DEMAND FOR ELECTRICITY ON THE GLOBAL ELECTRIFICATION FRONTIER

By

Robin Burgess, Michael Greenstone, Nicholas Ryan and Anant Sudarshan

February 2020

COWLES FOUNDATION DISCUSSION PAPER NO. 2222

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Demand for Electricity on the Global Electrification Frontier*

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

February 27, 2020

Abstract

Nearly a billion people, mostly in rural Africa and South Asia, do not have electricity at home. The advent of off-grid solar power means that many of these households, at the frontier of global electrification, have a choice between competing sources of electricity. This paper studies the demand for electricity with a discrete choice model wherein households choose between grid electricity, several off-grid electricity sources, and having no electricity at all. The model is estimated using a randomized experiment that varied the price of solar microgrids for a sample of villages in the state of Bihar, India, an outpost on the global electrification frontier. There are three main findings. First, households value electricity, but demand for any one electricity source is highly elastic, because several sources provide similar energy services at similar prices. Second, even in a relatively poor, rural sample, richer households greatly prefer grid electricity. Third, future growth in income will drive an increase in electrification due mainly to new grid connections, even if the cost of solar continues to fall. Our analysis suggests that off-grid solar power provides an important stop-gap technology, which fills the space between having no electricity at all and grid electricity, but that the future will run on the grid.

JEL Codes: O13, Q41, Q21, C93

1 Introduction

The global electrification frontier is the collection of places in the world where, at a given time, households are getting electricity for the first time. The steady movement of this frontier, in the United States from 1935 onwards, Brazil from the 1960s, China in the 1980s, and much of South Asia and sub-Saharan Africa in the 2000s up through today, is inseparable from economic growth.[1]

*We thank Manoj Sinha and Col. Baljit Singh of Husk Power Systems for their partnership in implementing the project. We thank Jennifer Allard, Shruti Bhimsaria, Menna Bishop, Catherine Che, Ananya Kotia, Rakesh Pandey, Aditya Petwal, Victor Quintas-Martinez, Johanna Rayl and Rashi Sabherwal for excellent research assistance. We thank the Shakti Sustainable Energy Foundation, the LGT Venture Philanthropy Foundation, USAID and the International Growth Center (IGC) for financial support. 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 comments. All views and 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 [Energy Policy Institute at Chicago, University of Chicago: anants@uchicago.edu]

[1]Electrification is also a cause of growth in terms of both economic output and human development (Lipscomb, Mobarak and Barham, 2013; Kitchens and Fishback, 2015; Lewis and Severnini, 2019).

Many developing countries and multilateral agencies are making huge investments in grid extension and in subsidies for household connections, to reach the roughly one billion people around the world who are not connected to the electricity grid (International Energy Agency, 2017).

The advent of solar power has changed the shape of the global electrification frontier. In the traditional mode of electrification, the frontier was a literal boundary, defined by the extent of the grid, with households filling in behind it (Lee et al., 2014). Solar panels, of course, can supply the grid, but, unlike other sources of power, they can also generate efficiently at a small scale, on the roof of a single, isolated household. The frontier today has therefore dissolved and permeated rural areas: a rapid decline in the cost of solar panels has opened up a new mode of electrification, whereby every household can individually choose whether to get solar power, regardless of whether the grid has reached them, and, even if the grid is present, households may choose another source of power instead. The ready nature and falling cost of solar have spurred hope of a faster, greener path to universal electrification.[2] This optimism has some justification, since off-grid solar's market share on the global electrification frontier has skyrocketed in the last decade (Figure 1).

This paper estimates the demand for electricity, from *all available electricity sources*, in the state of Bihar, India, an outpost on the global electrification frontier. Between 2000 and 2016, India contributed over 80% of the total gain in the number of households electrified in the world (International Energy Agency, 2017). We collect comprehensive data on the demand and supply sides of the electricity market in 100 villages, during a four year period in which there was an almost *40 percentage point* increase in electrification in our sample.[3] We document this surge and argue that it was due to the same two factors, namely investments in the grid and a sharp decline in solar prices, that are reshaping the global frontier (see Section 2).

We study this transformation of Bihar's electricity market in three steps. First, we introduce an experiment that varied prices for solar microgrids, at the village level for two and a half years, and use

[2] Former UN Secretary General Ban Ki-moon proclaimed "Developing countries can leapfrog conventional options in favor of cleaner energy solutions, just as they leapfrogged land-line based phone technologies in favor of mobile networks." ("Powering Sustainable Energy for All," *The New York Times*, January 11th, 2012. See also "Africa Unplugged: Small-scale Solar Power is Surging Ahead", *The Economist*, October 29th, 2016.) UN Sustainable Development Goal #7 is to "ensure access to affordable, reliable, sustainable, and modern energy" and targets increasing the share of renewable energy in the global energy mix in particular. Nearly all large-scale aid programs in the power sector include significant on-grid and off-grid components. USAID, for example, launched *Power Africa* in 2013 and DFID launched *Energy Africa* in 2015, both of which invest in off-grid renewable electricity.

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

the experiment to estimate the demand curve for microgrids alone (Section 3). Second, we specify a structural model of demand, for all sources of electricity, and estimate the model using the same experimental variation (Section 4). Third, we apply the model to run counterfactual simulations that study how the surplus from electrification depends on technology and policy (Section 5).

In the first step, we find that demand for solar microgrids is highly elastic. In the experiment, under the treatment arm that introduces solar microgrids at their market price, 6% of households buy them, while under the subsidized treatment arm, which offers microgrids at half of their market price, 19% of households buy. Despite the rapid and beneficial adoption of solar microgrids during the experiment, microgrid demand had collapsed by one year after the experiment ended, both in subsidized villages, where the removal of experimental subsidies caused prices to rise, and in market price villages, where prices remained nearly constant.

Why did households give up on solar microgrids? We argue that the reason has to do not with the product, but the market. In stark contrast to the developed world, electricity in Bihar is a differentiated product: there are many sources of electricity, which differ in price, load, hours of supply and other features. The main substitutes for microgrids, namely the grid and individual household solar systems, were improving over the period of our study, and may have taken microgrid market share. The dependence of microgrid demand on the availability of substitutes would affect both the external validity of our microgrid demand curve, and our interpretation, within the context of the experiment, of how much the poor value electricity.

In the second step, therefore, we estimate a model of household demand over all electricity sources. We model the demand for electricity connections as a discrete choice demand system (McFadden, 1974; Lancaster, 1971), where households choose between grid electricity, common diesel generators, solar microgrids, their own solar systems and an outside option of no electricity. We allow for substantial observed heterogeneity in source characteristics and household valuations for different sources. We also allow the unobserved quality of electricity sources to vary without restriction across villages and time (Berry, 1994). In our setting, this feature is essential, since we expect source quality is changing in ways that are not easy to measure. For example, the government made great efforts in our study period to sign up new households for grid connections, which might have lowered household connection costs. Solar systems were also improving on dimensions like price, availability and reliability.

We estimate the model using our household-level panel data for one hundred village-level markets over three survey waves spanning four years. We leverage the experimental treatment assignments as instruments for price, to account for the endogeneity of price to unobserved aspects of product quality. The availability of experimental price variation to estimate a discrete choice demand model is extraordinarily rare and removes the need to rely on traditional assumptions, about market conduct or the structure of demand shocks, to generate instrumental variables (Berry, Levinsohn and Pakes, 1995; Hausman, 1996; Nevo, 2001).

There are two main findings from the demand estimates. First, household demand for electricity is price elastic, especially for off-grid technologies like diesel and solar, for which we estimate own-price elasticities of around negative two.[4] An ancillary finding is that the experimental variation is necessary to recover unbiased and precise estimates of the price elasticity of demand.[5]

Second, even in our relatively poor sample in Bihar, richer households strongly prefer grid electricity. Whether a household has a solid roof is commonly used to measure wealth in development economics (Alatas et al., 2012; Haushofer and Shapiro, 2016). In a simple version of our demand model, with few covariates, we find that an indicator for solid roof ownership increases the probability that a household chooses grid electricity by 0.208 (standard error 0.039), or 21 percentage points (pp), on a base of 24 pp. This overwhelming preference of richer households for the grid seems sensible, since the grid is the only source that can support the higher-load appliances, like fans and televisions, that richer households demand. In our full demand model, we include many household observable characteristics, and find that household demand for the grid is increasing in a wide set of income and wealth proxies.

The third part of the paper applies the demand model to study counterfactual scenarios that vary aspects of technology, policy, and household characteristics, to predict their effects on electricity access and social surplus. We are interested in how changes in policy change the sources of electricity that households would choose to buy for themselves, as opposed to the sources their governments

[4] These large elasticities are found because, when the price of off-grid sources rise, households substitute to both the grid and the outside option. Grid demand is less elastic (-0.6), but still responsive to price in absolute terms: a INR 10 increase in the monthly price of grid services, enough to buy two cups of tea or three bananas, reduces the grid's market share by three percentage points (on a base of 24 percentage points).

[5] The experimental estimates of the price coefficient are negative, large and stable across specifications. If we instead estimate the price coefficient using ordinary least squares, it is negative and precisely estimated, but smaller than the experimental estimate by a factor of seven. Traditional instruments from the industrial organization literature are found to have no power in our setting (Berry, Levinsohn and Pakes, 1995; Hausman, 1996; Nevo, 2001).

may promote.

We draw three main findings from the counterfactual analysis. First, while households value electricity, no one source of electricity is indispensable. We counterfactually remove the grid, or all off-grid solar technologies, from the market, by turns, and find surplus losses in the neighborhood of 30% of the total value of electrification. These scenarios are extreme, but arguably have modest effects on surplus, because households can freely substitute between several sources that offer similar bundles of basic energy services at similar prices. Experimental estimates of microgrid demand therefore greatly understate the total surplus from all electricity sources: our estimate of total surplus from electricity is larger than our estimate of surplus from microgrids by a factor of 5. Moreover, our model estimates imply that this finding is not unique to microgrids—the surplus from *any* single source, including the grid, greatly understates the total surplus from electrification.

The second finding of our counterfactuals is that the state places a high value on electricity access and supply for the rural poor. In our setting, producer surplus is negative, since the state loses money from providing power below cost. The publicly-owned utility's losses, gross of state subsidies, are about as large as total consumer surplus from all sources of electricity. In counterfactuals, we find that intensive and extensive margin increases in supply, while raising electrification rates, increase the losses of the state utility by around 3 times as much as they increase household surplus. Yet the state, in our data, is rapidly expanding supply on both of these margins. Such an expansion lowers social surplus, but may be justified if the government has a preference for redistribution, valuing the surplus of the rural poor more than its own losses (Burgess et al., 2020), or expects benefits from electrification that are external to household demand (Lipscomb, Mobarak and Barham, 2013).

The third counterfactual finding concerns the medium-run future of electrification in Bihar. We consider a counterfactual that makes the following simultaneous changes: all households achieve at least the 80th percentile of current income and wealth in our sample; the grid expands to all villages; peak hours of supply increase to the feasible maximum of five hours; and solar costs decline to projected 2022 levels. These changes are large but plausible in our dynamic context; for example, per capita income in Bihar has already been growing by 10% per annum in the last five years. In this scenario, the electrification rate surges by about 40 pp. There are very large increases in both household surplus and producer losses. Our key finding from this forward-looking scenario is that, if the grid is available everywhere, growth in household income drives gains in electrification through the grid alone: *all* new electricity connections, on net, are grid connections, even as the cost of solar falls, because as households get richer their preference for the grid grows stronger. This finding argues that off-grid solar power is an important stop-gap technology, which occupies the space between having no electricity at all and grid electricity, but that the future will run on the grid.

Our paper makes two contributions to a vibrant literature on electricity access in developing countries. First, we study household demand, a revealed-preference measure of the value of electricity, whereas most of the literature has measured the impact of electricity access on a range of economic and welfare outcomes.[6] Second, we estimate how households value both grid and off-grid electricity together, in a single demand system, which allows us to study substitution between sources. Other work on the demand for electricity has considered electricity sources in isolation.[7] We show that substitution between sources is critical to understanding Bihar's electricity market. With both the grid and off-grid solar becoming ever more widely available in South Asia and Sub-Saharan Africa, we expect that such household choices, between electricity sources, are dictating the pace and type of electrification all along the global electrification frontier.

Our study joins a methodological movement in the development literature that combines structural models with experimental variation to aid in the interpretation and increase the external validity of experimental results.[8] Our experiment varied prices at the village level for two and a half years and we collected data over four years, removing concerns about the external validity of experiments that vary prices artificially, at the household level, or for only a short time. Our finding on the large gap between the value of electrification and the value of any individual electricity source shows the value of a structural model for placing experimental estimates of demand, for any particular product, in the context of the broader product market. Experimental work that estimates the

[6] Prior work has found that electricity access causes large increases in industrial output (Rud, 2012; Allcott, Collard-Wexler and O'Connell, 2016), labor supply (Dinkelman, 2011), and proxies for household welfare, such as the human development index, indoor air quality and kerosene expenditures (Lipscomb, Mobarak and Barham, 2013; Barron and Torero, 2017; Aklin et al., 2017). See Lee, Miguel and Wolfram (2020) for a recent review of the impacts of electrification on development. Our focus on demand is complementary to these analyses; demand is broader then any single measure of household well-being, though well-being may not be fully captured by household demand, for example, if households are not unitary, or there are externalities from electricity use.

[7] Lee, Miguel and Wolfram (2016) estimate household willingness to pay for grid connections in Kenya. Grimm et al. (2019), in contemporaneous work, estimate demand for off-grid solar technologies in Rwanda. Aklin et al. (2018) study how household characteristics predict solar take-up in India.

[8] There are several prominent examples of work that uses experiments to help estimate structural models of education, labor supply and migration, though not demand for electricity (Attanasio, Meghir and Santiago, 2011; Duflo, Hanna and Ryan, 2012; Bryan, Chowdhury and Mobarak, 2014; Galiani, Murphy and Pantano, 2015). Field experiments have lately gotten longer to address the realism and durability of effects (De Mel, McKenzie and Woodruff, 2013; Dupas and Robinson, 2013; Bandiera et al., 2017).

demand for any one product, like a particular bed net or loan, at a time, may similarly understate the demand for the categories to which these products belong.[9] The structural model enables us to decompose the gains from electrification, during a period of transformative change, and to study policy counterfactuals that extend beyond the boundaries of the experiment itself.

2 Background and Data: The Electricity Landscape in Bihar

This section introduces our setting and data sources and then uses that data to describe the electricity market in Bihar, India. Bihar is a state of 104 million people (Census of India, 2011), about the population of Ethiopia or the Philippines. Bihar is one of India's poorest states and, at the start of our study period, had very low access to electricity.

Table 1 juxtaposes the United States, India, sub-Saharan Africa and Bihar on the dimensions of GDP per capita, electricity consumption per capita and access to electricity circa 2012 (our baseline survey was conducted at the end of 2013). The electrification rate in Bihar at this time was only 25%, below the rate of 37% in sub-Saharan Africa and about one-third of the all-India rate of 79%. The average Bihari used just 122 kWh of electricity per year, less than one percent of the level in the United States (column 4, last row). At this low level of consumption, which averages over many households with no electricity at all, an individual can power two light bulbs totaling 60 watts for six hours per day year round. The low level of average consumption is an equilibrium outcome. Demand for electricity is low because many households are poor. Supply of electricity is limited, on both the extensive margin, since many villages are not on the grid, and the intensive margin, since supply is rationed.

Our study, luckily, was well-timed to capture two big changes in the electricity market. First, in response to persistently low rates of household electrification in states like Bihar, the Government of India funded large campaigns for village-level grid extension and household-level connections, respectively. In his 2015 independence day address, Indian Prime Minister Narendra Modi launched

[9] Experimental estimates of demand have been an enormous area of growth in development economics and are used both to test theories of behavior and to consider optimal policy. Preventive health products (Berry, Fischer and Guiteras, forthcoming; Peletz et al., 2017; Dupas, 2014) and financial services (Bertrand et al., 2010; Karlan and Zinman, 2018, 2019) are two prominent markets in which experiments have been used to estimate demand. Though many products in these markets arguably have close substitutes, few studies that experimentally estimate demand explicitly model substitution, as is necessary to value surplus from an entire category of 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.

a rural electrification program with a thousand-day deadline to electrify the remaining 18,452 census villages still without access, at an estimated cost of USD 11 billion.[10] The village-level goal was declared achieved ahead of schedule on April 28, 2018. 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 infill household connections in electrified villages.[11] States receive this federal money and use it to provide connections; nearly all grid electricity in India is supplied by state-run utilities. Besides the Government of India's efforts, the Bihar state government has made electricity access a priority (Kumar, 2019). Nitish Kumar, Bihar's six-time Chief Minister, promised universal household electrification as part of his reelection campaign (Business Today, 2017).

The second big change, beginning before and carrying on through our study period, was a result of the continued, dramatic decline in the price of solar panels. After 2012, this decline reached a point where off-grid solar, which had long been too expensive for rural customers, became a feasible alternative to grid power. The share of households with their own solar systems, as well as solar generation on the grid, began to grow exponentially (Figure 1).

a Data

We collect data from both the demand and supply sides of the market in a sample of 100 rural villages over a nearly four-year period.

Our sample consists of 100 villages in two districts in Bihar (Figure 2).[12] The study villages were chosen to have low rates of electricity access, along three criteria. First, they were not listed as electrified villages by the government, meaning that household grid electrification was below ten percent and at least one neighborhood of the village was not on the grid at all. Second, as we worked with a solar microgrid provider, Husk Power Systems (HPS), to offer solar microgrids, villages must not yet have been offered HPS microgrids. Third, to facilitate a possible expansion of microgrids,

[10] A village is defined as electrified once public spaces, such as schools and health centers, have access to electricity, along with a minimum of 10% of its households. The target is out of a total of almost 600,000 census villages in India. This program, the Deen Dayal Upadhyaya Gram Jyoti Yojana (DDUGJY), is a continuation, under a new name, of the prior government's Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY), which had similar objectives but fell short of reaching all villages (Government of India, 2015; Burlig and Preonas, 2016).

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

[12] 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.

villages were chosen to be reasonably close to existing HPS sites. We selected 100 villages that met these criteria, which have a total of 48,979 households.

We collected data from four sources. First, on the demand side, a household-level panel survey on the sources and uses of electricity. Second, on the supply side, household-level administrative data on customer enrollment and payments from a provider of solar microgrids. Third, on the supply side, village-level survey data from the operators of common diesel generators, an important off-grid source of electricity. Fourth, on the supply side, household-level administrative data from the state utility on customer billing and payments, as well as village-level electricity supply. We ran a separate, contemporaneous project with the state utility that allowed us to gain access to this data. The timing of collection for each source of data is illustrated in Appendix Figure A1. We now describe each of these data sources.

Household panel survey. Our household panel survey sampled 30 households per village to cover about 3 thousand households, containing about 18 thousand people, across the 100 sample villages. The sample was drawn to represent those with an interest in a microgrid solar connection, but, because this screening for interest was loose, in practice the sample is nearly representative of the population as a whole.[13] The survey has three rounds, two thick rounds, which we call baseline and endline, and one thin round, which we call follow-up. The baseline survey took place in November and December of 2013, the endline from May to July of 2016, and the follow-up in May 2017 (Appendix Figure A1 shows the timing of survey rounds).

The two thick rounds used nearly the same survey instrument and covered demographics, the sources and uses of electricity, and welfare outcomes likely to be influenced by electricity use. There are three main kinds of variables. First, demographics and household characteristics, such as household size and literacy, as well as wealth proxies, such as income, house type, roof type and ownership of agricultural land. We use these variables to predict electricity demand. Second, variables on electrification status, sources of electricity, source characteristics such as hours of supply, payments, and uses of electricity, including a complete appliance inventory. Third, variables on

[13] 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.

education and health. We gave children reading and math tests and asked households about any respiratory problems.[14]

The thin, follow-up survey round took place one year after the endline for the experiment and was not part of our original plan. The purpose of this round was to update household electricity sources and choices, in light of the massive changes we observed on the supply side; as such, the follow-up round left out household characteristics, education and health outcomes. The baseline and follow-up survey rounds are separated by nearly four years.

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

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

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

These sources of data allow us to see, on the demand side, a rich set of household characteristics and the sources and uses of electricity. On the supply side, we have data on all the competing

[14] Prior research suggests that substituting from kerosene to electricity reduces indoor air pollution (Barron and Torero, 2017). It has been widely hypothesized that children may benefit from using electric light to study (United Nations Department of Economic and Social Affairs, 2014).

sources in the marketplace, in some cases from both administrative sources and our household and operator surveys.

b Sources of electricity

In developed countries, electricity is the archetype of a homogeneous good. In Bihar, and many developing countries, electricity connections are differentiated products. This section describes the characteristics of different electricity sources. Unless otherwise specified, we describe source characteristics at the time of our baseline survey.

Table 2 reports the characteristics of electricity sources at baseline (columns 1 through 5), endline (columns 6 through 10) and follow-up (columns 11 through 15). The main source of data is the household survey. Grid supply hours come from the administrative data.

Grid electricity. Grid electricity, the traditional mode, has a desirable bundle of characteristics (Table 2, Panel A, column 1). The grid price of INR 72, the mean monthly payment reported by grid-connected households, is roughly tied for the lowest price of any electricity source. Households on the grid have a mean connected load of 322 watts, 30% larger than the second-highest load source (the first set of columns, comparing columns 1 through 4).[15] Grid households, correspondingly, own more and larger appliances (Table 2, Panel B). Nearly all surveyed households, and across most sources of electricity, own light bulbs and mobile phones, which is to be expected as lighting and phone charging form the essential bundle of energy services that households want from electricity. Among grid-connected households, in addition, 22 percent own a fan and 15 percent a television; the ownership of these appliances for households with other electricity sources is much lower or even negligible.

We take the grid price to be the self-reported monthly payment for grid electricity, averaged across formal and informal households on the grid. 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 full grid price of INR 153 per month at baseline, if everyone paid their bills, would place it amongst the most costly sources, while at INR 72 per month, it is one of the cheapest.

[15] Properly, connected load is not a characteristic of a source, but depends on household appliance purchases. We describe connected load as if it were a source characteristic, because the connected load for all sources but the grid is effectively capped by the load a source can support.

There are two main drawbacks of the grid. First, the grid is only present in 29% of our sample villages at baseline. Second, for households on the grid, the lights are often not on at night. The grid supplied power for an average of 10.9 hours per day at baseline, longer than any other source. However, mean supply in the peak hours, from 5 to 10 pm — when people are at home and want power the most — was only 2 hours per day at baseline and endline, increasing to 3 hours at follow-up. While this is the average level of supply, on many days there is no grid supply at all (Appendix Figure A2 shows the distributions of hours of supply for the grid). All alternative sources of power provide more supply during the peak hours in all survey waves.

Diesel generators. Diesel generators are too big for households to buy on their own. Entrepreneurs buy generators and wire certain villages, where they anticipate enough demand, to supply diesel power on fixed-price plans. At baseline, 57% of villages had generators. Generators run on a predictable schedule during the evening peak hours for an average daily supply of 3.4 hours (Table 2, Panel A, column 2). Diesel generators have a mean price of about INR 100 per month, which is cheaper than the full price of the grid, but a third more expensive than the effective grid price. The modal diesel plan offered, by far, was a 100 watt connection for INR 100 per month. The 100 watt limit, enforced by wiring a fuse onto the service wire, is enough for consumers to power several light bulbs and charge their mobile phones (Panel B, column 2).

Own solar. Households in all villages can choose to buy their own solar systems in private markets. Households pay for own solar systems up front and would usually have to travel to a market town to buy a system. A system consists of a panel, a battery, and sometimes a socket to plug in and switch appliances. We refer to this source as "own solar" to distinguish it from solar microgrids, which serve several households together. On observable characteristics, own solar is a fairly close substitute to diesel electricity (Table 2, Panel A, column 3). Own solar systems have a similar price to diesel and power appliances of similar load.[16] Solar systems at baseline run 8 hours per day, longer than diesel generators, but somewhat less than the grid.

Microgrid solar. A solar microgrid has the same basic components as an own solar system but operates at a slightly larger scale; it serves six to nine households at a time. Households can therefore only connect if they also have interested neighbors (as with a diesel generator, but

[16] Once purchased, own solar systems 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 20% interest rate.

even more locally).[17] The microgrid offers broadly similar energy services to other off-grid sources. Because of their low voltage, microgrids are unable to power small appliances such as fans (Table 2, Panel B, column 4). The prevailing price of microgrids at the start of our study was INR 200, making a microgrid the most expensive source of electricity (Table 2, Panel A, column 4). This high price was later cut, in some places, to INR 160 and was further subsidized (to INR 100) as part of our experiment (Section 3 a). One advantage of microgrids, relative to own solar systems, is that the provider, rather than the household, is responsible for set-up and maintenance.

c Bihar's electricity transformation

The electricity landscape in Bihar was transformed during our study, by a surge in grid extension and connections and a fall in the cost of solar power. Here we describe how these disruptions, on the supply side, changed household electricity access and technology choice.

These two disruptions are plainly visible in our data. The government, under the electrification programs described above, extended the grid from 29% to 72% of sample villages, held camps to connect more households, and heavily subsidized connections, including by offering connections for free to all households designated Below the Poverty Line (BPL). With regard to solar, by 2022, the US National Renewable Energy Laboratory projects a 55% reduction in the cost of solar photovoltaics and the US Department of Energy targets a 75% reduction in the cost of batteries (Feldman, Margolis and Denholm, 2016). Our data reflects these trends. The price of own solar systems fell 10%, from INR 80 at baseline to INR 72 at follow-up. An important caveat is that this lower price is likely not for the same energy service, but a better one. Solar vendors penetrated smaller towns closer to villages, effectively lowering connection costs. Solar panels may also have gotten more efficient and batteries more reliable. We return to discuss changes in source quality with the demand model estimates.

Figure 3 shows the market shares of all electricity sources over time. Each stacked bar gives the share of households, from bottom to top, that use grid electricity, diesel generators, solar microgrids, own solar systems or no electricity. These market shares are calculated with respect to the total

[17] The microgrids offered by HPS consist of a 240 watt panel and a separate, 3.2 volt rechargeable battery and meter for each household. Households have a key pad to secure access to the battery and must purchase codes on a monthly basis to keep using the system. Each household on the microgrid gets 25 to 40 watts of power. To compensate for the small load, the system is bundled with two high-efficiency light bulbs and an electrical outlet, typically used for mobile phone charging, and therefore provides very similar energy services to diesel and own solar systems.

sample, without regard for whether a source is available in a village or not; in a village where the grid is not present, for example, the grid necessarily has a zero share. There are three clusters of bars, for shares in the baseline, endline and follow-up survey waves. Within each cluster of bars, the three bars from left to right give the market shares amongst households that do not have a solid roof, all households, and households that do have a solid roof, respectively.

The figure shows that the two disruptions passed through to household choices. Consider the middle bar in each group, for all households. The electrification rate, from any source, is the sum of the colored bar stacks, or one minus the share of households with no electricity. The electrification rate increased 37 pp, from 27% to 64%, in somewhat less than four years, a transformation in energy access.

The net gain in electrification conceals an enormous churning of market shares across sources. 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. In other words, the grid electrification rate almost doubled in the one year period between our endline and follow-up 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 microgrids (third from the bottom, in yellow) also increased their share, from nothing to 9% at endline, when subsidies were still offered as part of our experiment, but fell back down a year later. Own solar systems (top colored bar, in orange) picked up the slack, rising from a 5% share at baseline to a 15% share at follow-up, with all of their growth coming between the endline and follow-up rounds.

Figure 3 also shows significant heterogeneity in household access to electricity within a given survey wave (cluster of bars). 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.

The transformation of Bihar's electricity sector thus has three aspects. First, a surge in aggregate electrification. Second, a compositional shift, away from diesel and towards solar power and especially grid electricity. Third, heterogeneity in household demand, with richer households more likely to have electricity from the grid.

3 Demand for Solar Microgrids

This section introduces our experiment and uses it to estimate demand for solar microgrids. The demand for microgrids is important, in its own right, because off-grid solar technology has emerged as a widespread substitute for grid electricity on the global electrification frontier. We start by estimating demand for this new good, to describe our experimental variation in microgrid prices and to allow us to quantify how much the poor value microgrids.

The demand for one source of electricity may be a bad proxy for the demand for electricity, on the whole, if there are close substitutes available for any given source, as we have argued is the case in Bihar's electricity market. In the following sections, therefore, we will specify a discrete choice demand model that covers all electricity sources, and estimate the model using the same experimental variation described here.

a Experimental design

We partnered with Husk Power Systems (HPS) to vary the availability and price of solar microgrids in a randomized control trial.[18] The experimental design is a cluster-randomized control trial at the village level. We randomly assigned sample villages into one of three arms: a control arm (34 villages) where HPS did not offer microgrids, a normal price arm (33 villages) where HPS offered microgrids at the prevailing price, initially INR 200 per month, and a subsidized price arm (33 villages) where HPS offered microgrids at a reduced price of INR 100 per month. The normal price arm can be thought of as providing the service at or slightly above cost and the subsidized arm at perhaps 40% below cost.[19] Within each treatment village, all households were offered the same HPS

[18]HPS was founded in 2007 to provide electricity in rural areas using biomass gasifiers as generators to obtain power from agricultural waste, such as rice husks (hence the name of the company). These biomass plants could only serve a village if demand was sufficiently broad and were subject to fuel supply disruptions. HPS made a strategic decision to add a solar microgrid product to its portfolio as a means of reaching a wider set of customers.

[19]We estimate the capital and installation costs of a microgrid to be INR 105 per household per month (Appendix Figure D3). 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 cost in the range of INR 160 to INR 200 per month, the range of prices offered in our normal

connection and pricing, regardless of whether they had previously expressed interest in a microgrid or participated in our baseline survey. Sales of solar 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, during our study period, for two reasons. First, the normal or prevailing 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, endogenously 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 at INR 170 per month. This price adjustment meant that 22 normal price villages experienced price declines of INR 30 (from 200 to 170); 11 normal price villages experienced a INR 10 increase; and all 33 subsidized price villages saw a substantial increase of INR 70 (from 100 to 170). HPS did not enter the control villages at any point during our entire study period. In the demand analysis, we use treatment assignments, and their interactions with survey wave indicators, as exogenous instruments for price.

Table 3 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.

Table 3, column 1 gives household characteristics in the control group. Because of the sample selection criteria, our sample is poorer than the population of Bihar as a whole. Self-reported household incomes imply mean per capita daily income of INR 43 (USD PPP 2.5) at baseline, compared to mean per capita daily income of INR 99 (USD PPP 5.8) across the state.[20] Two-thirds of households own agricultural land and less than half have a solid roof, constructed of an impermeable material like metal. House characteristics like building materials are commonly used in proxy means tests as indicators of household wealth and poverty (Alatas et al., 2012; Haushofer and Shapiro, 2016). The average household has 3.3 adults living in 2.4 rooms.

price arm.

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

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. Households in subsidy villages are more likely to have solid houses and solid roofs than control households. The overall rate of electrification does not differ by arm (an F-test for the joint equality of "Any elec source (= 1)" across treatment arms has p-value 0.54), but households in the subsidy treatment arm are more likely to have electricity from the grid and somewhat less likely to have it from a diesel generator. We address this slight imbalance by including household covariates as controls in our demand estimates.

b Results

We use the experiment to estimate the demand for solar microgrid connections. Table 4 presents estimates of intention to treat effects from a regression of whether a household had a solar microgrid on village-level treatment assignments. 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. These estimates are based on the survey data and reported at baseline (November 2013), endline (May 2016) and follow-up (May 2017) in columns 1 through 3, respectively.

The first main finding in Table 4 is that the experiment increased solar microgrid penetration. At baseline, in column 1, the estimated constant, representing take-up in the control group, and the estimated normal price and subsidized treatment coefficients are very small and statistically not different from zero. We expect there should be zero take-up at the baseline, because microgrids were a new product, about to be launched. At endline, in column 2, the estimated constant was 2.3 pp (standard error 0.5 pp), and 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.[21] We find a similar gap in estimated demand when using

[21] Virtually all of the take-up that is observed in the normal price treatment arm occurred at a price of INR 160, to

administrative measures of household payments, rather than surveys, to measure take-up.[22]

We build a demand curve from microgrid market shares, using the experimental treatment assignments as instrumental variables, to form a first-cut estimate of the value of microgrids to households. Appendix Table C5 shows estimates of linear and log-linear aggregate demand curves for microgrids. With these demand curve estimates, we calculate that removing microgrids from the market would reduce surplus by INR 83 to INR 129 per household per annum, averaging over all households in the population, regardless of whether they chose microgrids.[23] The surplus of INR 91, calculated from the linear demand curve estimates, is quite small, at 1.6% of household energy expenditure and a mere 0.1% of mean household income in our sample. The following section will consider the robustness of these surplus estimates and place them in the context of the value of alternative electricity sources.

Our household survey allows us to measure the sources of these surplus gains. The higher rates of solar microgrid take-up in the treatment villages lead to treated households getting more energy services. Appendix Table B2 shows that households in treatment villages use more hours of electricity per day, own light bulbs and mobile phones at higher rates, and pay less to charge their phones. Microgrids therefore provide the basic energy services that they promised. We also test for whether microgrids have impacts on a broader range of socio-economic outcomes, including children's test scores and respiratory distress. We do not see any statistically significant improvements on these welfare measures, though our analysis is underpowered for these outcomes, especially for test scores (Appendix Table B3).

The second main finding in Table 4, which motivates much of the subsequent analysis, is that

which HPS endogenously cut prices in 11 of the 33 normal price villages during the experiment. Indeed, take-up at INR 200, the normal price prevailing at the start of the study, was nearly zero throughout, which sets a tight range on the choke price, the price beyond which no household buys microgrid.

[22]We have administrative data from Husk Power that contains the monthly payment history of all eligible households. Appendix Table C11 repeats the demand analysis from Table 4 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.

[23]All surplus numbers we report are per household in the population, regardless of whether or not they choose a source. The surplus numbers for the hypothetical removal of microgrids are also understated in the sense that they give the average effect of removal relative to the status quo scenario at the time of the endline survey. In this status quo, microgrids are not present in the control group, one-third of sample villages, to begin with. Thus the removal of microgrids, by design, has no effect on surplus in those villages.

solar microgrid demand collapsed between endline and follow-up. 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%). This collapse in market shares may have several explanations. Plainly, in the subsidized arm, the increase in price after the experiment ended must have had a large effect. However, the proportionate decline in the normal price arm was just as large, suggesting that price changes alone do not explain the collapse in microgrid market shares.

Features of both the product and the market, in principle, may have contributed to the fall in microgrid market shares. On the product, it could be that microgrids are an experience good, which households discovered was no better than kerosene, or that microgrids were poorly maintained, leading to growing disuse over time (as Hanna, Duflo and Greenstone (2016) find for cook stoves). On the market, the descriptive evidence, from Section 2, that own solar and the grid grew rapidly during our study argues that these sources may have stolen market share from microgrids.

This substitution between multiple sources of electricity would affect both the external validity and the interpretation of our demand estimates. As a matter of external validity, household demand for the microgrid product might have been different, perhaps drastically so, if households had faced a different choice set, for instance, if the government had not made a big push for the grid or if the price of alternatives like own solar had not declined. On interpretation, our internally-valid estimates of microgrid demand cannot tell us household willingness to pay for electricity, even within the context of the experiment, when such close substitutes are available. Households may value electricity, but not buy microgrids, if they can instead buy another source of electricity they prefer. The short-lived market share gains we find for microgrids thus illustrate the risk in interpreting experimental estimates of demand, for a given product, in isolation from the rest of that product's market. To address these concerns, the next section introduces a model to jointly estimate household demand over all sources of electricity.

4 Model of Demand for All Electricity Sources

We model consumer demand for electricity using a discrete choice demand model over electricity sources. We specify a nested logit model (McFadden, 1978, 1980; Goldberg, 1995).

Several aspects of our empirical setting allow for an especially rich specification of the model and unbiased estimation of its parameters. First, we have a household-level panel survey of electricity choices and demographic and economic characteristics, which allow demand to depend on time-varying source and household characteristics. Second, we allow the unobserved quality of all electricity sources to vary without restriction across villages and time (Berry, 1994). Third, we observe one hundred separate markets in three time periods, and experimentally vary the price of one product, microgrids, across markets. We use the experimental variation to estimate the sensitivity of households to prices.

a Specification

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

$$U_{ijtv} = \delta_{jtv} + z_{it}'\gamma_j + \epsilon_{ijt} \tag{1}$$

$$= V_{ijtv} + \epsilon_{ijt} \tag{2}$$

The term V_{ijtv} is the strict utility of a choice for a household absent their idiosyncratic taste shock ϵ_{ijt}. The strict utility depends on the average utility of a source δ_{jtv} as well as a vector z_{it} of observable household characteristics. These characteristics affect household utility through source-specific coefficients γ_j. For example, households with higher incomes may have a greater preference for grid electricity, but an unchanged preference for diesel.

The term δ_{jtv} represents the mean utility of an electricity source j in village v at survey wave t. Mean utility depends on observable source characteristics x_{jtv} and unobserved source quality ξ_{jtv}, according to

$$\delta_{jtv} = x_{jtv}'\overline{\beta} + \xi_{jtv} \tag{3}$$

The vector x_{jtv} of observable source characteristics includes price, hours of supply on-peak (from five to ten pm), and hours of supply off-peak. We refer to ξ_{jtv} as unobserved quality or just quality. Unobserved quality is known to households but not the econometrician. It may include both unmeasured physical characteristics, such as the capacity of a solar system battery, as well as characteristics of the service, such as the monetary or hassle costs to obtain a connection from a given source.

The choice probabilities in the nested logit model take a simple form.[24] Each electricity source j belongs to a nest, indexed by g. The parameters σ_g measure the similarity of sources within a nest. The inclusive value of nest g is defined as

$$IV_{igtv} = \ln \sum_{j \in \mathcal{J}_g} e^{V_{ijtv}/(1-\sigma_g)}$$

which is the expected indirect utility when maximizing utility across sources in nest g. The probability of i choosing a source j in nest g_j is then

$$\Pr(y_{it} = j | z_{it}) = \frac{e^{V_{ijtv}/(1-\sigma_{g_j})}}{e^{IV_{ig_j tv}\sigma_{g_j}} \sum_g e^{IV_{igtv}(1-\sigma_g)}} \tag{4}$$

Choice probabilities differ by household because they depend on household characteristics z_{it} via the strict utility term V_{ijtv}. Market shares in the model are defined as the average of household choice probabilities across households in a village.

b Estimation

We estimate the model in two stages. The first, non-linear stage estimates the parameters of equation 1 via maximum likelihood with data from all three surveys. The second, linear stage estimation uses the $\hat{\delta}_{jtv}$ from the first stage as the dependent variable to estimate equation 3 using two-stage least squares. This two-step procedure is common to address endogeneity in the estimation of random coefficients logit models (Berry, Levinsohn and Pakes, 1995, 2004), of which the nested logit is a simple case. The key idea is to invert market shares to solve for mean indirect utilities, which then allows for linear IV estimates in the second stage that are unbiased in the presence of the endogeneity of price to quality (Berry, 1994).

[24] The nested logit assumption imposes that households' idiosyncratic tastes for electricity sources are distributed iid across households and survey waves with the joint distribution

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

As σ_g approaches one, idiosyncratic variance in utilities comes mostly from the nest level, not from distinctions between sources within a nest. Under the restriction $\sigma_g = 0$ there is no within-nest correlation and the model becomes a multinomial logit model.

Non-linear estimation of the first stage. In the first stage, we use maximum likelihood to estimate the parameters δ, γ and σ using equation 4. Let y_{itj} indicate that household i in survey t chose product j. The log-likelihood of the sample is

$$\log \mathcal{L}(\gamma, \sigma | y, z) = \sum_{i=1}^{N} \sum_{t=1}^{T} \log \Pr(y_{itj} | z_{it}; \gamma, \sigma, \delta(\gamma, \sigma)). \tag{5}$$

We write $\delta(\gamma, \sigma)$ to show that we concentrate the δ parameters out of the log-likelihood (Berry, Levinsohn and Pakes, 1995). For every candidate parameter vector (γ, σ) we solve for the δ that exactly fits the aggregate market shares.[25] This greatly reduces the dimensionality of the non-linear search, as the δ vector has up to 1200 elements (= 4 sources $\times$ 100 villages $\times$ 3 surveys), if every source were available in every village.

Linear estimation of the second stage. We can now use equation 3 to recover the $\overline{\beta}$ vector via a linear regression of the estimated $\widehat{\delta}_{jtv}$ on the observable characteristics x_{jtv} of electricity sources at the survey-by-village level. Let $\xi_{jtv} = \overline{\xi}_{jt} + \tilde{\xi}_{jtv}$ be the sum of a survey-wave average quality, $\overline{\xi}_{jt}$, for each source, and the deviation $\tilde{\xi}_{jtv}$ of the quality of a source in a village from that average. The main concern with estimation of equation 3 is that the error term $\tilde{\xi}_{jtv}$ measures the unobserved quality at the source by survey by wave level, inferred from market shares. If a source is very good in a particular village at a particular time, for example a diesel operator allows higher loads, then the price of that source may endogenously be set higher, implying $\mathbb{E}[\tilde{\xi}_{jtv} | x_{jtvk}] \neq 0$. A second concern is measurement error, since our price and characteristic measures are derived from surveys. Classical measurement error will attenuate the estimated price coefficient towards zero.

The traditional solution to these concerns is to instrument for price or other characteristics that may be endogenous to quality. In our setting, the solar microgrid experiment offers instruments that are exclude-able and likely to be powerful, given that the microgrid treatment changed demand (Table 4). Let T_{Normal} and $T_{Subsidized}$ be the same experimentally assigned village-level treatment

[25] 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.

indicator variables used in the estimation of demand for solar microgrids (from Table 4), and $\mathbf{1}\{Endline\}$ be an indicator for the endline survey wave. We estimate equation 3 by two stage least squares, where the two stages are:

$$\widehat{\delta}_{jtv} = \sum_{k \neq price} x'_{jtvk}\overline{\beta}_k + x_{jtv,price}\overline{\beta}_{price} + \overline{\xi}_{jt} + \tilde{\xi}_{jtv} \tag{6}$$

$$x_{jtv,price} = \pi_0 + \pi_1 T_{Normal}\mathbf{1}\{Endline\} + \pi_2 T_{Subsidized}\mathbf{1}\{Endline\} + \nu_{jtv}. \tag{7}$$

The $\overline{\xi}_{jt}$ are source by time fixed effects. The first stage, equation 7, uses the experimental treatment assignments, interacted with a dummy for the endline survey, when the experiment was ongoing, as instrumental variables. Thus, the instrumental variable is equal to zero in the baseline and follow-up periods, when the experimentally assigned prices did not apply.

As a basis of 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 should affect source mark-ups and prices under oligopolistic competition (Berry, Levinsohn and Pakes, 1995). Second, the average price for a given source in the nearest three villages where that source is available, which will covary with source price due to common supply shocks (Hausman, 1996; Nevo, 2001).

The hours of supply on the grid may also be endogenous to demand. To account for this possibility, in some specifications we also instrument the hours in a village (both on- and off-peak) by the supply predicted using the supply in villages nearby, using a similar logic as Nevo (2001). In our setting, basing a supply instrument on nearby supply is sensible, because the structure of the distribution grid physically links the supply decisions within a region; villages that are served by the same substation would be similarly affected by the distribution companies' application of the same rationing or "load shedding" rule. The exclusion restriction is that supply of electricity in nearby villages is not correlated with the determinants of demand in a given village. An example where this restriction would be violated is if there are common unobserved demand shocks across nearby villages, conditional on our rich set of household observables. Appendix A c details the construction of the instrument.

Having estimated equation 6, the fitted residuals allow us to recover unobserved quality as

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

With these estimates, we can observe how the quality of electricity sources varies across sources, villages and time.

Counterfactual surplus. With the parameters of the demand model we can calculate household choices and surplus under counterfactuals that vary the availability and characteristics of electricity sources. The aggregate market share of electricity source j is the choice probability for that source averaged over households:

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

The expected household-level indirect utility from a choice set $\mathcal{J}$ is the log of the sum over nests of a term dependent on nest inclusive value

$$\widehat{\mathbb{E}}\left[\max_j U_{ijtv} \,\middle|\, \mathcal{J}\right] = \ln \sum_g e^{(1-\widehat{\sigma}_g)\widehat{IV}_{igtv}}$$

We run counterfactuals by considering a restricted set of choices $\mathcal{J}'$ or by using the estimated coefficients to calculate new $\hat{\delta}_{jvt}$ associated with changed source characteristics. The willingness to pay for a scenario that alters the choice set or choice characteristics is:

$$\widehat{WTP} = -\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\mathbb{E}}\left[\max_j U_{ijtv} \,\middle|\, \mathcal{J}'\right] - \widehat{\mathbb{E}}\left[\max_j U_{ijtv} \,\middle|\, \mathcal{J}\right]\right) \Big/ \widehat{\overline{\beta}}_{price}$$

The main objects of interest in the counterfactuals are predicted market shares and household willingness to pay.

c Results

This section reports estimates of household demand for electricity sources. The full demand model has 1,031 parameters: 999 source-by-village-by-survey mean indirect utility parameters, backed out from the first-stage demand model, 28 parameters governing household heterogeneity, 3 parameters on the average effects of source characteristics and a parameter governing correlation of the source-specific utility shocks. We therefore report only select parameters, to give a sense of how the model represents household electricity choices. First, we report the linear estimates of the average effects of source characteristics, from the second stage. Second, we present estimates, from the non-linear first stage, of how household characteristics affect choice probabilities. Third, we present distributions of source quality.

Second stage estimates: Mean effect of source characteristics. We begin with estimates of equation 7, which is the first stage of the linear part (second stage) of the broader model. Appendix Table C4 gives the estimates with several different instrumental variables strategies. Our preferred specification is column 2, which uses our experiment and instruments for both price and supply hours. We find that the experimental treatment assignments have a modest, but highly statistically significant, effect on price. The mean price of *all* sources at endline in the control group is INR 95. In villages assigned to the normal price treatment, the mean price was INR 6.4 (standard error INR 2.8) higher per month; in villages assigned to the subsidized price treatment, this was INR 16 (standard error INR 2.1) lower per month. These effects in the first stage are averaged over all sources, though the experiment only directly changed the prices of microgrids. The first-stage F-statistic for a test of the null that the instruments do not affect price ranges from 21 to 42, depending on whether we instrument for price and hours simultaneously (column 2), or only for price (column 1). Appendix Table C4 additionally shows that our predicted supply instruments, based on supply to nearby villages, predict hours of supply both during peak hours and off-peak hours.

Alternative instrument sets lack power to predict price in the first stage. Neither the BLP (F-statistic 0.4) nor Hausman (F-statistic 1.0) instruments have much predicted 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, since sources like own solar are perfectly competitively supplied and the government's objective, in pricing grid electricity, is clearly not to maximize profits.

Table 5 reports estimates of the linear part of the demand model, equation 6. Column 1 reports results from ordinary least squares estimates, as a straw man, since we expect OLS will be biased. Columns 2 and 3 report instrumental variables estimates using the first stage from the experiment, instrumenting either for price or for both price and hours. Columns 4 and 5 replace the experimental variables in the instrument set with alternate instruments for price.

The OLS estimate of the effect of price on mean utility is negative and fairly precisely estimated, at -0.25 (standard error 0.12). This estimate would imply inelastic demand, as we discuss below. The experimental instrumental variables estimates of the price coefficient are larger by a factor of seven. We find a coefficient of -1.70 (standard error 0.63) on price (column 2), which is unchanged if we additionally instrument for hours of supply (column 3). Price has a much greater effect on mean utility than the OLS estimates would imply, consistent with bias from some combination of endogeneity and measurement error. Estimates of the price coefficient using alternative sets of instruments are imprecise. One point estimate is half as large as the experimental estimate, another is twice as large, and we cannot reject their equality with any of the experimental estimates, the OLS estimates, or a zero coefficient on price. We therefore conclude that the experiment is necessary to recover unbiased and precise estimates of the price coefficient in our setting.

We calculate the price elasticities implied by these coefficients using our preferred, column 3 estimates. (The elasticities also depend on the other parameters of the demand model, including household tastes and quality, that we discuss below.) Table 6 presents these aggregate own- and cross-price elasticities by source. The demand elasticity for grid electricity is estimated to be -0.58. We view households as very price sensitive in absolute terms. The average probability of choosing the grid is 24%, and the model estimates imply that a INR 10 increase in the grid price (17% of the mean price of INR 60) decreases grid market share by 2.9 pp (12% of the average share). Though INR 10 is just enough money to buy two cups of tea or three bananas, raising the grid price by this amount in a month cuts market share by a noticeable 3 pp. Demand is even more elastic for off-grid electricity sources, which have smaller market shares. Diesel, own solar and microgrid solar electricity have own-price elasticities of -1.83, -1.91 and -1.58, respectively.

We also estimate the effect of supply hours on household mean utility. We find a positive but statistically insignificant effect of peak hours of supply on mean utility and a smaller, negative, and borderline statistically significant coefficient for off-peak hours. Our estimate for the value of peak hours is not precise, but agrees with the idea that agricultural households, who may be away during the day, mainly value power in the evening hours. We proceed with the column 3 estimates, instrumenting for both price and hours, as our main specification for counterfactual analysis. The IV specification is preferred on the *a priori* grounds that supply may be rationed in part based on village demand. In practice, the IV estimate of the value of peak hours (column 3) is similar to the coefficients in both the IV specification where only price is instrumented for (column 2) and the OLS specification (column 1).

First stage estimates: Heterogeneity in demand across households. Table 7 reports the effects of household characteristics on choice probabilities in the demand model. Choice probabilities depend on both the mean level of utility from each source and household observables, via equation 1. We estimate two instances of the model. First, to provide a simple univariate proxy for wealth, we estimate a model that includes as covariates only the number of adults in the household and a dummy variable for whether the household has a solid roof (columns 1 through 5). Second, we estimate the full model, which includes five additional observable proxies for household demand: whether the household has a solid house, the number of rooms in the house, household income, whether the household owns agricultural land, and the education level of the household head. The effects of household characteristics are non-linear. The table therefore reports marginal effects evaluated for a "poor" household, which lacks the binary indicators of wealth and has an income at the 20th percentile of our sample distribution.[26] The marginal effects are not strictly marginal; for binary variables we report the effect on each given choice probability of changing the value from zero to one, and for continuous variables the effect of an one standard deviation increase.

The main finding of the table is that richer households, by any measure, have stronger preferences for grid electricity over all other sources. Consider the simple model specification (columns 1 to 5). The baseline probability of grid choice is 24 percent. On top of this base, a household with a solid roof is 21 pp (standard error 3.9 pp) more likely to choose grid electricity. Nearly all of this effect comes from a reduction in the choice of the outside option (no electricity), rather than substitution from other sources.

In the full model we add additional covariates (columns 6 to 10). The effect of having a solid roof on grid choice declines, since it is correlated with other measures of wealth, but remains large (11 pp). We find that each of our seven observable demand proxies has a positive, economically meaningful and statistically significant effect on a household's probability of choosing grid electricity, and also reduces the probability that a household chooses no electricity (the outside option). For example, a

[26] The profile of a poor household is defined as a household of two adults living in an one-room house, without a solid roof or solid walls, and no agricultural land ownership. The full profile of a poor household's characteristics is described in Appendix Table C6.

household that owns agricultural land is 4.9 pp (standard error 1.8 pp) more likely to choose the grid. These demand proxies have much smaller effects on the choice probabilities for other sources, though some do significantly affect demand; for example, higher-income households are slightly more likely to choose microgrids. Table 2 offers a natural interpretation of this heterogeneity: grid electricity offers higher load, and many more households on the grid can run a fan or a television.[27] Richer households want the energy services that these devices bring. A likelihood ratio test, reported at the bottom of Table 7, easily rejects the simple demand model in favor of the full model (p-value $<$ 0.001). We use the full model for counterfactuals.

Unobserved source quality. The demand model flexibly allows for changes in ξ_{jvt}, the unobserved quality of each electricity source in each village and survey wave. In the model, unobserved quality is the residual from the estimation of 6. This residual fits the source market shares exactly, conditional on source-specific observable characteristics, and is naturally recovered only for source-village-wave combinations in which the source was available in the market. Quality for a source will be lower if there are unobserved costs of using that source (e.g., connection fees or hassle) and higher if there are unobserved benefits (e.g., reliable service).

Figure 4 summarizes the evolution of source quality. Each row shows one source and each column one survey wave. Within each source and wave, the histogram shows the distribution of quality across villages. We also plot the median source quality as a horizontal line in each histogram.

The main finding from the figure is we observe large changes in quality for sources that we have *a priori* grounds to expect quality improvements, but not otherwise. The two disruptions in the market came from grid electricity and own solar systems. The estimated quality of the grid improved greatly over the span of our data. The median grid quality increases from -1.1 at baseline to -0.3 at endline and $+0.7$ at follow-up (Figure 4, row 1). These improvements are likely the result of a government drive to increase household connections, which decreased connection costs by subsidizing poor households. The estimated quality of own solar also improved, especially between the endline and follow-up waves (from -2.3 to -1.2) (Figure 4, row 4). These increases could be due to improvements in technical factors such as battery capacity and load, which we do not observe directly, or to a broader reach of marketing and distribution networks for these systems,

[27] In our baseline survey, very few households use microgrids. As a result, the appliance ownership summary statistics for microgrids at baseline are subject to large standard errors.

which would have lowered connection costs.

Quality for diesel generators and solar microgrids, by contrast, stagnated. The distribution of diesel generator quality, for example, is about the same in all three survey waves (though there is truncation at the bottom, due to exit, as generators were driven out of the market). Our microgrid partner, HPS, did not offer its product in many villages at baseline, by design, and did not change its product during our study. This stagnation is apparent in the figure, as the median quality of microgrids is unchanged across survey waves (Figure 4, row 3). Our estimates of quality therefore do not support explanations for the fall in microgrid market share that stem from changes in the product itself, such as households learning they did not like microgrids or microgrid maintenance being poor. Instead, the estimates imply that changes in the price of microgrids and in the quality and availability of other products rationalize falling microgrid market shares, for the same product that was offered all along.

Overall, we find a remarkable concordance between our prior understanding of changes in quality for each technology and the qualities inferred from the demand model (Figure 4). The figure reveals how the landscape of electrification in Bihar has shifted, with the grid and own solar systems rapidly improving and other technologies stagnating.

d Modeling choices

The model casts household electricity demand as a static differentiated choice problem. Here we discuss several of our modeling choices.

Nested logit. We use a nested logit model instead of a random coefficients (mixed) logit model for three reasons. First, we have especially rich observable household data that allows for complex patterns of substitution, even without random coefficients.[28] Second, we find that introducing a small amount of unobservable correlation in tastes, via the nested logit assumption, has negligible effects on the estimates.[29] Third, the nested logit model can be estimated efficiently by maximum

[28] We have household-level panel data with very detailed observable household characteristics, which we show have large effects on demand, and a small number of product choices. Therefore, the aggregate patterns of substitution in the model will not be tied to simple patterns like the independence of irrelevant alternatives, even within nests, because individual households make their own decisions. A mixed logit model would allow patterns of substitution to be richer still (Berry, Levinsohn and Pakes, 2004). We believe the gains from a more complex model would be larger in settings where one had fewer observable proxies for demand and a larger number of product choices.

[29] Appendix Table C12 shows that the coefficients on observable characteristics and the fit of the model barely change at all when varying the nest structure, or using a multinomial logit model with no nest at all. Nested logit is a simple case of a mixed logit model where the random coefficients are on group-specific dummy variables (Cardell,

likelihood without simulation.

Substitutes. The model's structure assumes that sources are substitutes and that households cannot choose bundles of sources. In some settings, for example in cities, households may have diesel generators or solar power to provide power during grid outages, making the technologies complementary. We see very little bundling in our sample, perhaps because households are too poor. At the time of our endline survey, which is used as the point of departure for counterfactuals, only 1.4% of households held multiple sources (Appendix Table A1). For these few cases, we set a priority order in which households are assumed to have chosen the grid if it is part of their chosen bundle.

Static model. We use a static model instead of a dynamic model, where households hold sources as assets, or condition future choices on past decisions. We took this route for two main reasons. First, in our context, three of the four sources we study are paid 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. We see total dis-adoption of diesel, and adoption and then dis-adoption of microgrids, within our study period, and massive changes in shares from one year to the next. These fluid aggregate movements suggest that households do not show a stickiness in their connection to a particular source. Our model does allow for unobserved adoption or connection costs, via the quality terms.

5 Counterfactuals on the Value of Electrification

The transformation of Bihar's electricity landscape happened on several dimensions at once. The demand model we have estimated now allows us to break down the household surplus from the different changes that Bihar went through, and from counterfactual policy changes. With administrative data from the public utility, we can also study the impacts of counterfactual policies on the losses of the utility.

We specify counterfactual scenarios to address four broad questions. The first three concern the supply side. First, how much have innovations in solar power benefited households? Second, what has been the value of government investment to expand grid access? Third, how would further

1997).

changes in electricity supply, with respect to investment, pricing and quality, affect household access and surplus? This question is acutely policy relevant, because the *status quo* electricity policy in Bihar, as in many developing economies, is somewhat of a puzzle: governments invest large amounts of capital to extend the grid to rural areas, but then provide low quality service and ration supply to the poor customers they sought to reach (Burgess et al., 2020). The fourth question concerns demand. We ask how economic growth may affect electricity access and surplus, either with no further supply changes, or along with continued and simultaneous investment in the grid.

Table 8 reports the counterfactual results. The table treats the endline survey (mid-2016) as the *status quo* level against which alternatives are compared, so, unless otherwise noted, we hold the characteristics of households and the availability and characteristics of sources at endline survey levels. For example, the counterfactuals assume that the village-level availability of sources are set at their endline values of 53% for grid, 18% for diesel, 100% for own solar, and 66% for solar microgrids, so that when a source is counterfactually removed from the market, it may not have been present in all villages to begin with. Columns 1 through 5 report source market shares, including the outside option of no electricity. Columns 6 to 8 report consumer, producer and total surplus. All surplus measures are per household per year across the entire population, including households who choose the outside option of no electricity. The estimate of producer surplus is the surplus from operating the grid only. 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.

Table 8, panel A summarizes the endline data in row 1 and reports on the fit of the model in row 2. Unobserved source quality terms allow the model to exactly fit the actual market shares. In the data, 57% of households have no electricity access. The grid is the most popular source of electricity, with a 24% share, though the two types of solar are not far behind, with a combined 17% share. The average household surplus from all electricity sources (the *status quo* choice set) is INR 528 per year. For context, household income is on average INR 92,261 per year and energy expenditures INR 5,520 per year. Producer surplus is INR −497 per household, hence the total surplus from electricity supply is roughly zero (a paltry INR 31).

It is apparent that the state places a high priority on access, despite the significant losses it incurs from serving poor households. A natural interpretation is that the state places a higher weight on the surplus of the rural poor than on utility losses, so that subsidized electricity may lead to higher welfare, if not higher social surplus, than not supplying electricity at all. Or, the state may believe that there are externalities from energy supply that our demand estimates do not capture. Because of these considerations, we find it difficult to interpret total surplus, and will focus on how counterfactuals change the separate components of consumer and producer surplus.

a Innovation in solar power

The first set of counterfactuals considers the value of innovation in solar power. Figure 5 shows how the market shares of all sources respond to changes in the price of solar systems. Again, the characteristics and availability of sources are held constant at endline survey levels. The price of solar microgrids, on the horizontal axis, ranges from INR 50 per month, up through the range of our experimental treatments, from INR 100 to INR 200 per month, to a maximum price of INR 300 per month. Based on Department of Energy and National Renewable Energy Laboratory projections, we forecast a 2022 microgrid price of INR 120, slightly more expensive than our subsidy treatment of INR 100 (Feldman, Margolis and Denholm, 2016; Howell et al., 2016).[30] While the horizontal axis is in terms of the price of microgrids, in the construction of this figure, we vary own solar prices proportionally with microgrid prices, since solar panels form the main capital costs for both of these sources.[31]

There are three main results from the figure. First, echoing the reduced-form demand estimates for microgrids (Table 4), we see that microgrid demand is only 6 pp at INR 200, and negligible by INR 300. In other words, at prices prevailing before our study period, off-grid solar power could not

[30]For solar PV, we assume a 55% reduction in cost (Feldman, Margolis and Denholm, 2016). For batteries, we assume a 75% reduction in cost, in accord with the US Department of Energy's 2022 goal (Howell et al., 2016). Since the panel and batteries only make up a part of the system, these changes imply a reduction in total cost of 30%, or INR 50 (USD 0.83) (See Appendix Figure D3 for a breakdown of costs). We assume that this decline in cost passes through completely to microgrid prices, thereby lowering prices from a market price of INR 170 to INR 120.

[31]We apply the same proportional reduction in cost to own solar, which assumes that the 55% of total cost due to panel and batteries observed for microgrid applies to own solar as well. This is conservative because the HPS product involves monthly recharge costs that do not apply to the own solar product, so the cost of own solar, being made up mostly of capital, may in fact fall more than we forecast.

gain a meaningful market share among this population. Second, as prices rise, households substitute away from solar power, mostly to no electricity at all (solid line), but also partly to grid electricity (dashed line). At higher prices of solar, above INR 200, most households who have grid electricity in their village (53%), and can switch, have switched to the grid, so the grid market share is flat above this level. As solar prices rise even further, households switch to the outside option of no electricity. Third, cutting prices to the projected 2022 level implies that the share of households with any source of electricity (one minus the share with no electricity) would increase by 13 pp, relative to electrification at a price of INR 300. Hence, solar serves as a technological stop-gap between grid electricity and kerosene, the outside option.

Table 8, panel B measures how the surplus from the advent of solar power depends on the availability of substitutes. We model the advent of solar power as a reduction in solar prices from infinity to the observed 2016 endline prices. If solar were removed altogether (panel B, row 1), four pp (or 24%) of the households displaced from solar would connect to the grid, and the share of households without electricity would increase by 11 pp, to 68%. The total household surplus from electrification would fall by just 29% (from INR 528 to INR 377). Panel B, row 2 expands the supply of solar microgrids to the 34 villages that did not have them at the endline. Given that microgrids are relatively costly, this leads to an almost imperceptible rise in total consumer surplus. Row 3 captures the effect of innovation in solar bringing prices down to projected 2022 levels; household surplus would rise by 13%, to INR 594 per household, and the market share of off-grid solar sources would jump by 9 pp to 26 pp. Across panel B, large changes in solar availability and costs have fairly large effects on solar market shares but muted effects on the overall rate of electrification and total household surplus from electricity.

Although Section 3's finding that WTP for *solar microgrids* is low foreshadowed this section's finding that off-grid solar has a modest impact on household welfare, the demand model is a major advance in understanding households' preferences for *electrification*. In particular, the model estimates show that studying demand for microgrids alone, without considering other sources of power, can lead to a gross misunderstanding of household willingness to pay for electricity. In the full structural model, the loss in surplus from removing microgrids from the market is INR 93 per household per annum, nearly the same as what we found with the reduced-form demand curve (Section 3). In the model, however, we can additionally estimate that the total household surplus from electrification is INR 528, greater than the surplus from microgrids by roughly a factor of five. Demand for microgrids is low and elastic because households can easily move to other sources of electricity. Moreover, our model estimates imply that this finding of low demand for a specific source is not restricted to microgrids; similar gaps between the value of any one source and the value of electricity apply to all sources, including the grid (see Table 8, panel C, row 2). The general point is that the demand curve for a single product is a poor measure of household surplus in a supply environment where several similar products compete.

The competition between sources also implies that the advent of solar power reduces the state utility's losses due to grid power supply. In the status quo, producer surplus is INR −497 per household per year. If solar power disappeared (row 1), producer surplus would fall to INR −581; if the cost of off-grid solar power declined to projected 2022 levels (row 3), producer surplus would rise to INR −433. The advent of solar power, from being out of the market to 2022 prices, therefore reduces the state's own financial losses from grid power supply by 25%. This finding provides a novel justification, aside from environmental externalities, for why developing country governments that subsidize the grid may wish to subsidize household solar adoption also.

b Improving grid access and quality

In our setting, grid service is less than complete: the grid is present in only 53% of villages at endline and, where it is present, the average supply is 11 hours per day. Table 8, panel C reports counterfactual market shares and surplus based on changes in grid availability and supply quality.

If the grid were to be removed from all villages, the share of households without electricity from any source would rise modestly, from 57% to 66%, because solar would gain 13 pp of market share (panel C, row 2). Household consumer surplus from all sources would fall by only 32%, because solar serves as a backstop or substitute for the grid, at these households' current income and wealth levels.[32] If the grid were to be extended to all villages (panel C, row 3), the electrification rate would rise by 10 pp, to 47%. In this scenario household surplus would increase by 22% relative to

[32] For all counterfactuals, we do not model supply responses from other sources, even though in practice diesel generators may enter and exit. Our assumption of no supply response is probably accurate for counterfactuals that increase the quality or extent of the grid, due to diesel's already low market share of 3% at the endline. However, for counterfactuals that make the grid worse, our counterfactual results probably overstate the losses in consumer surplus and electrification rates, since we expect diesel providers would enter if the utility did not. At baseline, diesel generators held 17% market share.

the status quo (panel A, row 2), or 78% relative to grid nowhere (panel C, row 1). Producer surplus becomes more negative, with the utility losing 62% more per household as it increases the number of villages served (panel C, row 3 compared to panel A, row 2). Grid expansion thus acts as an ongoing commitment by the state government to future losses on energy supply.

The model can also consider changes on the intensive margin of grid quality, such as changes in peak supply hours. We counterfactually increase peak supply hours by two hours a day, up to a maximum of five hours, which is the full duration of the evening peak. The better supply draws 5 pp more households onto the grid, nearly all from no electricity, holding constant the number of electrified villages (panel C, row 4). The status quo poor quality of grid supply, therefore, has a meaningful, but modest, effect on whether households choose to get a grid connection. Producer surplus, at the same time, falls from INR −497 (panel A, row 2) to INR −775 (panel C, row 4). The utility's losses from serving each customer increase by 56%, nearly as large as the cost for extending grid services to all villages, and far larger than the corresponding gain in household surplus.

Two themes emerge from these counterfactuals. First, even after either universal village-level electrification, or large changes in the quality of supply, nearly half of all households would remain unelectrified. This electrification gap is the justification for the government's program to subsidize infill connections, which we suspect is reflected in the higher unobserved grid quality by the follow-up survey (Figure 4, row 1, column 3). Low demand is a major reason that progress in connecting poor, rural households in developing countries has been slow. Second, both extensive and intensive margin reforms to encourage access would increase utility losses markedly. At the extensive margin, we estimate that the monetary loss to the utility from grid extension is larger, by a factor of 2.6, than the additional willingness to pay by households in villages presently without grid electricity.

c Grid pricing reforms

Electricity in Bihar is subsidized and the state tolerates informality, which further lowers effective prices. Here we use our demand model to consider alternative approaches to supply. The goal is to understand why the government has chosen this particular energy service bundle of low quality and low price, which entails large producer losses.

Table 8, panel D reports counterfactual results for grid reforms. Row 1 removes theft, which we model by increasing the grid price from the observed mean monthly payment of INR 60 to INR 128 per month (see Section 2 b). This price increase would devastate the grid's market share, which falls from 24% to 8%; this decline is a direct consequence of the experimental estimates showing elastic demand in this range of prices. Although the number of households with grid electricity would decrease by 16 pp, solar soaks up half of these households, so that, on net, the electrification rate falls by only 6 pp. Household surplus falls by a modest 24% (INR 128) as solar's availability provides a substitute, while producer losses fall plummet 79% (INR 392). Households are sufficiently price sensitive that the equilibrium outcome with a grid priced roughly at cost resembles the outcome under a scenario where the grid does not exist at all (panel C, row 2).

Panel D, row 2 reports on a more targeted reform that restricts the higher counterfactual price of INR 128 per month to households with incomes above the poverty line and leaves the price at INR 60 for below poverty line households. Such a reform would preserve the goals of improving the distribution company's finances and subsidizing poorer households, but may be infeasible, since the utility serves above and below poverty line households on the same distribution wires, which makes preventing theft difficult. If it could be done, such a reform would reduce the grid's market share by 5 pp, but most of the targeted APL households would switch to other sources, such that the overall electrification rate falls by just 2 pp. There is clear value in targeting: household surplus falls by a mere 7% while grid losses are cut 28%.

A less drastic reform would try to move from the present low price, low quality bundle to a higher price, higher quality bundle; in other words, compensate households for additional payments with better energy services. Many policy observers have recommended reforms of this kind (Zhang, 2018). We use the model to consider a budget neutral "grand bargain": the government increases supply during peak hours and pays for that increase by raising prices. We calculate that a price increase to INR 103 per month would be sufficient to pay for increasing peak hours of supply to the maximum of five.

Table 8, panel D, row 4 reports that this ambitious grand bargain has minimal benefit: it leaves household surplus slightly lower and producer surplus, by construction, the same as in the status quo. The grid electrification rate ticks down by 3 pp, but due to rounding, overall electrification remains essentially unchanged. The interpretation that the grand bargain roughly breaks even depends on the model specification we use to a greater extent than do the other counterfactual findings.[33] The upshot of this scenario, across all specifications, is that the government has no route to increase quality and access without also markedly increasing its own losses.

We interpret the collection of counterfactual results in panels C and D as indicative of the state's substantial willingness to pay, via large distribution company losses, for greater grid energy access, albeit bundled with relatively poor quality. There are two sets of findings that support this view. First, grid expansion (panel C, row 3) and supply increases (panel C, row 4) are political priorities, even though their net effect is a reduction in total surplus: while they do increase consumer surplus, they increase utility losses by around thrice the increase in consumer surplus. In our sample period, the state rapidly expanded grid access, from 29% to 72% at the village level, and increased hours of supply as well. Second, the state could substantially reduce its losses through some combination of charging higher prices and reducing theft (panel D, rows 1 and 2). Either these reforms are infeasible to implement, or the state judges the gain in its finances to be worth less than the loss in household surplus. We conclude that the current policy bundle can be rationalized by the state placing a higher weight on the surplus of the rural poor than on its treasury, or, alternatively, by the state perceiving benefits from rural electrification that are not captured by household willingness to pay.

d Higher incomes and predicting the future

The final set of counterfactuals consider changes in demand due to increases in income. Within our sample, we observe significant heterogeneity in household demand based on observable characteristics; recall, we found that an indicator for solid roof ownership increases the probability that a household chooses grid electricity by 21 pp, on a base of 24 pp. Here we use our full demand model to better capture the effects of income growth and wealth accumulation on demand for electricity. In Bihar, per capita income grew at an annualized rate of 10.2% from 2011 through 2017, implying a doubling of income in roughly 7 years, so any realistic projection of the future electricity market needs to account for demand growth.

[33]We estimate the value of price to consumers using our experiment, but do not have as strong an instrument for hours of supply. Because our estimate of the value of peak supply is imprecise, the conclusion depends on the exact specification of the demand system. For example, if we were to instrument for price but not for hours, in the second, linear stage of the demand model, then the grand bargain would decrease total surplus more, due to lower household willingness-to-pay for additional supply in that specification. Hence, the finding that the grand bargain breaks even, for consumer surplus, is perhaps slightly optimistic.

To model demand growth, we create profiles that represent "poor," "median" and "rich" households within our sample. These profiles are vectors of household observable characteristics where each element roughly corresponds to the 20th, 50th and 80th percentiles for each wealth and income proxy (see Appendix Table C6 for a complete description of the profiles). In the counterfactual scenarios, we raise the levels for all households to the minimum of their current observables and the median (or rich) profile, and study how these shifts change demand.

Table 8, panel E reports on the results of these counterfactuals. Raising all households at least to the level of the median household increases electrification rates from 43% to 49% (row 1), and raising them to the level of the rich household to 60% (row 2). Of the total 18 pp gain in electrification, 11 pp (61%) comes from the grid, somewhat larger than the grid's status quo inside market share (56%), despite that the grid is only present in about half of all villages. Household surplus from electrification increases by 60% when all households have at least the rich (80th percentile) profile.

In a final scenario, (panel E, row 3), we take a shot at predicting the medium-run future. We make the following simultaneous changes: all households achieve the "rich" profile; the grid expands to all villages; peak hours of supply increase to the feasible maximum of five peak hours; and solar costs decline to projected 2022 levels. While these changes may seem profound, they are consistent with recent growth and grid investment in Bihar. In this scenario, the electrification rate surges to 81% and household surplus from electrification shoots up by 148%, relative to the status quo. The electrification rate rises 38 pp, which consists of a 42 pp increase from the grid, 110% of the total, and a 4 pp *decrease* from other sources of electricity. When the grid is present, therefore, income growth leads to gains in electrification *solely* from new grid connections, even though we simultaneously project further declines in the cost of off-grid solar systems. This finding is a consequence of the heterogeneity across households in our demand model, which implies that increases in demand due to income growth have a larger marginal effect on grid market shares than on the shares for other sources. In short, off-grid solar is a stop-gap technology, which fills the product space between no electricity at all and grid electricity, but the future will run on the grid.

6 Conclusion

Electricity markets on the global electrification frontier are undergoing radical changes, driven by a combination of a traditional "big push" for grid electricity and technological innovation that has brought off-grid solar generation within the economic reach of the poor. Governments, multilateral agencies, and private companies are setting the shape of the frontier and the pace of expansion, with investments in and subsidies for both grid and off-grid technologies. Yet there is little evidence on which of these technologies the people on the global frontier would choose for themselves.

To study how households value electricity, we collect data on household electricity demand for *all* sources and model household choices with a discrete choice demand system. We use experimental variation in price to estimate the model and then draw on rich data, on both sides of the market, to apply the model to a broad range of questions about the value of electrification. In particular, we study how substitution between grid and off-grid electricity sources affects household surplus from a range of investments, technological changes and policy reforms.

There are three main findings. First, households value electricity, but demand for any individual electricity source is very elastic, because several sources provide comparable services at similar prices. Household surplus from electricity, for example, is 5 times greater than the surplus from just solar microgrids. Second, even in a relatively poor, rural population, higher-income households greatly prefer grid electricity, presumably because it can support higher load appliances like fans and televisions. Households do adopt solar, when their choices are limited, and so solar may continue to gain market share in areas where the grid has not reached, such as in many rural parts of sub-Saharan Africa, or where energy on the grid is priced at a prohibitive level for the poor. Third, when the grid is ubiquitous, income growth leads to gains in electrification that come mainly from new grid connections. Off-grid solar power is a stop-gap technology; the future will run on the grid.

Our findings help to rationalize a system of simultaneously subsidizing and rationing grid electricity, which we observe in Bihar and which is common in developing countries (Burgess et al., 2020). The state is willing to bear large losses for greater energy access provided at low quality. Our demand estimates show that households prefer this bundle to feasible alternatives. Current policy, to redistribute surplus via grid electricity, can be rationalized by the state placing a higher weight on the surplus of the rural poor than on the distribution company's losses, or by external returns to grid electrification (Lipscomb, Mobarak and Barham, 2013). Reforms that raise the price for households to ensure higher quality appear, by our estimates, to be utterly incompatible with wide access to energy. A meaningful direction for reform may be targeting, for example, by bringing all subsidies explicitly onto the books and redistributing only through lifeline tariffs, which would replace indiscriminate rationing with a limited amount of cheap power available to each household.

The rationed and unreliable supply we observe is far from the guaranteed 24-hour supply that developed countries enjoy, and which is arguably necessary for many kinds of productive business (Allcott, Collard-Wexler and O'Connell, 2016). Future research may better break down the costs of this low quality equilibrium, in terms of foregone economic growth, and find ways for governments to move out of it without decimating energy access on the global electrification frontier.

References

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

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

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

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

Attanasio, Orazio P, Costas Meghir, and Ana Santiago. 2011. "Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresa." *The Review of Economic Studies*, 79(1): 37–66.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2017. "Labor markets and poverty in village economies." *The Quarterly Journal of Economics*, 132(2): 811–870.

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

Berry, James, Gregory Fischer, and Raymond Guiteras. forthcoming. "Eliciting and utilizing willingness-to-pay: evidence from field trials in northern Ghana." *Journal of Political Economy.*

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

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

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

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

Bihar State Government. 2015. "Bihar Report Card." Bihar State Government.

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

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

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

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

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

Cardell, N Scott. 1997. "Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity." *Econometric Theory*, 13(2): 185–213.

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

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

Duflo, Esther, Rema Hanna, and Stephen P 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.

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

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

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

Goldberg, Pinelopi Koujianou. 1995. “Product differentiation and oligopoly in international markets: The case of the US automobile industry.” *Econometrica: Journal of the Econometric Society*, 891–951.

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. 2019. “Demand for Off-grid Solar Electricity: Experimental Evidence from Rwanda.” *United States Associate for Energy Economics Working Paper*, 19(411).

Hanna, Rema, Esther Duflo, and Michael Greenstone. 2016. “Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves.” *American Economic Journal: Economic Policy*, 8(1): 80–114.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

McFadden, Daniel. 1978. “Modeling the choice of residential location.” *Transportation Research Record*, , (673): 72–77.

McFadden, Daniel. 1980. “Econometric models for probabilistic choice among products.” *Journal of Business*, S13–S29.

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

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

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

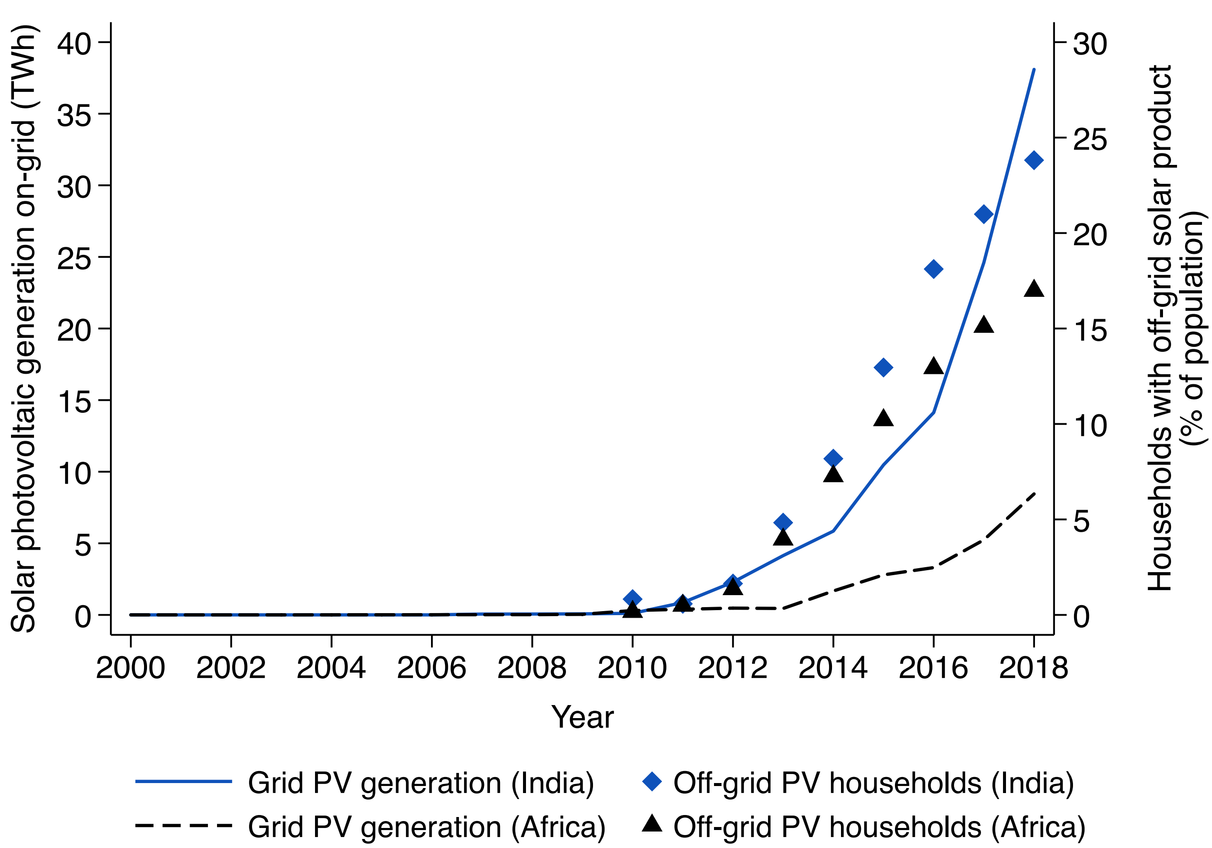
United Nations Department of Economic and Social Affairs. 2014. “Electricity and education: The benefits, barriers, and recommendations for achieving the electrification of primary and secondary schools.” United Nations Department of Economic and Social Affairs.

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

Zhang, Fan. 2018. *In the Dark: How Much Do Power Sector Distortions Cost South Asia?* The World Bank.

7 Figures

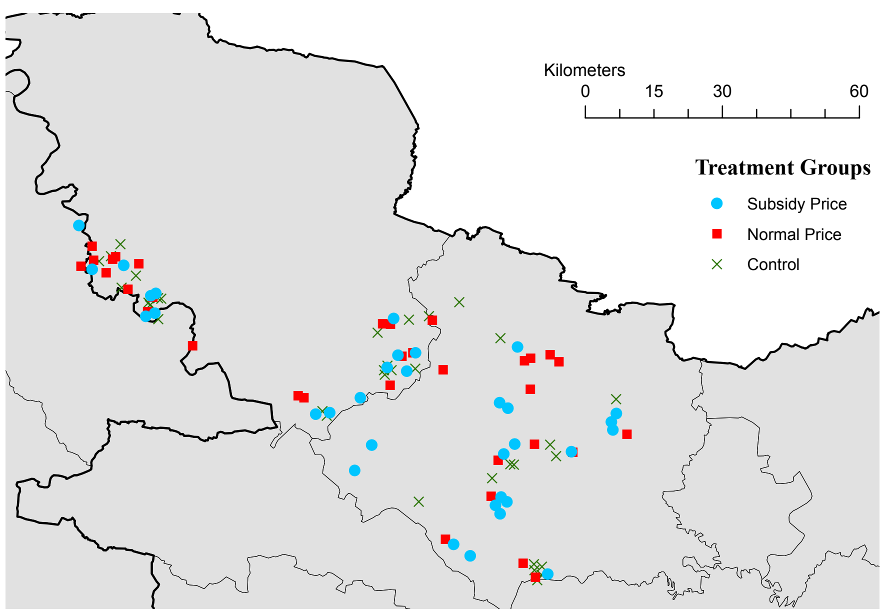
Figure 1: Growth of Solar Power in Developing Countries



The figure shows the growth of solar power in India and Africa, which account for most of the global population without electricity. The line series, measured against the left axis, show grid electricity generation from solar photovoltaics for India and the African continent. Generation data comes from the International Energy Agency, IRENA and the Central Electricity Authority, Government of India. The marker series, measured against the right axis, denote the percentage of households using off-grid solar systems in India and Africa. We estimate cumulative household market shares using data on solar system sales from GOGLA. To calculate the stock of market shares from flow data on sales, we assume that each household owns only one system and the number of systems in use is the sum of systems sold less a 10% annual depreciation of the prior stock. We divide by the population of households using population and household size data from the UN Population Division, the World Bank and the Indian Census.

Figure 2: Maps of Study Area

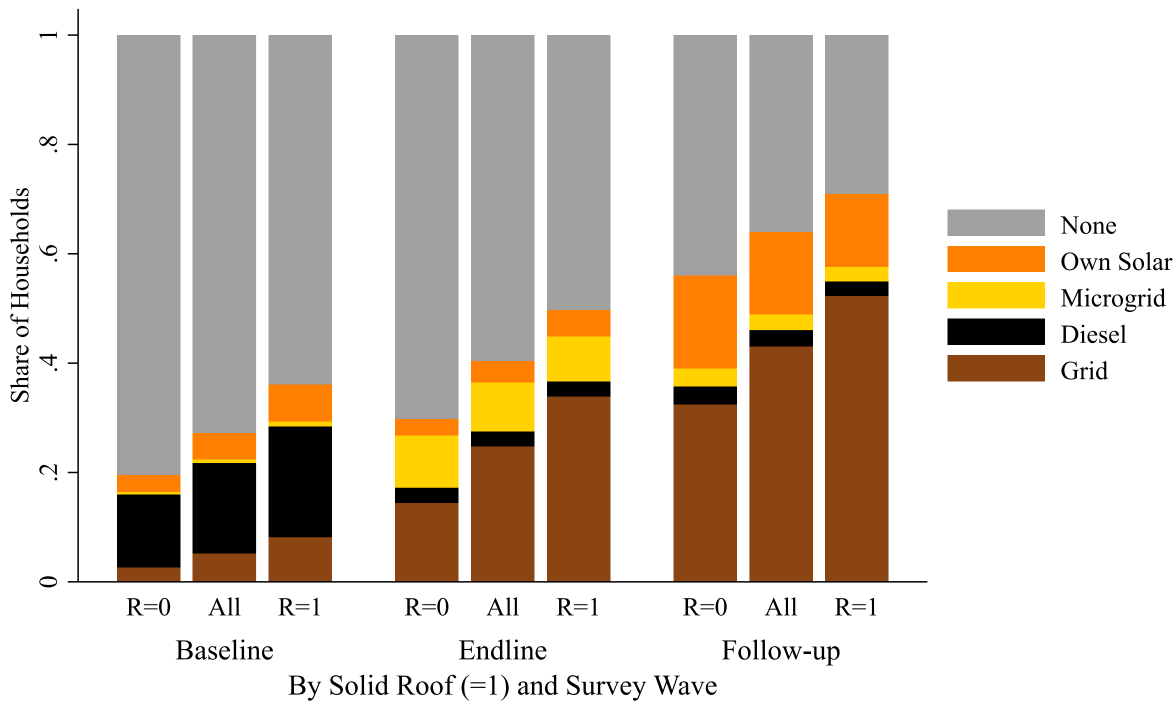
A Study districts within the state of Bihar, India



B Sample villages within study districts

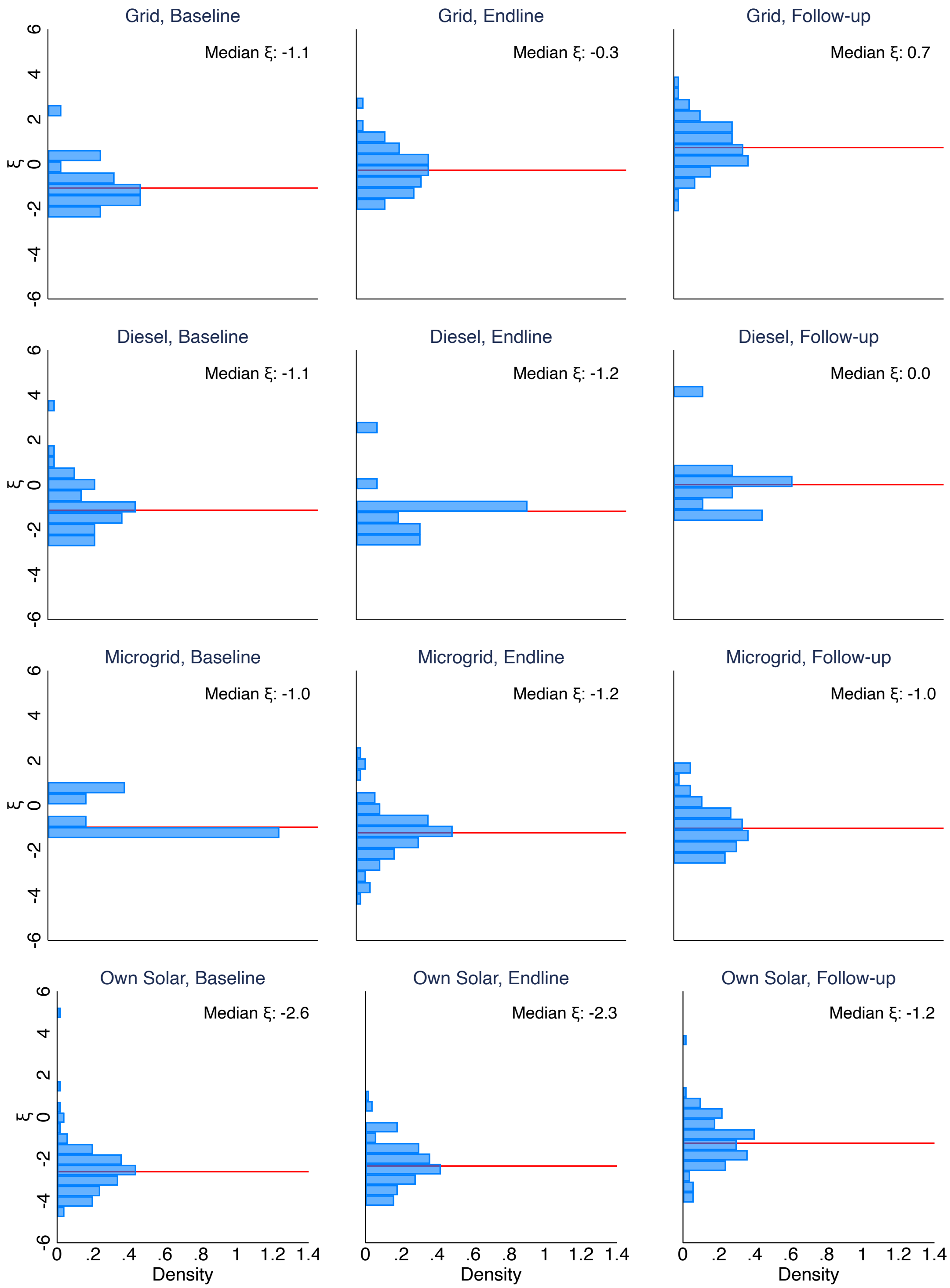
The figure shows the study area. Panel A highlights the two districts of West Champaran and East Champaran, in the northwest corner of Bihar, which contain the study villages. Panel B shows, within the two study districts, the locations of sample villages and their treatment assignments. The nearest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

Figure 3: Household Electricity Sources Over Time



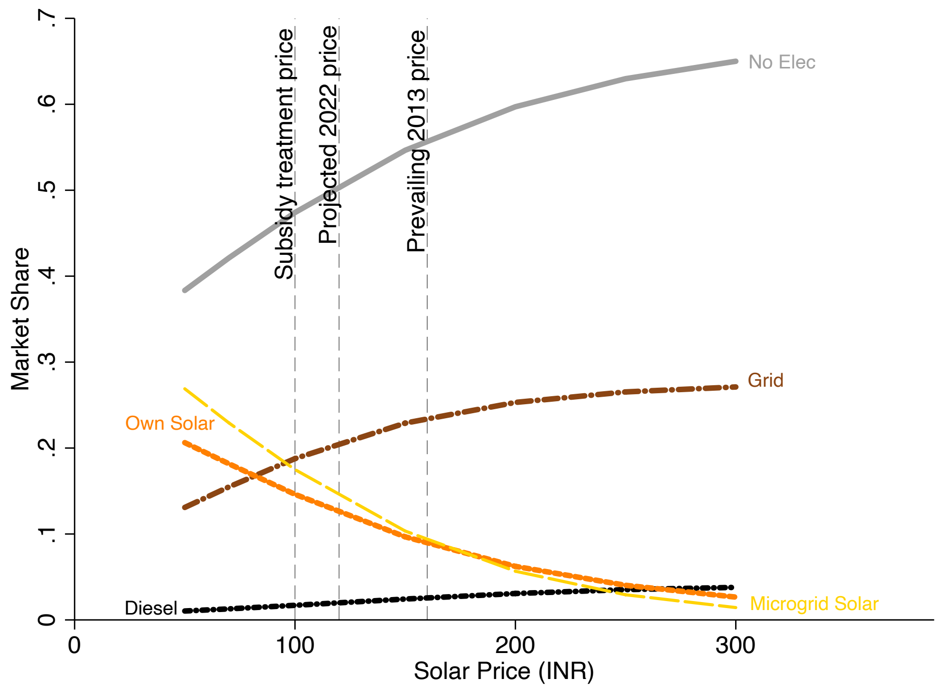
The figure shows the market shares of different sources of electricity over time. Each stacked bar gives the share of households, from bottom to top, that use grid electricity, diesel generators, solar microgrids, own solar systems or no electricity. These market shares are calculated with respect to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

Figure 4: 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.

Figure 5: Market Shares under Varying Solar Prices



The figure shows the market shares for all electricity sources as we counterfactually vary the price of solar power in the model. Each curve is the predicted market share of an electricity source. The horizontal axis gives the price of a solar microgrid. While the horizontal axis shows the price of a solar microgrid only, we vary the price of own solar systems proportionately with the microgrid, on the grounds that capital cost reductions in solar photovoltaic panels or batteries would affect the prices of both kinds of solar system. Household and source characteristics and the availability of all sources are fixed at their endline (mid-2016) levels. The three vertical lines indicate, from left to right: the price under our subsidy treatment arm in the experiment; the projected 2022 price (Feldman, Margolis and Denholm, 2016; Howell et al., 2016); and the prevailing 2013-2016 price.

8 Tables

Table 1: Electrification Around the World

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

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

Table 2: Summary of Electricity Sources

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

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

Table 3: 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. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Solar Microgrid Demand by Village Treatment Arm

	Survey		
	Baseline	Endline	Follow-up
	(1)	(2)	(3)
Treatment: Subsidized price	-0.001	0.193***	0.081***
	(0.005)	(0.049)	(0.027)
Treatment: Normal price	0.009	0.060**	0.020*
	(0.010)	(0.028)	(0.012)
Constant	0.006	0.023***	0.002
	(0.004)	(0.005)	(0.002)
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 survey data, which measures whether households report having the source. 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 at one of the three survey waves. Standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Demand for Electricity: Estimates of Linear Stage

	OLS	Price IV	Price & Hours IV		
		RCT	RCT	BLP	Hausman
	(1)	(2)	(3)	(4)	(5)
Price (Rs. 100)	-0.25**	-1.70***	-1.70***	-0.93	-3.24
	(0.12)	(0.63)	(0.63)	(5.80)	(2.80)
Hours of peak supply	0.20	0.11	0.21	0.25	0.14
	(0.21)	(0.21)	(0.27)	(0.32)	(0.34)
Hours of off-peak supply	-0.092*	-0.078*	-0.11*	-0.12	-0.097
	(0.047)	(0.046)	(0.058)	(0.074)	(0.073)
ξ_{tj} mean effects	Yes	Yes	Yes	Yes	Yes
Observations	999	999	999	945	989
First-stage F-Stat		42.1	21.1	0.4	1.0

The table presents estimates of the second, linear stage of our demand system (equation 6). The dependent variable is mean indirect utility at the market-by-survey wave level, estimated in the non-linear first stage. Peak hours refers to electricity supply during the evening, from 5 to 10 pm, and off-peak to other hours of the day. The columns estimate the same equation either by ordinary least squares (column 1) or instrumental variables (columns 2 to 5). Each column uses a different set of instruments. In column 2, we use the experimental treatment assignments interacted with a dummy for the endline survey as instruments (equation 7). In column 3, we additionally instrument for hours of supply, on-peak and off-peak, using the predicted hours of supply based on supply in nearby villages. In columns 4 and 5 we replace the experimental instruments with instruments from the industrial organization literature (Berry, Levinsohn and Pakes, 1995; Nevo, 2001; Hausman, 1996). Column 4 uses the average characteristics of the other products available in a given village as instruments (Berry, Levinsohn and Pakes, 1995). The characteristics we use are hours of supply and load. Column 5 uses the average price of each product in the nearest three villages as instrument for its price in a given village (Nevo, 2001; Hausman, 1996). All regressions control for wave-by-source fixed effects. The final row of the table reports the first-stage F-statistic from the price equation. Standard errors are clustered at the village-level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Price Elasticities of Electricity Source Demand

	Elasticity of share for source:				
			Own	Micro-	
	Grid	Diesel	solar	grid	None
with respect to price of source:	(1)	(2)	(3)	(4)	(5)
Grid	-0.58	0.31	0.54	0.14	0.14
Diesel	0.06	-1.83	0.24	0.03	0.04
Own solar	0.20	0.44	-1.91	0.12	0.09
Microgrid	0.15	0.16	0.32	-1.58	0.18

The table presents aggregate own- and cross-price elasticities of demand by electricity source. The arc elasticities are calculated using a 10% increase in each source's price from its mean endline price. The elasticities are calculated for the market share of each column source with respect to the price of each row source.

Table 7: Electricity Source Choice Probabilities by Household Characteristics

	Simple Model					Full Model				
	Grid	Diesel	Own Solar	Micro-grid	None	Grid	Diesel	Own Solar	Micro-grid	None
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Number of adults	0.032	0.004	0.003	0.007	-0.045	0.036	0.002	0.001	0.005	-0.045
	(0.006)	(0.004)	(0.002)	(0.003)	(0.004)	(0.009)	(0.006)	(0.001)	(0.004)	(0.008)
Solid roof (=1)	0.208	-0.016	0.005	-0.015	-0.183	0.107	-0.007	0.005	-0.007	-0.098
	(0.039)	(0.019)	(0.008)	(0.008)	(0.017)	(0.025)	(0.013)	(0.003)	(0.007)	(0.018)
Solid house (=1)						0.077	-0.004	-0.002	-0.005	-0.066
						(0.023)	(0.013)	(0.003)	(0.008)	(0.019)
Number of rooms						0.026	0.011	0.003	0.003	-0.043
						(0.008)	(0.006)	(0.001)	(0.004)	(0.008)
Owns ag. land (=1)						0.049	-0.023	0.003	0.008	-0.036
						(0.018)	(0.010)	(0.003)	(0.009)	(0.016)
Education of household head (1-8)						0.026	0.008	-0.001	0.002	-0.036
						(0.008)	(0.005)	(0.001)	(0.004)	(0.007)
Household income						0.016	0.002	0.001	0.014	-0.034
						(0.008)	(0.005)	(0.001)	(0.004)	(0.008)
Observations			8822					8822		
Log-likelihood			-5884.7					-5791.4		
LR index			0.031					0.047		
LR test statistic								186.6		
LR test p-value								0.000		

The table shows the effects of household characteristics on the probability of a household choosing a given electricity source. The table reports the results of two models. A simple model, reported in columns 1 through 5, includes as covariates the number of adults in the household and a dummy variable for whether the household has a solid roof. Our full model, reported in columns 6 through 10, includes five additional observable proxies for household demand: whether the household has a solid house, the number of rooms in the house, household income, whether the household owns agricultural land, and years of education of the household head. The effects of household characteristics are nonlinear. The table therefore reports "marginal" effects evaluated for a "poor" household, lacking the binary indicators of wealth and with an income at the 20th percentile. The profile of a poor household is defined as a household of two adults living in an one-room house, without a solid roof or walls, and lacking agricultural land ownership. See Appendix Table C6 for the characteristics of a poor household. The marginal effects are not truly marginal; for binary variables, we report the effect on choice probability of changing the value from zero to one, and for continuous variables the effect of an one standard deviation increase in that variable. To assess the goodness-of-fit of each model, we report a likelihood ratio index, which is defined as $\rho = 1 - \mathrm{LL}(\hat{\beta})/\mathrm{LL}(0)$, where $\mathrm{LL}(\hat{\beta})$ is the log-likelihood at the estimated parameters and LL(0) is the log-likelihood of a null model, where we constrain all the household characteristic coefficients to be zero. We also report the maximized value of the log-likelihood for both models and a likelihood ratio test statistic, distributed χ^2 with 20 degrees of freedom, from a test of the restriction that the coefficients on the covariates added in the full model are jointly zero.

Table 8: The Value of Electrification under Counterfactual Policies

	Market shares					Surplus (INR per household per year)		
	Grid (1)	Diesel (2)	Own solar (3)	Micro-grid (4)	None (5)	Consumer (6)	Producer (7)	Total (8)
Panel A. Fit of model								
Data	24	3	7	10	57			
Model	24	3	7	10	57	528	-497	31
Panel B. Value of solar innovation								
Solar nowhere	28	4	0	0	68	377	-581	-204
Solar everywhere	24	3	6	11	56	537	-495	42
Solar cost falls	21	2	12	14	51	594	-433	161
Solar cost falls and grid nowhere	0	3	23	16	57	471	-0	471
Solar subsidies	24	2	9	12	53	559	-497	62
Panel C. Grid extension								
No grid or solar	0	6	0	0	94	58	-0	58
Grid nowhere	0	4	18	12	66	361	-0	361
Grid everywhere	39	2	3	9	47	644	-803	-159
Increase peak hours	29	3	5	10	53	607	-775	-168
Grid everywhere and solar cost falls	34	2	7	13	44	692	-715	-23
Grid everywhere, increase peak hours and solar cost falls	45	1	5	11	38	810	-1171	-361
Panel D. Grid reforms								
Remove theft by raising grid price	8	4	13	12	63	400	-105	294
Remove theft for APL	19	3	8	11	59	490	-359	131
Remove theft using Proxy Means Test	13	4	11	11	61	433	-209	224
Maximum peak supply and raise prices	21	3	7	10	58	512	-497	15
Panel E. Demand growth								
All households at least median	27	3	8	10	51	616	-575	40
All households at least rich	35	4	10	11	40	843	-735	108
All households at least rich and big push	66	1	5	10	19	1311	-1714	-403

The table presents market shares and surplus under counterfactual changes in the electricity market. The counterfactual scenarios are laid out in Section 5 of the text and the detailed assumptions behind the counterfactuals are in Appendix Table D14. All counterfactuals are calculated using the full demand model estimates of Table 7, columns 6 through 10. For each counterfactual, columns 1 to 5 give the market shares of each source, and columns 6 through 8 give consumer, producer and total surplus. Consumer surplus is the amount in INR per household per year that households would be willing to pay for a given choice set, relative to having only the outside option of no electricity. The amounts of both consumer and producer surplus are averaged over the entire sample of consumers, regardless of their choice. Producer surplus is the variable profit of the state utility that provides grid electricity.

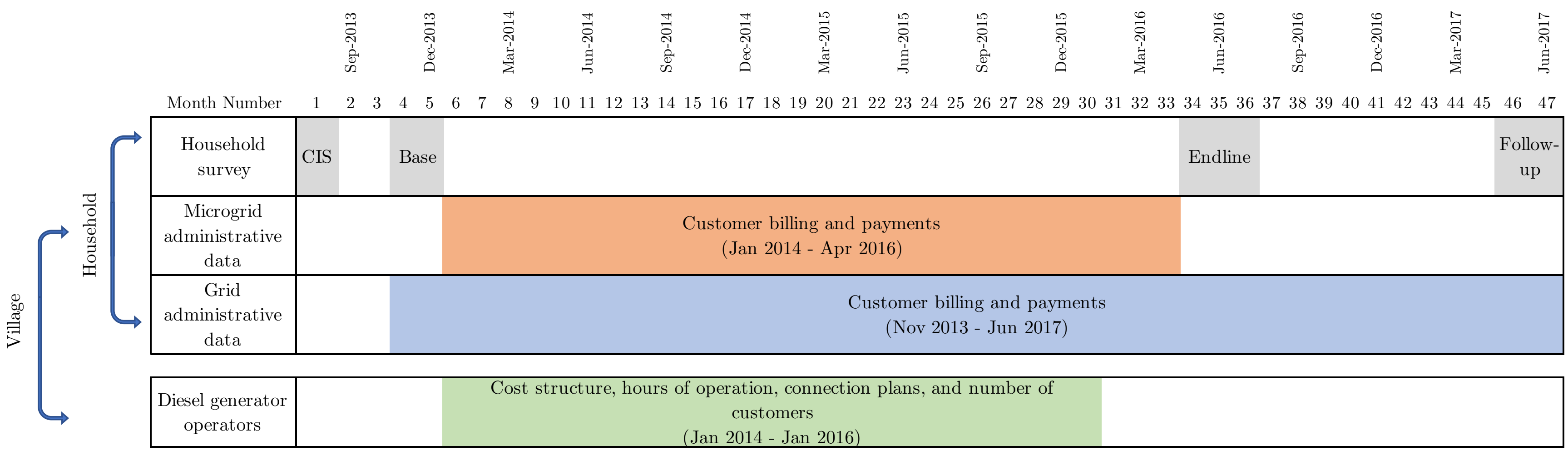
A Appendix: Data

This Appendix describes our data collection and the construction of instrumental variables for hours of electricity supply. We also provide additional summary statistics.

a Sampling and timeline

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

Figure A1: Data Collection Timeline



The figure shows the collection of data during the study from several sources: a household survey, microgrid administrative data, grid administrative data and a diesel operator survey. See Section 2 a 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.

b Construction of hours of supply for grid

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

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

c 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.

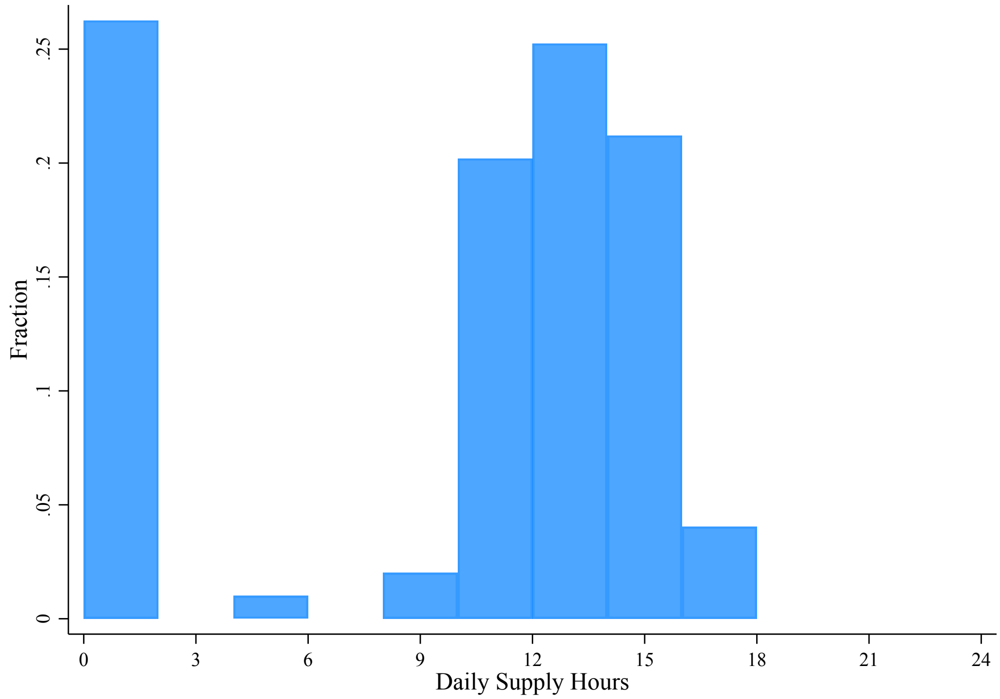
The instruments for peak and off-peak hours of supply are the predicted peak and off-peak hours of supply for a given village based upon hours of supply to nearby villages, as described in Appendix Section b above. We expect hours of supply to nearby villages to be correlated since they are served by 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.

d Summary statistics

Appendix Figure A2 shows the distribution of daily hours of electricity supply on the grid. Appendix Table A1 shows the market shares of electricity sources at endline accounting for the possibility of ownership of multiple sources.

Figure A2: Daily Hours of Supply on the Grid



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

Table A1: Electricity Source Ownership at Endline

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

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

B Appendix: Impact Analysis of Solar Microgrids

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

We estimate the impact of access to microgrid electricity on two sets of welfare outcomes, corresponding to anticipated direct and indirect effects of microgrid access. We begin by estimating the following specification:

$$y_{ivt} = \alpha + \beta_N T_{Normal,v} + \beta_S T_{Subsidized,v} + \mathbf{x}_i'\delta + \epsilon_{itv} \tag{8}$$

where y_{ivt} is the outcome of interest for household i in village v at time t, $T_{Normal,v}$ and $T_{Subsidized,v}$ are indicators equal to one when village v was assigned to be offered solar microgrids at normal and subsidized prices, respectively, and $\mathbf{x}_i$ is a vector of household-level controls from the baseline survey.

Table B2: Household Electricity Use Outcomes

	Daily Hours of Electricity Use (1)	Light Bulb Ownership (=1) (2)	Mobile Phone Ownership (=1) (3)	Price of Full Charge (Rs.) (4)
Treatment: Subsidized price	0.94***	0.15***	0.034**	-0.67***
	(0.24)	(0.047)	(0.014)	(0.24)
Treatment: Normal price	0.52**	0.098**	0.022	-0.46*
	(0.20)	(0.044)	(0.013)	(0.23)
Baseline controls	Yes	Yes	Yes	Yes
Control mean	1.16	0.32	0.88	4.72
Observations	2868	3001	3001	964

The table shows the impact of treatment status on household level electricity use, light bulb ownership, mobile phone ownership, and cost of a full charge of a mobile phone. The specifications include baseline electricity source indicators, baseline monthly income, and baseline equivalent of outcome variable as controls. Standard errors are clustered at the village-level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2 reports the estimates of the effects of microgrid treatment assignments on direct measures of electricity access and use. Micrograms power two low-wattage LED lamps and one mobile charging point, which was provided with every connection, so we expect microgrids to have effects

Table B3: Household Income, Education and Health Outcomes

	Monthly Income	Standardized Test Score		Respiratory Problems (=1)	
	(INR '000s)	Reading	Math	Adults	Children
	(1)	(2)	(3)	(4)	(5)
Panel A. Reduced Form					
Treatment: Subsidized price	0.18	0.11*	0.095	0.026	0.012
	(0.31)	(0.061)	(0.065)	(0.021)	(0.0082)
Treatment: Normal price	0.63*	0.020	0.071	0.017	0.0041
	(0.33)	(0.061)	(0.062)	(0.018)	(0.0082)
Panel B. Instrumental Variables					
Hours of electricity	0.15	0.22	0.21	0.027	0.014
	(0.35)	(0.24)	(0.23)	(0.027)	(0.012)
Baseline controls	Yes	Yes	Yes	Yes	Yes
Control mean	7.4	0	0	0.14	0.024
Observations	2692	646	637	2710	2669

The table shows the effects of provision of solar microgrids on social and economic outcomes, for health, education and test scores. Panel A of the table is the reduced-form or intent-to-treat effect of solar microgrids for these outcomes, and Panel B is the instrumental variable estimate of the coefficient on hours of electricity using the two treatment assignment dummies as instruments. We find no evidence that respiratory problems decrease for adults or children (Panel B, columns 4 and 5). The predominant source of indoor air-pollution comes from cooking, which is unaffected by the provision of microgrids, and we do not find significant declines in kerosene expenditure (not reported). Effects on standardized test scores are positive but imprecisely estimated (columns 2 and 3). For example, we estimate that an hour of additional electricity use increase children's reading scores by 0.22 standard deviations (standard error 0.24 standard deviations). This is a fairly large standardized effect but imprecise due to low first-stage take-up and the children tested being only a subsample of the overall experiment. We cannot rule out a zero effect or a significant positive effect of lighting on child test scores. Finally we find that electricity has a null effect on household income of INR 150 per month (standard error 350), which is small compared to baseline income of INR 7,400 per month. Test score results are at the child level. The regressions include baseline electricity source indicators, baseline monthly income, and baseline equivalent of outcome variable controls. Standard errors are clustered at the village-level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

mainly on lighting and mobile phone charging and use. Subsidy treatment households increased hours of electricity use by an estimated 0.94 hours per day (standard error 0.24 hours per day) relative to 1.16 hours in the control group at baseline (column 1). We find that assignment to a subsidy treatment village increases light bulb ownership by 15 pp (standard error 4.7 pp) relative to an ownership rate of 32 pp in the control group at baseline (column 2). The effects of being assigned to a normal price village are smaller but have the same sign and are also statistically significant. Households assigned to a subsidy treatment village are also more likely to own a mobile phone, by 3.4 pp relative to an already high control group ownership rate of 88 pp (column 3) (control households are thus 2.75 times as likely to own a mobile phone as a light bulb). Finally, assignment to a subsidy treatment village also decreases the amount of money households spend charging their mobile phones. Households without electricity typically charge their phones at a shop, at a higher unit cost of energy than would apply if they can charge at home.

Table B3 reports estimates of the effects of microgrids on welfare measures that may be indirectly affected by microgrid access. We consider three measures. First, household income (column 1). Second, children's test scores (columns 2 and 3). We administered basic reading and math tests, drawn from Pratham's Annual Status of Education Report survey instrument, to the children of treated households at both baseline and endline. Third, self-reported respiratory distress among adults and children (columns 4 and 5). Panel A of the table gives the intent-to-treat estimates and panel B gives instrumental variables estimates where the microgrid treatment assignments instrument for the hours of daily electricity use in the household.

Across all of these indirect welfare measures, we do not find evidence that solar microgrids improved outcomes. For example, the estimated effect of microgrids on household income is INR 0.15 thousand (standard error INR 0.35 thousand) per month, on a base of INR 7.4 thousand per month. The main caveat to this uniform finding of no indirect welfare effects is that our estimates for children's test scores are very imprecise. Because the demand for microgrids was modest, only a limited number of households took them up, and so the instrumental variables estimates are imprecise. The point estimates for both reading and math test scores are that one additional hour of electricity supply leads to about a 0.2 standard deviation increase in scores; the ITT effect on reading scores of being in a subsidized price village is 0.11 standard deviations (standard error 0.061 standard deviations). We therefore believe that the possible links between electrification and children's study habits and learning deserve further study.

In short, we find that when solar microgrids are made available, households use more electricity, purchase more mobile phones and light bulbs, and spend less money charging their phones. These effects are in line with the services provided by microgrids and are more pronounced in the subsidized treatment arm, where microgrids achieved higher market shares. We do not find significant effects of microgrids on measures of household welfare that may be indirectly affected by microgrid access.

C Appendix: Additional Results on Demand

This section presents additional results on demand. Subsection a presents estimates of the first stage from the estimation of the second, linear part of our structural demand model. Subsection b shows the calculation of microgrid surplus using several alternative functional forms to specify the aggregate demand for microgrids. Subsection c 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 d provides additional estimates to check the robustness of the structural demand estimates to alternative nest structures in the nested logit model.

a First stage estimates for structural demand model

Table C4 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; Nevo, 2001).

Table C4: First-Stage of Linear Estimation of Demand for Electricity

	Price IV	Price & hours IV			Price & hours IV BLP	Price & hours IV Hausman
	Price	Price	Peak hours	Off-peak hours	Price	Price
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment normal price X Endline	0.064**	0.064**	0.0050	-0.0046		
	(0.029)	(0.028)	(0.0051)	(0.030)		
Treatment subsidy price X Endline	-0.16***	-0.16***	0.0075	0.014		
	(0.021)	(0.021)	(0.0063)	(0.031)		
Hours of peak supply	-0.050					
	(0.049)					
Hours of off-peak supply	0.0081					
	(0.013)					
Peak hours instrument		-0.032	0.94***	0.19	-0.040	-0.041
		(0.045)	(0.063)	(0.15)	(0.043)	(0.043)
Off-peak hours instrument		0.0040	0.032**	0.88***	0.0057	0.0058
		(0.0094)	(0.013)	(0.030)	(0.0090)	(0.0089)
Total supply of competing products					-0.0017	
					(0.0060)	
Load of competing products					0.0079	
					(0.011)	
Avg price in nearby villages						-0.30
						(0.25)
ξ_{tj} mean effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	999	999	999	999	945	989
First-stage F-Stat	42.1	21.1	524.1	1057.2	0.4	1.0
Control mean	0.95	0.95	4.09	3.07	0.95	0.95

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

b Calculating the surplus from microgrids under alternative demand specifications

Table C5 shows estimates of the surplus from the presence of microgrids in the market (equivalently, the loss from the removal of microgrids) under different assumptions on the functional form of the demand curve. The first two rows of the table show estimates of the demand, or are omitted, when the estimated demand is too complex to be summarized in a couple of coefficients. Columns 1 through 3 show different reduced-form estimates of the demand curve, using either a piecewise-linear, linear, or log-linear demand specification. Column 4 shows the surplus calculations using the full structural model. For columns 2 and 3, we present the estimated price coefficient and constant. The dependent variable for the estimated demand curves is the market share of microgrids. The price of microgrids is instrumented by the experimental treatment assignment.

Row 3 and onwards in the table show the surplus gains from microgrids offered at various prices; that is, the area under each estimated demand curve, scaled to units of INR per household per year. We take the area under the demand curve up to a maximum price of INR 500, at which point we assume that demand extends horizontally to the vertical axis, for all functional forms. We present surplus estimates at four sets of prices: INR 100, the subsidized price in the experiment, INR 160 and INR 200, the normal prices in the experiment, and the endline prices. The prices at endline were a combination of INR 100 in subsidized price villages, INR 160 or 200 in normal price villages, and effectively infinity (we use INR 500) in villages where microgrids were not offered. Therefore, the surplus calculations in the final set of rows are for introducing microgrids at the prices that they were sold at in the experiment, after accounting for no availability in control villages.

Table C5: Consumer Surplus from Microgrid

	Piecewise-linear (1)	Linear (2)	Log-linear (3)	Structural (4)
Price		-0.001**	-0.997***	
		(0.001)	(0.386)	
Constant		0.347***	2.511	
		(0.091)	(1.929)	
Surplus 100 INR	225	222	242	215
Microgrid share 100 INR	22	22	13	23
Surplus 160 INR	68	92	172	94
Microgrid share 160 INR	22	14	8	11
Surplus 200 INR	4	37	138	52
Microgrid share 200 INR	5	9	6	7
Surplus endline	83	91	129	93
Microgrid share endline	11	11	6	10

The table compares various reduced-form estimates of consumer surplus from microgrid vs the structural estimate. Columns 1 through 3 apply a piecewise-linear, linear, and log-linear specficiation to the endline survey data, which measures whether households report having microgrid. In each linear specification, we instrument for price using a treatment dummy for being in the subsidized treatment arm. The final column corresponds to our structural model estimate of consumer surplus, which is calculated as the difference in surplus from having a microgrid price of INR 100 (or INR 160 or INR 200) vs the chock price, which we estimate to be INR 500. The reported surplus is the average surplus in INR per household per year. Standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The main finding in the table is that the reduced-form and structural estimates of demand yield very similar surplus values for microgrids. Our reduced-form estimates, at endline prices, range from a surplus of INR 83 to INR 129, depending on the functional form of the demand curve. Our structural estimate is INR 93, within this range and very close to the estimate of INR 91 obtained using a linear demand curve.

c Marginal effects for alternative household profiles

Table 7 presents the marginal effects of household characteristics on electricity choice probabilities for a "poor" household. Section 5 shows the results of counterfactuals where we increase the income and wealth of households from "poor" to "median" and "rich" levels. This subsection defines these household profiles and shows marginal effects for alternative household profiles to complement the estimates in the main text.

Table C6 shows the characteristics of households that are used to create the three profiles of household covariates. The number of adults (column 1) is integer valued, house characteristics (2 and 3) are indicator variables, the number of rooms is integer valued (4), agricultural land ownership is an indicator variable (5), literacy is integer valued (6) and income is continuous. Each row gives the values that these variables take on for each of the three household profiles we use to calculate marginal effects and to run counterfactuals.

The levels of these variables were chosen in order to roughly place a household, on an univariate basis, at the 20th, 50th and 80th percentile of the income or wealth distributions. Table C8 shows detailed summary statistics for the household covariates that enter our demand model in order to place the household profiles in context.

To calculate the marginal effects of these covariates on choice probabilities, we change their values by either one unit, for dummy variables, or one standard deviation, for integer valued and continuous variables. Table C7 shows the changes that this entails for each household covariate that enters the profiles. For the binary variables (Pukka, roof, land), this approach necessarily means that we cannot calculate the discrete effect for these variables when they are already equal to one in a given profile. For example, we cannot calculate the impact of having a roof for a median household, as a median household already has a roof. We therefore omit these entries from the corresponding tables of marginal effects.

Table C6: Profile Details

	Adults (1)	Pukka (2)	Roof (3)	Rooms (4)	Land (5)	Liter-acy (6)	Income (INR) (7)
Poor	2	0	0	1	0	1	3750
Median	3	0	1	2	1	1	6000
Rich	5	1	1	3	1	5	9500

This table details the characteristics for a poor, median, and rich household. Each profile was constructed by independently taking a fixed percentile of each column attribute. The fixed percentiles corresponding to poor, median, and rich are 20, 50, and 80, respectively. For example, a poor household lives in a 1-room dwelling, which corresponds to the 20th percentile of households in our sample across all survey waves.

Table C7: Definition of Household Characteristics and Magnitude of Marginal Change

Characteristic	Definition	Marginal Change (Poor)
Adults	Adults in the household	1 SD (1.82 persons)
Pukka	Indicator for Pukka house	0 to 1
Roof	Indicator for solid roof	0 to 1
Rooms	Number of rooms in the house	1 SD (1.32 rooms)
Land	Indicator for agricultural land	0 to 1
Education	Education of household head (1-8)	1 SD (2 levels)
Income	Monthly income	1 SD (INR 6486)

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 7). Base profiles for a representative poor, median, and rich household can be found in Table C6. 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.

Table C8: Summary Statistics of Household Characteristics

	Mean (1)	Median (2)	Q1 (3)	Q3 (4)	SD (5)	Min (6)	Max (7)
Adults in the household	3.67	3	2	5	1.83	1	15
Indicator for pukka house	0.32	0	0	1	0.47	0	1
Indicator for solid roof	0.51	1	0	1	0.50	0	1
Number of rooms in the house	2.45	2	2	3	1.32	1	11
Indicator for agricultural land	0.63	1	0	1	0.48	0	1
Education of household head (1-8)	2.48	1	1	4	2.04	1	8
Monthly income (INR)	7576	6000	4000	8500	6486	0	65000
Observations	8822	8822	8822	8822	8822	8822	8822

The table summarizes each of the household covariates used in our structural demand estimation. Each observation is for a household in a specific survey wave.

Tables C9 and C10 show the estimated discrete effects for a median and rich household, respectively, to be compared to Table 7 in the main text. The main finding is that, at all levels of household income, the discrete effects of increasing income or wealth proxies is to increase the demand for grid electricity and decrease, or barely alter, the demand for other sources of electricity. The discrete effects of household characteristics on choice probabilities are slightly smaller for rich than for poor households on some measures (e.g., the effect of income on grid choice), though these differences are small and not generally statistically significant. This relative lack of attenuation may reflect that even rich households, in our sample, have far from complete take-up of any electricity source.

Table C9: Impact of Household Characteristics on Choice Probabilities (Median Household)

	Grid (1)	Diesel (2)	Own Solar (3)	Microgrid (4)	None (5)
Number of adults	0.047	-0.001	0.001	0.003	-0.049
	(0.010)	(0.003)	(0.002)	(0.004)	(0.008)
Household income	0.019	0.000	0.001	0.014	-0.034
	(0.010)	(0.003)	(0.002)	(0.005)	(0.008)
Household owns land	-	-	-	-	-
Education of household head	0.036	0.004	-0.002	0.000	-0.039
	(0.010)	(0.003)	(0.002)	(0.004)	(0.008)
Pukka (solid) house	0.098	-0.006	-0.006	-0.009	-0.076
	(0.025)	(0.008)	(0.005)	(0.009)	(0.020)
Solid roof	-	-	-	-	-
Number of rooms	0.035	0.005	0.004	0.001	-0.046
	(0.010)	(0.004)	(0.002)	(0.004)	(0.009)

The table shows the discrete effects of changes in household observable characteristics (in rows) on the probability the household will purchase different electricity sources (in columns). The household characteristics are from our survey. The changes in choice probabilities are calculated with the demand model, for which the estimated coefficients are presented in Appendix Table C12, column 2. Each cell entry is the change in choice probability for a median household from increasing the row characteristics. For discrete household characteristics, the increase is from zero to one. For continuous household characteristics, the increase is for one standard deviation. Appendix Table C6 describes the statistical profile of a poor, median, and rich household. Standard errors are constructed using the delta method.

Table C10: Impact of Household Characteristics on Choice Probabilities (Rich Household)

	Grid (1)	Diesel (2)	Own Solar (3)	Microgrid (4)	None (5)
Number of adults	0.036	-0.003	-0.000	0.000	-0.033
	(0.008)	(0.003)	(0.001)	(0.004)	(0.005)
Household income	0.011	-0.001	-0.000	0.011	-0.022
	(0.008)	(0.003)	(0.001)	(0.005)	(0.006)
Household owns land	-	-	-	-	-
Education of household head	0.028	0.002	-0.002	-0.001	-0.026
	(0.008)	(0.003)	(0.001)	(0.003)	(0.004)
Pukka (solid) house	-	-	-	-	-
Solid roof	-	-	-	-	-
Number of rooms	0.026	0.003	0.002	-0.001	-0.030
	(0.009)	(0.004)	(0.002)	(0.004)	(0.006)

The table shows the discrete effects of changes in household observable characteristics (in rows) on the probability the household will purchase different electricity sources (in columns). The household characteristics are from our survey. The changes in choice probabilities are calculated with the demand model, for which the estimated coefficients are presented in Appendix Table C12, column 2. Each cell entry is the change in choice probability for a rich household from increasing the row characteristics. For discrete household characteristics, the increase is from zero to one. For continuous household characteristics, the increase is for one standard deviation. Appendix Table C6 describes the statistical profile of a poor, median, and rich household. Standard errors are constructed using the delta method.

d Robustness of demand estimates

Table C11 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 4 estimates in the main text but use administrative data on payments rather than survey data on source usage as the measure of demand. The estimated market share in subsidized price villages is very similar across both data sources, while the estimated market share in normal price villages is higher in the survey data than in the payments data. Payments for microgrids may differ from survey reports due to measurement error or because households still use microgrids, for a time, even after they have stopped paying the monthly price. We understood from our field work that the pace at which HPS repossessed systems for non-payment was slow.

Table C11: Solar Microgrid Demand by Village Treatment Arm

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

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

We specify a nested logit demand model, which requires an ex ante choice of nest structure. Since different sources of electricity differ on multiple dimensions, it is not obvious for what sources demand should be expected to unobservably correlate. Table C12 presents coefficients from the non-linear part of the demand model for alternative nest structures. Column 1 gives coefficients from a multinomial logit model, in which there is no correlation between the unobserved taste shocks for different sources. Columns 2 and onwards then give coefficients from nested logit models with varying nest structures. The last two coefficients in the table are the parameters that govern the correlation of the logit shocks within nests.

There are two main points from the table. First, we have chosen the column 2 specification as our main specification, since it achieves the highest log likelihood. Our main estimates therefore come from using a nest structure that assigns microgrids and the outside option to their own nests, and grid, diesel and own solar to a third nest. Second, while these estimates achieve the highest log likelihood, the assumed nest structure has very small effects on both the log likelihood of the model and the coefficients on household characteristics. All of the models, including the multinomial logit model, yield very similar coefficients on household observables.

Since the nest structure may affect the coefficients on observable characteristics, in the non-linear part of the structural model estimation, it may then change the dependent variable and estimates in the second, linear part of the structural model also. Table C13 shows estimates of the second, linear part of the structural demand model, where alternative nest structures have been used in the first stage. The coefficient on price is very close to our main estimate across all specifications. The coefficient on hours of peak supply is somewhat more variable, but imprecise, and not significantly different from our main estimate in any alternative specification.

Table C12: First-Stage Estimation Results for Alternative Nest Specifications

γ_{jr}	Multinomial Logit (1)	{Grid, Diesel, Own solar} & {Microgrid} (2)	{Grid, Diesel, Microgrid} & {Own solar} (3)	{Grid, Own solar, Microgrid} & {Diesel} (4)	{Grid} & {Non-Grid} (5)	{Grid, Own solar} & {Diesel, Microgrid} (6)	{Grid, Diesel} & {Solars} (7)
Grid × Income	0.21	0.19	0.21	0.21	0.20	0.21	0.21
	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)
Diesel × Income	0.14	0.16	0.14	0.14	0.21	0.14	0.14
	(0.08)	(0.06)	(0.09)	(0.08)	(0.05)	(0.14)	(0.08)
Own solar × Income	0.16	0.18	0.16	0.17	0.20	0.16	0.17
	(0.07)	(0.06)	(0.07)	(0.07)	(0.05)	(0.13)	(0.20)
Microgrid × Income	0.50	0.49	0.50	0.48	0.25	0.50	0.49
	(0.11)	(0.11)	(0.14)	(0.13)	(0.00)	(0.13)	(0.49)
Grid × Land	0.24	0.20	0.24	0.24	0.24	0.24	0.24
	(0.09)	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Diesel × Land	-0.15	-0.09	-0.15	-0.15	-0.01	-0.15	-0.15
	(0.12)	(0.11)	(0.12)	(0.12)	(0.00)	(0.15)	(0.12)
Own solar × Land	0.15	0.18	0.15	0.16	0.11	0.15	0.16
	(0.11)	(0.09)	(0.11)	(0.12)	(0.00)	(0.17)	(0.22)
Microgrid × Land	0.21	0.20	0.21	0.20	0.08	0.21	0.19
	(0.17)	(0.17)	(0.17)	(0.16)	(0.06)	(0.17)	(0.42)
Grid × Adults	0.12	0.12	0.12	0.12	0.12	0.12	0.12
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Diesel × Adults	0.09	0.09	0.09	0.09	0.08	0.09	0.09
	(0.04)	(0.03)	(0.04)	(0.04)	(0.01)	(0.06)	(0.03)
Own solar × Adults	0.09	0.10	0.09	0.09	0.10	0.09	0.09
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.06)	(0.03)
Microgrid × Adults	0.09	0.09	0.09	0.09	0.10	0.09	0.09
	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)	(0.04)	(0.05)
Grid × Pukka	0.42	0.36	0.42	0.41	0.42	0.42	0.42
	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
Diesel × Pukka	0.13	0.20	0.13	0.13	0.09	0.13	0.15
	(0.14)	(0.11)	(0.14)	(0.14)	(0.10)	(0.16)	(0.15)
Own solar × Pukka	0.11	0.16	0.11	0.11	0.09	0.11	0.11
	(0.12)	(0.10)	(0.12)	(0.12)	(0.09)	(0.15)	(0.32)
Microgrid × Pukka	-0.01	-0.00	-0.00	0.02	0.12	-0.00	0.02
	(0.19)	(0.19)	(0.21)	(0.20)	(0.10)	(0.19)	(0.86)
Grid × Lit	0.10	0.08	0.10	0.09	0.10	0.10	0.10
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Diesel × Lit	0.09	0.08	0.09	0.09	0.06	0.09	0.09
	(0.03)	(0.02)	(0.03)	(0.03)	(0.00)	(0.06)	(0.03)
Own solar × Lit	0.02	0.05	0.02	0.02	0.04	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.03)	(0.00)	(0.03)	(0.02)
Microgrid × Lit	0.05	0.05	0.05	0.05	0.06	0.05	0.05
	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)	(0.04)	(0.04)
Grid × Roof	0.58	0.52	0.58	0.57	0.58	0.58	0.57
	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
Diesel × Roof	0.20	0.28	0.20	0.20	0.23	0.20	0.20
	(0.13)	(0.11)	(0.13)	(0.13)	(0.05)	(0.17)	(0.14)
Own solar × Roof	0.42	0.44	0.42	0.41	0.32	0.42	0.40
	(0.12)	(0.09)	(0.12)	(0.11)	(0.00)	(0.29)	(0.38)
Microgrid × Roof	0.00	0.00	0.00	0.03	0.22	0.00	0.02
	(0.18)	(0.18)	(0.19)	(0.20)	(0.00)	(0.18)	(0.84)
Grid × Rooms	0.13	0.14	0.13	0.14	0.13	0.13	0.13
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Diesel × Rooms	0.15	0.15	0.15	0.15	0.16	0.15	0.15
	(0.05)	(0.03)	(0.05)	(0.05)	(0.03)	(0.10)	(0.05)
Own solar × Rooms	0.18	0.17	0.18	0.18	0.16	0.18	0.18
	(0.04)	(0.03)	(0.04)	(0.04)	(0.03)	(0.11)	(0.06)
Microgrid × Rooms	0.10	0.10	0.10	0.10	0.15	0.10	0.11
	(0.07)	(0.07)	(0.07)	(0.06)	(0.03)	(0.07)	(0.15)
σ_1	-	0.55	0.01	0.10	0.82	0.01	0.10
		(0.20)	(0.28)	(0.32)	(0.00)	(0.43)	(0.42)
σ_2	-	-	-	-	-	0.01	0.08
						(0.39)	(5.06)
Observations	8822	8822	8822	8822	8822	8822	8822
Log likelihood	-5793.31	-5791.36	-5793.35	-5793.27	-5791.67	-5793.37	-5793.12
LR test statistic	-	3.91	-0.07	0.08	3.29	-0.10	0.38
LR test p value	-	0.05	1.00	0.78	0.07	1.00	0.83

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

Table C13: Linear Estimation of Demand for Electricity (Alternative Nests)

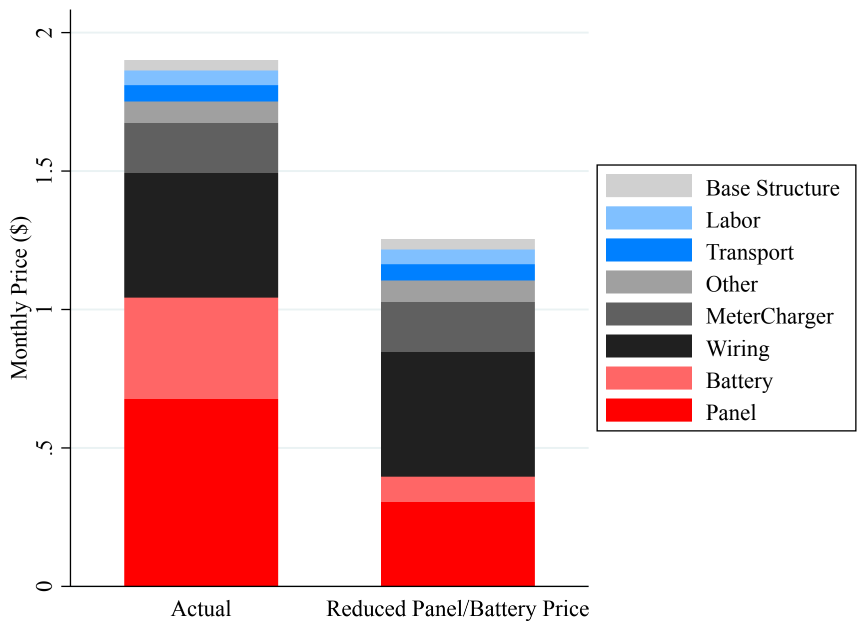
	Nested Logit			
	{Microgrid} {Non-Microgrid} (1)	{Solar} {Non-Solar} (2)	{Grid} {Non-Grid} (3)	Multinomial Logit (4)
Price (Rs. 100)	-1.70***	-1.60**	-1.74**	-1.58**
	(0.63)	(0.64)	(0.72)	(0.64)
Hours of peak supply	0.21	0.44	0.47	0.47
	(0.27)	(0.30)	(0.30)	(0.30)
Hours of off-peak supply	-0.11*	-0.16**	-0.17**	-0.17**
	(0.058)	(0.065)	(0.066)	(0.065)
ξ_{tj} mean effects	Yes	Yes	Yes	Yes
Observations	999	999	999	999
First-stage F-Stat	21.1	21.1	21.1	21.1

The table presents linear estimation of our demand system, using alternative nest structures in the non-linear estimation. The dependent variable are the mean indirect utilities, specific to each village and survey wave, which come from the non-linear first-stage estimation. The first column uses our preferred nest structure of grouping grid, diesel, and own solar in one nest and microgrid in its own nest. (The estimates in the first column are the same as those in column 3 of Table 5.) The second column uses a nest structure with grid and diesel in one nest and both solar technologies in another. In the third column, we group grid in its own nest and all non-grid technologies in a second nest. In the last column, we use the mean indirect utilities derived from a multinomial logit first-stage estimation. For all second-stage linear estimations, we instrument for price, peak hours, and off-peak hours. Peak hours refers to electricity supply during the evening (5 - 10pm). All regressions control for wave × source mean effects. Standard errors are clustered at the village-level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Appendix: Counterfactual Scenarios

This section gives additional details on our counterfactual scenarios. Figure D3 shows the breakdown of costs for solar microgrids, which we use to forecast the effects of declines in the prices of solar photovoltaic panels and batteries on solar systems. Because the capital cost of PV panels and batteries make up a significant, but incomplete, share of total cost, this breakdown is necessary to calculate the effect that any proportional decline in capital costs will have on the total costs of solar systems.

Figure D3: Microgrid Solar Cost Structure: Current and Predicted



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

For reference, Table D14 enumerates the assumptions in our counterfactual scenarios from Table 8. Column 1 gives the name of each scenario and columns 2 through 4 detail assumptions made in the row scenario regarding source availability, supply hours, pricing and subsidies, and any additional details.

Table D14: Counterfactual Analysis: Assumptions

Scenario	Source availability	Source hours (peak)	Other notes
Solar nowhere	Endline for grid and diesel, solar nowhere	Endline	Endline
Solar everywhere	Endline for grid and diesel, solar everywhere	Endline	Endline
Solar cost falls	Endline	Endline	Reduction in microgrid price from INR 170 to INR 120 (based on 2022 projection), proportional (30%) reduction in own solar price
Solar subsidies	Endline	Endline	Microgrid subsidy of INR 50 (INR 170 minus INR 120) for households in villages without grid access, proportional subsidy for own solar
Grid nowhere	Endline for diesel and solar, grid nowhere	Endline	Endline
Grid everywhere	Endline for diesel and solar, grid everywhere	Endline	Endline
Increase peak hours	Endline	Two additional peak hours for grid (capped at 5 hours), endline peak hours for all other sources	Endline
Remove theft by raising grid price	Endline	Endline	Grid at full price of INR 128, all else at endline. This removes the "theft adjustment", which results from applying the percent of households that answer affirmatively to "Do you pay your bill?" to the reported grid bill amount (INR 128)
Remove theft for APL	Endline	Endline	Grid at full price of INR 128 for households who are above the poverty line
Remove theft using Proxy Means Test	Endline	Endline	Grid at full price of INR 128 for households who have a solid roof
Maximum peak supply and raise prices	Endline	Grid peak hour of 5 everywhere, all else at endline	Grid at INR 103 everywhere, all else at endline. INR 103 is the price needed to maintain the same producer surplus as under the status quo (model)
All households at least X	Endline	Endline	Each household covariate is at least as large as it is under profile X where $X \subset \{\text{Median}, \text{Rich}\}$ (details on each profile can be found in Appendix Table C6)
Big push			This is an abbreviation for the combination of grid everywhere, increase peak hours, and solar cost falls