up Green Energy." *National Bureau of Economic Research Working Paper #29154*.

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.

Trotter, Philipp A. 2016. "Rural Electrification, Electrification Inequality and Democratic Institutions in Sub-Saharan Africa." *Energy for Sustainable Development*, 34: 111–129.

U.S. Energy Information Administration. 2024. "Monthly Energy Review." Accessed: 2025-01-02.

World Bank. 2024. "World Development Indicators." Accessed: 2025-01-17.

Online Appendix

Electricity Demand and Supply on the Global Electrification Frontier

Robin Burgess, Michael Greenstone, Nicholas Ryan and Anant Sudarshan

A Appendix: Data

This Appendix describes our data collection and the construction of instrumental variables for hours of electricity supply. Section A.1 describes our original data from Bihar. Section A.2 describes the Living Standards Measurement Survey data we use for seven African countries: Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania and Uganda, as well as the similar household survey data we use from the Kenya National Bureau of Statistics.

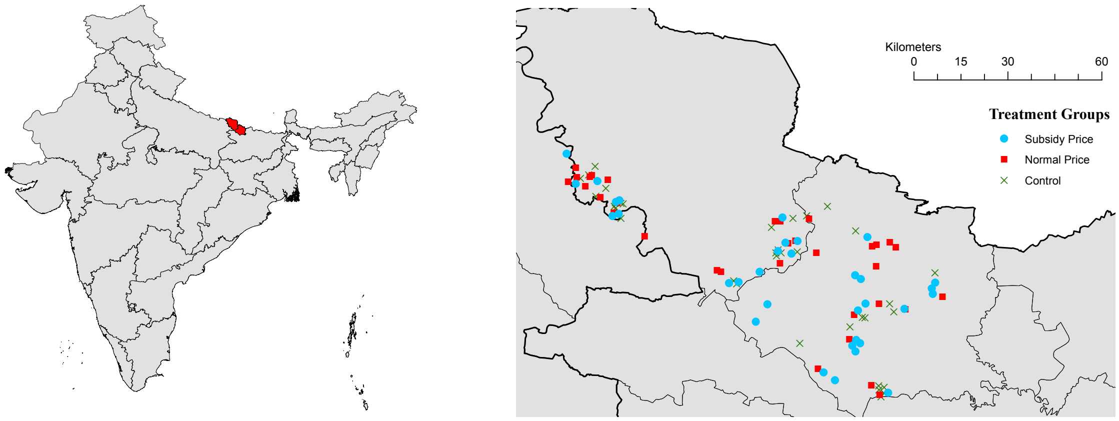
A.1 Bihar experimental data

We describe our panel survey in Section 1.2. 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 2). 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 logbooks. The data sources (i) and (ii) are matched at the customer level to our survey respondent households.

Figure A1: 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 contains 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.

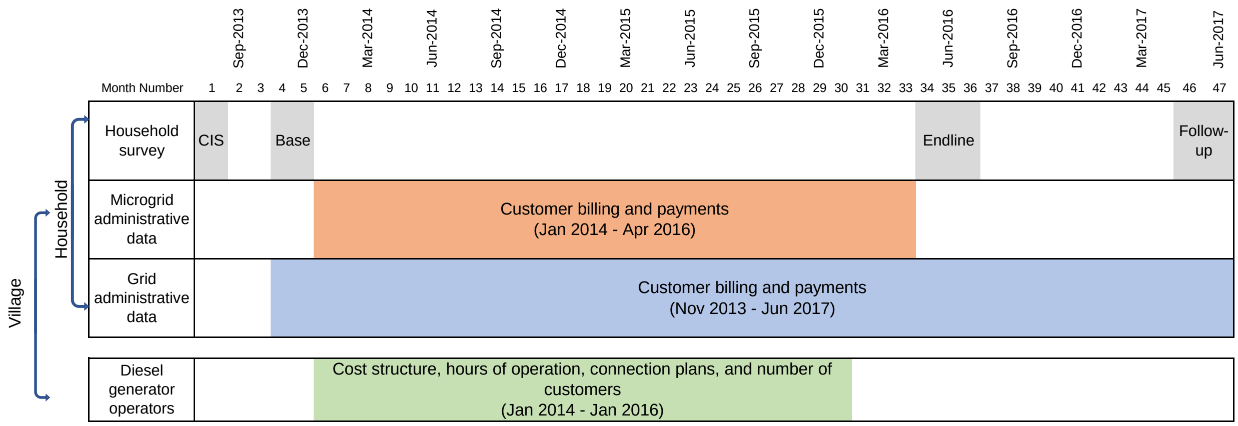
Many households using the grid in the survey are not matched to the administrative database, as there are high rates of informal connections, i.e. electricity theft, in Bihar. We can measure informal connections by designating households as informal if they could not provide a customer ID from their 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 two maps of Bihar within India (panel A) and the sample villages within Bihar (panel B).

Figure A2: 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 1 1.2 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.

Figure A2 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.

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 the power supply. We aggregated this data from the hourly level to compute the 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 latitude and longitude as covariates, division fixed effects, and their interactions. The root mean squared error of our prediction, for villages where data is available, is 1.9 hours.

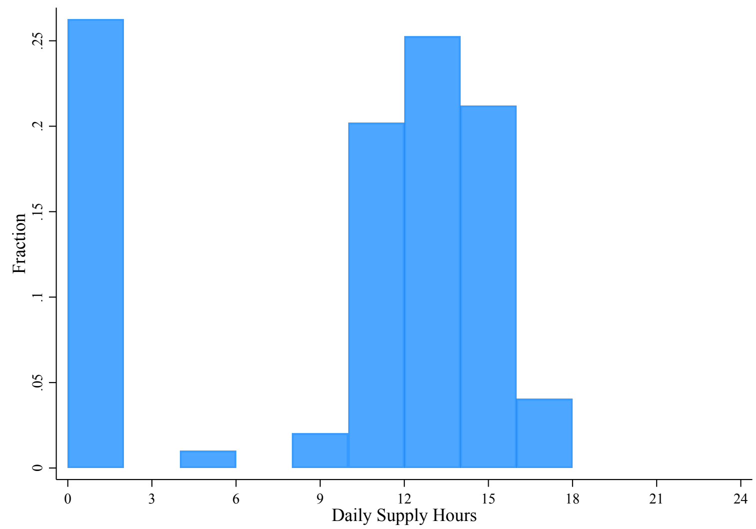
Construction of instrument for hours of supply .—The experiment provides instruments for price but not for other product characteristics, which in principle may also be endogenous to demand: for example, a high-demand village may be given more supply by the distribution company. Our preferred specification 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.

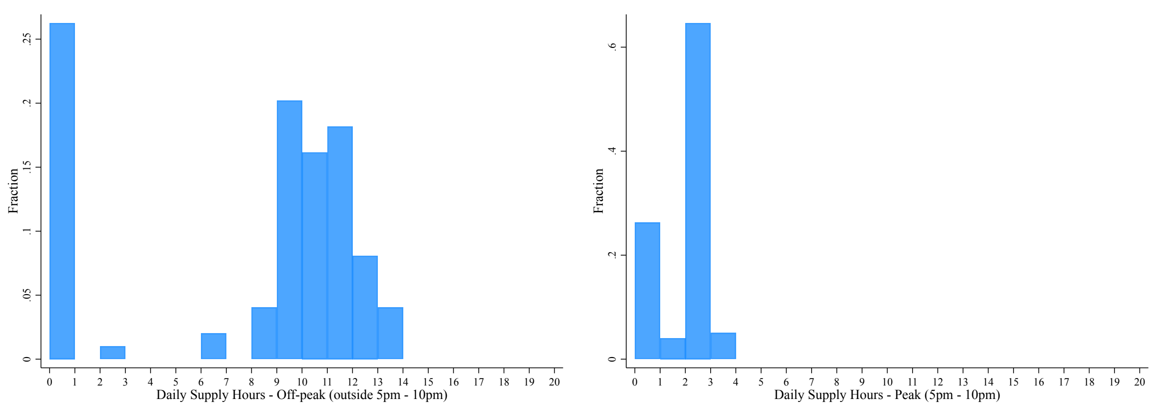
Summary statistics .—Appendix Figure A3 shows the distribution of daily hours of electricity supply on the grid and Figure A4 the distributions of supply hours during off-peak and on-peak times.

Figure A3: 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.

Figure A4: 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.

A.2 Living Standards Measurement Surveys

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-ISA), conducted by the World Bank. The specific sampling design of the surveys varies, 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.

The data on Kenya is sourced from two surveys by the Kenya National Bureau of Statistics: Kenya Integrated Household Budget Survey (KIHBS) in 2015 and Kenya Continuous Household Survey Programme (KCHSP) in 2019. Both surveys cover all counties and are designed to be nationally representative. Generally our treatment and cleaning of these data sources follows that for the LSMS surveys and so we discuss them together in this section.

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.

Variable construction and imputation.—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 household characteristics in the LSMS data are displayed in Table A1.

Source characteristics were aggregated at the market 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.

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 the grid price is only available for one wave, but not the other, the 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 the 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[31]. 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 A2 summarizes supply side characteristics in the LSMS data from Africa. Table A3 compares average values of characteristics on the demand and supply sides of the market between Bihar and the African survey data.

[31]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 A1: Summary of household characteristics in LSMS and KNBS 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)	Mobile (11)	TV (12)	Fridge (13)	Bulb (14)
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.

Table A2: 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	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Ethiopia	2011	49.55	169.21	443.28	10.32	0.00	0.00	9.42	7.52	3.90	3.36	2.55	2.10	3,466
Ethiopia	2018	49.55	123.29	443.28	17.12		12.86	7.13	7.53	3.91	2.55	2.55	2.11	3,029
Kenya	2015	259.66	142.97	363.32				9.05	7.53	3.93	3.23	2.55	2.11	13,087
Kenya	2019	259.66	116.73	363.32				9.05	7.53	3.93	3.23	2.55	2.11	12,598
Malawi	2010	342.12	73.17	87.54	16.72			9.18	7.53	3.93	3.28	2.55	2.11	10,038
Malawi	2016	352.32	30.97	100.79	18.62			9.17	7.53	3.93	3.27	2.55	2.11	10,113
Mali	2014	422.91	149.53	339.02	10.38			9.22	7.53	3.93	3.29	2.55	2.11	1,786
Mali	2017	422.91	129.85	339.02				9.05	7.53	3.93	3.23	2.55	2.11	6,841
Niger	2011	234.28	169.21	39.46	20.31			9.10	7.53	3.93	3.25	2.55	2.11	2,430
Niger	2014	251.03	149.53	21.49	16.68			9.18	7.53	3.93	3.28	2.55	2.11	2,261
Nigeria	2010	149.87	48.11	225.66	23.59		22.55	3.73	7.53	3.23	1.33	2.55	1.74	3,351
Nigeria	2018	218.99	123.29	226.13	19.36	25.71		2.76	7.51	3.47	2.44	2.55	2.51	3,376
Tanzania	2008	183.91	188.89	805.45				9.05	7.53	3.93	3.23	2.55	2.11	2,039
Tanzania	2019	183.91	116.73	805.45				9.05	7.53	3.93	3.23	2.55	2.11	671
Uganda	2009	598.71	257.70	647.76				13.45	7.53	3.93	4.80	2.55	2.11	2,086
Uganda	2019	352.37	56.39	983.88				14.04	7.53	3.93	5.01	2.55	2.11	2,233

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.

Table A3: Comparison of Demand and Supply-Side Characteristics between Bihar and Africa (Selected Countries)

	N	Bihar Means	Africa Means	Δ Means	p-value
Panel A. Demand Side Characteristics					
HH Head Education (1–8)	41319	2.55	4.94	2.39	< 0.01
Number of HH Members	41319	5.96	5.50	-0.46	< 0.01
Number of HH Adult Members	41319	3.43	2.53	-0.90	< 0.01
Household Income (INR/month)	41319	7356.65	5584.14	-1772.52	< 0.01
Owns Agricultural Land (=1)	41319	0.68	0.74	0.06	0.02
Number of Rooms	41319	2.49	2.92	0.43	< 0.01
Solid House (=1)	41319	0.27	0.36	0.09	< 0.01
Solid Roof (=1)	41319	0.46	0.67	0.20	< 0.01
Own Mobile (=1)	37845	0.78	0.43	-0.35	< 0.01
Own Fan (=1)	20641	0.02	0.20	0.18	< 0.01
Own Television (=1)	41196	0.02	0.14	0.12	< 0.01
Panel B. Supply Side Characteristics					
Grid Available (=1)	41319	0.35	0.82	0.46	< 0.01
Grid Price	41319	72.04	197.26	125.21	< 0.01
Grid Hours (Peak)	41319	1.98	2.63	0.65	< 0.01
Grid Hours (Off-Peak)	41319	8.61	7.37	-1.24	< 0.01
Diesel Price	41319	98.75	367.16	268.41	< 0.01
Diesel Hours (Peak)	41319	3.38	1.96	-1.41	< 0.01
Diesel Hours (Off-Peak)	41319	0.00	3.65	3.65	< 0.01
Solar Price	41319	140.24	123.11	-17.13	< 0.01
Solar Hours (Peak)	41319	4.51	2.55	-1.96	< 0.01
Solar Hours (Off-Peak)	41319	1.86	7.53	5.67	< 0.01

The table reports the balance of covariates between Bihar and African sample used in our analysis. The Bihar data is constructed using our own survey data and administrative data from HPS and the state utility, with detail in section A.1. The African data is constructed by Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda data from World Bank LSMS - ISA and Kenya data from KIHBS, with detail in section A.2. We restrict the year in the balance to the earliest year available for each country or region. To account for the survey weight that differs by country and year in the African data, we use World Bank (2024) population data to reweight and construct an African-wide weight. There is only one type of solar in the African data, whereas the Bihar solar variables are taken from the average of the HPS microgrid and own solar. Standard errors are clustered at the village level.

A.3 Emission Calculation

The CO_2 emissions in Table 6 are the sum of the emissions from the grid, diesel, own solar, and solar microgrid. We do not calculate the emission of the outside option of not using electricity. The emission of each source is calculated using the Bihar electricity consumption per capita, multiplied by the average number of people in a household, the market share of the source, and the emission factors summarized in Table A4. The grid emission factor is the weighted average emission rate of the grid in India at 0.82 metric tons of CO_2 per MWh (Central Electricity Authority, 2021). We also multiply the household electricity consumption by 1.2 to account for a potential increase in emissions between the grid generation and consumer use. We use Alsema (2000)'s diesel emission factor at 3.13 kg CO_2 per kg diesel, multiplied by the diesel volume conversion ratio of 0.84 kg per liter and by an assumption of 0.4 liters per kWh diesel use. Both solar sources are assumed to emit no CO_2.

Table A4: Emission Factor for Different Sources

Source	**Emission Factor (ton CO_2 per MWh)**
Diesel	1.1
Grid	0.8
Own Solar	0.0
Solar Microgrid	0.0

The table compares the emission factors used to calculate emissions from each electricity source. The grid emission factor is the weighted average emission rate for India, calculated by dividing the absolute CO_2 emissions of all power stations by the total net generation (Central Electricity Authority, 2021). Diesel emission factor is calculated using Alsema (2000)'s estimation of 3.13 kg CO_2 per kg diesel consumed, assuming a conversion ratio of 0.84 kg per liter and 0.4 liter per kWh efficiency. Both solar options are assumed to not emit any CO_2.

B Appendix: Model

B.1 Estimation moments in detail

This section derives the moments used in estimation.

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

$$\Pr\left(j'_{it}, x_{vt}, \beta, \delta_{vt}\right).$$

This prediction relies on an integral over the distribution of ν_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 ν_i once for a household and hold them fixed across simulations. These draws are used to form predicted market shares as

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

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

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

where $\tilde{s}_{vtj}$ are the observed market shares in the data, slightly corrected for sampling error.[32] We write $\delta(\beta)$ because we concentrate the δ out of estimation using the contraction mapping of Berry (1994). 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.

[32] 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.

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

Covariances of household characteristics and chosen sources.—The second set of moments is based on covariances between household characteristics and indicators for the electricity source they chose. Let d^i_{jvt} equal one if household i in village v and time t chose source j. Let $z_i = (z_{i1} \ldots z_{iT})$ be an $R \times T$ matrix of household characteristics and $d^i_{vk} = (d^i_{jv1} \ldots d^i_{jvT})'$ be a $T \times 1$ vector stacking the sources 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 choices are:

$$\underbrace{G_i\left(\tilde{\beta}, \delta_v\right)}_{R\times 1} = z_i \underbrace{\left(d^i_{jvt} - \mathbb{E}\left[d^i_{jvt} \mid z_i, \tilde{\beta}, \delta_v\right]\right)}_{T\times 1}$$

This moment gives the product of the household characteristic and the deviation of the household's chosen source from their expected choice probability for that source in the model. Expected choice probabilities for each household and source are formed in the model using Gauss-Legendre quadrature as:

$$\mathbb{E}\left[d^i_{jvt} \mid \mathbf{z_{it}}, \tilde{\beta}, \delta_v\right] \approx \sum_q w_q \Pr\left(j' \mid \mathbf{z_{it}}, \mathbf{x_{vt}}, \beta, \delta_{vt}, \nu_{iq}\right).$$

The expected source choice probabilities 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 $j = 1, \ldots, J$ inside sources. We form a vector of moment conditions

$$\underbrace{G_i^2\left(\tilde{\beta},\delta_v\right)}_{RJ\times 1}=\begin{bmatrix} G_{i1}^2\left(\tilde{\beta},\delta_v\right) \\ G_{i2}^2\left(\tilde{\beta},\delta_v\right) \\ \vdots \\ G_{iJ}^2\left(\tilde{\beta},\delta_v\right) \end{bmatrix}=\begin{bmatrix} G_{i,j=1,r=1}^2 \\ G_{i,j=1,r=2}^2 \\ \vdots \\ G_{i,j=2,r=1}^2 \\ G_{i,j=2,r=2}^2 \\ \vdots \\ G_{i,j=J,r=R}^2 \end{bmatrix}.$$

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

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, ℓ may denote grid electricity, off-grid electricity, and no electricity as different groups of possible choices. Then the transition matrix between groups is

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

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.

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

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

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

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_i^3 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

$$Pr(\ell_t, \ell'_{t+1}|z, x, \beta, j_t) \approx \frac{1}{S}\sum_s \Pr(\ell_t \mid z_{it}, x_{vt}, \beta, \delta_{vt}, \nu_{is}) . \Pr(\ell'_{t+1} \mid z_{i,t+1}, x_{v,t+1}, \beta, \delta_{v,t+1}, \nu_{is}) .$$

The fact that draws of unobserved tastes ν 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.

B.2 Estimation procedure and objective

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

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

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

   (a) Use G^1 and the BLP contraction mapping to solve for δ_{vtj} such that predicted market shares equal observed shares. Berry (1994) shows that

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

   is a contraction mapping. For any $\tilde{\beta}$, iterating until convergence recovers the unique

δ_{jvt} such that the predicted market shares of product j in village v at time t exactly equal observed shares. Let $\delta(\tilde{\beta})$ be the $(V \times T \times J) \times 1$ vector of such shares.

(b) Stack the household-level moments as

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

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 $G(\tilde{\beta})$ be the $M \times 1$ vector with the row means

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

(c) Form the objective as

$$\hat{Q}\left(\tilde{\beta}, \delta, W\right) = G(\tilde{\beta})' W G(\tilde{\beta}) \tag{15}$$

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

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

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

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

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

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

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

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

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

C Appendix: Additional Results

This section presents additional results on demand. Subsection C.1 reconciles market shares in the raw data, with Laplace correction, and as predicted by our structural model. Subsection C.2 presents estimates of the first stage from the estimation of the second, linear part of our structural demand model. Subsection C.3 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 C.5 provides additional estimates to check the robustness of the structural demand estimates to alternative nest structures in the nested logit model.

C.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 32), 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 effect on market shares (Table C1, panels A through C, row 3 versus row 2).

Table C1: Structural Model Fit versus Data

	Market shares				
	Grid (1)	Diesel (2)	Own solar (3)	Micro-grid (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	1	31
Panel B. Endline					
Raw data	25	3	4	9	40
Data with Laplace correction	24	3	7	10	43
Model	24	3	7	10	43
Panel C. Follow-up					
Raw data	43	3	15	3	64
Data with Laplace correction	40	3	17	5	65
Model	40	3	17	5	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.

C.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-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).

Table C2: First-Stage Estimates for Price and Hours of Supply IV

	Price & Hours IV		
	Price	Peak hours	Off-peak hours
	(1)	(2)	(3)
Treatment normal price × Endline (=1)	0.064**	0.0050	-0.0046
	(0.029)	(0.0051)	(0.030)
Treatment subsidy price × Endline (=1)	-0.16***	0.0075	0.014
	(0.021)	(0.0063)	(0.031)
Mean peak hours in nearby villages	-0.032	0.94***	0.19
	(0.045)	(0.063)	(0.15)
Mean off-peak hours in nearby villages	0.0040	0.032**	0.88***
	(0.0094)	(0.013)	(0.030)
ξ_{tj} mean effects	Yes	Yes	Yes
Observations	999	999	999
F-Stat	21.6	524.1	1057.2

This table presents the first stage of our preferred IV specification, corresponding to column 5 of Table 4. Column 1 shows first-stage regression results for price. Columns 2 and 3 present result for the first stage regressions on peak and off-peak hours of supply, respectively Each outcome variable is an endogenous variable that we instrument for in the IV estimations. 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.3 Estimates of heterogeneity in demand by household characteristics

Table C3 presents the marginal effects of household characteristics on electricity choice probabilities. The observable variables enter the model for household choices through equation (6). The marginal effects calculated will of course also depend on the other model parameters. We

Table C3: Electricity Source Choice Probabilities by Household Characteristics - Median Households

	Full Model				
	Grid (1)	Diesel (2)	Own Solar (3)	Microgrid (4)	None (5)
Panel A. Marginal effects of household characteristics					
Number of adults	0.228	0.013	-0.017	-0.019	-0.205
	(0.042)	(0.007)	(0.008)	(0.014)	(0.031)
Solid roof (=1)	0.174	0.008	0.006	-0.022	-0.165
	(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	3.09				
	(0.173)				
Observations	8822				
Objective Value	.0965				
Wald Test Statistic (p-value)	600.66 (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 the full model, which includes as covariates the number of adults in the household, a dummy variable for whether the household has a solid roof, whether the household owns agricultural land, years of education of the household head and household income. The model also includes a random coefficient on price with a uniform distribution. The estimated dispersion parameter of the price distribution (equal 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 endline availabilities, qualities, and prices of each good. 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 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.

evaluated marginal effects for a "median" household, evaluated at the time of the endline survey. The marginal effects are not strictly marginal; for binary variables, we report the effect on each choice probability of changing the value from zero to one, and for continuous variables, the effect of a one-standard-deviation increase.

Table C4 shows detailed summary statistics for the household covariates that enter our demand

Table C4: 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 ('000s INR)	7.69	6.26	0	4.25	6	8.75	65	2917

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

model. A median household has 3 adults, a solid roof, owns land, and has a household income of INR 6,000 per month. Because some of these variables are binary, and the value for the median household is already one, we do not evaluate marginal effects by increasing their value from 1 to 2 (which is infeasible) but rather from 0 to 1. Table C5 gives the changes in household covariates that we use to evaluate "marginal" effects, accounting for the discreteness of these variables.

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

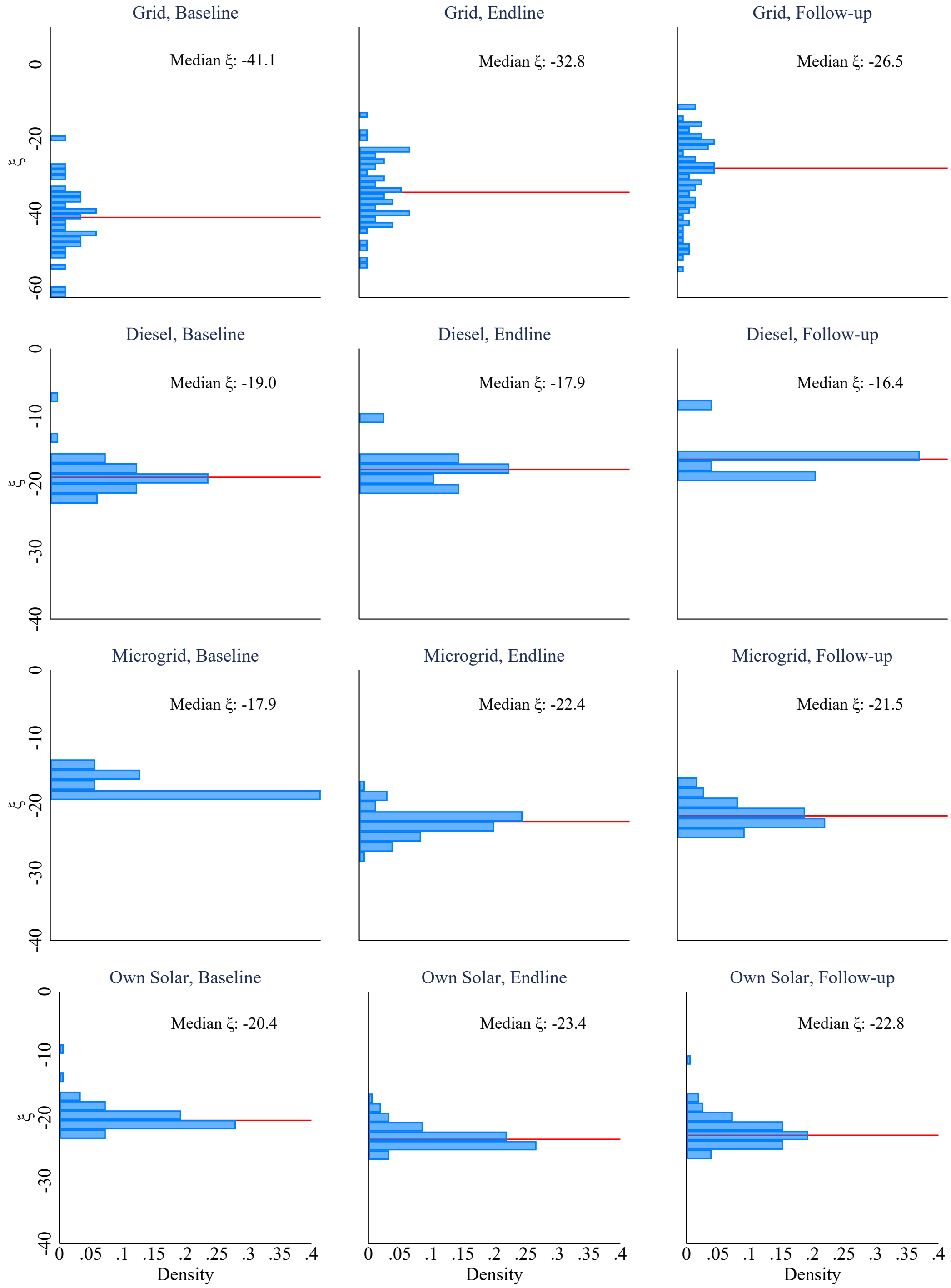
Characteristic	Definition	Marginal Change (Poor)
Adults	Adults in the household	1 SD (1.88 persons)
Roof	Indicator for solid roof	0 to 1
Land	Indicator for agricultural land	0 to 1
Education	Education of household head (1-8)	1 SD (2 levels)
Income	Monthly household income	1 SD (INR 6259)

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 C3). 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.

C.4 Estimates of the unobserved quality of electricity sources

Figure C1 plots the distributions of estimates of the unobserved quality of each source across survey waves. We measure the unobserved quality as the residual from Equation 5 that fits source market shares given the observed characteristics of each source j in village v at survey wave t. We normalize the value of the mean unobserved utility by setting the mean utility of the outside option to zero. Because these are the mean *unobserved* components of quality, the values do not account for observed differences in the mean utility of a source driven by price or hours of supply. The unobserved quality of most off-grid sources is roughly stable over the sample, but that of the grid is steeply increasing. This increase in the appeal of the grid may reflect the government's outreach and subsidies for new grid connections as part of the Saubhagya electrification campaign.

Figure C1: Evolution of Electricity Supply Quality by Source



The figure plots the estimated distributions of unobserved source quality for all electricity sources over time. The four rows are for different electricity sources, from top to bottom: grid electricity, diesel, solar microgrids, and own solar systems. The three columns are for the survey waves, from left to right: baseline (starting November 2013), endline (starting May 2016) and follow-up (starting May 2017). Each histogram in the figure shows the distribution across villages v of unobserved mean quality $\widehat{\xi}_{jtv}$ for the row source j during the column survey wave t. The vertical axis is the value of mean unobserved quality, where the outside option is normalized to zero, and the horizontal axis is the density. The mean unobserved quality is estimated in the demand model as the residual that fits source market shares given the observed characteristics of each source.

C.5 Robustness of demand estimates using payment data

Table C6 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 fieldwork that the pace at which HPS repossessed systems for non-payment was slow.

Table C6: Solar Microgrid Demand by Village Treatment Arm

	Administrative		
	Baseline (1)	Endline (2)	Paid ever (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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.6 Additional counterfactual results

Table C7: Breakdown of Counterfactual Analysis

	Market shares					Surplus (INR per hh per year)		
	Grid (1)	Diesel (2)	Own solar (3)	Microgrid (4)	All (5)	Consumer (6)	Producer (7)	Total (8)
Panel A. Model market shares and surplus by survey wave								
1. Model at baseline	6	17	7	1	31	385	-108	277
2. Model at endline	24	3	7	10	43	1064	-492	572
3. Model at follow-up	40	3	17	5	65	1886	-842	1045
Panel B. Disruption due to the improvement of solar power, relative to baseline								
1. Model at baseline, no solar	6	18	0	0	24	345	-109	236
2. ... + solar available (follow-up ξ)	5	14	28	1	48	536	-105	431
3. ... + solar observables improved	5	13	23	8	50	570	-104	467
Panel C. Disruption due to the improvement of grid electricity, relative to baseline								
1. Model at baseline, no grid	0	19	8	1	28	246		
2. ... + grid extent improved (follow-up ξ)	16	14	6	1	37	840	-308	532
3. ... + grid observables improved	37	9	4	1	51	1605	-766	839

The table presents market shares and surplus under additional counterfactual scenarios that break down improvements in solar and grid electricity in Panel B and Panel C in Table 6. Panel A presents market shares and surplus for each survey wave at data for reference. Panel B presents changes in market shares and surplus due to 1. the introduction of solar (Panel B, row 2) keeping price and hours of supply constant and 2. improvement in price and hours of supply (Panel B, row 3, compare to Panel B, row 2, in Table 6). Similarly, Panel C presents changes in market share and surplus due to 1. improvement in grid extent (Panel C, row 2) keeping price and hours of supply constant and 2. improvement in price and hours of supply (Panel C, row 3, compare to Panel C, row 2, in Table 6). For Panel B, row 2, and Panel C, row 2, we assign all villages unobservable quality (ξ_{jvt}) of solar and grid, respectively, at the follow-up level. For villages, where a given source was not available at baseline, we assign price and hours of supply at the baseline median for that source (Panel B, row 2, and Panel C, row 2). All counterfactuals are calculated using the full demand model estimates of Table C3, columns 6 through 10. Levels rows are unindented, whereas changes rows (where the numbers displayed are differences in two counterfactual scenarios) are indented.