

# The Value of Grid Reliability: Evidence from a Natural Experiment in Senegal

Abdoulaye Cissé\*

This version: August 12, 2025

[Latest version is available here](#)

## Abstract

Governments and donors invest heavily in expanding access to electricity in low-income countries. However, recent studies document small impacts of access to grid electricity in several low-income countries. This paper provides new evidence on the impact of grid reliability on the willingness to pay (WTP) for grid electricity based on large-scale reliability improvement projects in Senegal's electricity supply network. I use high-frequency geocoded administrative data on outages, customer-level bills covering the universe of electricity users, and rich geocoded data on the economic activities of households and firms to conduct two analyses. First, I estimate the causal effect of improved reliability on consumers through a difference-in-differences approach that leverages the variation in the timing of reliability projects. I find that these projects reduced outage duration by 40%, increased electricity consumption by 8%, and led to greater appliance ownership, a reallocation of time away from household chores to wage labor, and lower disconnection rates. Second, I estimate a quantitative demand model and a random-coefficient discrete choice model of energy sources. I find that reliability projects increase the WTP for an additional hour of grid electricity by 10%, from an average of \$0.20. This increase varies over peak and off-peak hours and differs significantly across customer types (firms versus households) and by establishment size (small versus large). This finding corresponds to an internal rate of return for reliability projects exceeding 30%. My findings suggest that reliability improvements can significantly increase the value of grid electricity, even in contexts where baseline reliability and the WTP for grid electricity are low.

---

\*e-mail: [cisse@berkeley.edu](mailto:cisse@berkeley.edu). I am grateful to the members of my dissertation committee, namely, Severin Borenstein, Meredith Fowle, and Ethan Ligon, for their continuous guidance and feedback. Special thanks to Hunt Allcott, Francis Annan, Bocar Ba, Pierre Biscaye, Kirill Borusyak, Sofia Villa-Boas, Steve Cicala, Chi Man Cheung, Lucas Davis, Alain de Janvry, Joshua Deutschmann, Binta Zahra Diop, Kwabena Donkor, Pascaline Dupas, Gabriel Englander, Thibault Fally, Marco Gonzalez-Navarro, Kelsey Jack, Megan Lang, Jacob Lefler, Aprajit Mahajan, Edward Miguel, Melanie Morten, Abdoulaye Ndiaye, Khadidja Ngom, Leïla Njee Bugha, Elisabeth Sadoulet, James Sayre, Fatou Kiné Thioune, and Catherine Wolfram for their helpful comments and suggestions. Special thanks to the attendees of the UC Berkeley ARE Development Workshop, the UC Berkeley Development Seminar, the UC Berkeley ERE Seminar, the Pacific Development Conference, the LSE Environment Week, the LSE Environment Camp, the Columbia University Interdisciplinary Workshop in Sustainable Development, the UC Berkeley Energy Camp, the STEG Early Career Workshop, and NEUDC for their helpful discussions. I would like to thank Société Nationale d'Electricité du Sénégal (SENELEC) and Agence Nationale de la Statistique et de la Démographie (ANSD) for generously sharing their data and answering my questions. This paper is dedicated to the memory of Sara Johns. All errors are my own.

# 1 Introduction

Access to grid electricity has increased significantly over the last few decades in low-income countries. In sub-Saharan Africa (SSA), the share of the population with access to grid electricity increased from 26% in 2000 to 51% in 2022 owing to large-scale grid expansion programs such as Power Africa and Mission 300 (IEA et al., 2023). Development practitioners and donors generally have strong priors that access to grid electricity is a key factor for economic growth and rising standards of living. However, recent studies find minimal impacts of access to grid electricity on multiple economic outcomes (Lee et al., 2020; Burlig and Preonas, 2024b; Masselus et al., 2024).

Why does existing evidence not support the beliefs regarding the importance of electricity access? To explain this discrepancy, I study unreliable electricity supply. Indeed, despite the laudable efforts to achieve universal access to electricity, electricity reliability remains a major issue faced by firms and households: 39% of firms<sup>1</sup> in SSA identify electricity reliability as a major constraint, 72% experience major shortages (The World Bank, 2022), and 54% of connected households in SSA report regular electricity outages (Afrobarometer, 2019).<sup>2</sup>

This paper estimates the effects of electricity reliability on the willingness to pay (WTP) for grid electricity in Senegal, where the utility company has implemented reliability projects over time. Estimating the WTP for reliability is challenging in low-income countries due to the need for detailed data on consumers and the need for exogenous variation in electricity reliability, two conditions rarely found together. I use high-frequency, geocoded administrative data covering the universe of outages and electricity consumption—including over 125 million bills for more than 2 million clients across Senegal’s distribution network from 2012 to 2024—along with nationally representative geocoded household surveys covering 7,156 households. Mapping these datasets to the distribution network and leveraging the variation in the roll-out of reliability projects, I find that improvements in reliability increased electricity usage, appliance ownership, and the WTP for grid electricity.

To establish these results, this paper proceeds in two steps. The first step of my analysis estimates the causal impact of electricity reliability projects on the duration of outages and on various economic outcomes at the consumer level (electricity expenditures, location choice, bill payment, time use, and appliance ownership). These reliability projects involved installing remote terminal units (RTUs) on distribution stations, enabling the utility to pinpoint network incidents and resolve them more quickly.<sup>3</sup> I con-

---

<sup>1</sup>These are firms included in the World Bank Enterprise Surveys, which are limited to formal (registered) companies with five or more workers, along with other industry restrictions and ownership structures.

<sup>2</sup>These statistics are pooled indicators for the 45 SSA countries covered by the World Bank Enterprise Surveys and the 43 SSA countries covered by the Afrobarometer surveys.

<sup>3</sup>Stations are the last downstream element of the electricity distribution network before consumers.

struct the station-level treatment timing using detailed digitized maps of the electricity distribution network from 2015 to 2024 that show the gradual roll-out of reliability devices. My empirical analysis faces two primary challenges: (i) the potential spillovers of reliability projects on untreated stations that can threaten the validity of the use of untreated stations as a control and (ii) the selection and endogeneity of treatment. I therefore structure the empirical analysis in three parts.

First, I estimate the direct effects of reliability projects on outages at treated stations with a difference-in-differences design that compares early-treated stations with late-treated stations and with stations located on untreated feeders.<sup>4</sup> Identification requires the selection of distribution stations for reliability projects to be uncorrelated with time-varying unobserved determinants of outages. Second, to assess spillovers on untreated stations, I employ a triple difference-in-differences (DDD) design in which I compare untreated stations located upstream of early-treated stations to untreated stations located downstream of early-treated stations, before and after the reliability projects, relative to upstream and downstream untreated stations on late-treated feeders. Third, to address concerns about potential simultaneity between economic activity and reliability changes, I exploit the within-district variation in the treatment status of nearby stations that belong to different feeders, some of which have at least one station treated early while others have only late-treated stations or have no treated stations. For this third part, identification requires the economic outcomes of clients within the same district but served by stations belonging to different adjacent feeders to follow similar trends in the absence of reliability interventions.<sup>5</sup>

My empirical analysis builds on recent advancements in difference-in-differences methods for settings with multiple treatment periods. I specifically adopt the approach of [Callaway and Sant'Anna \(2021\)](#), who account for varying control group sizes over time and address likely serial correlation in error terms. To capture broader grid developments and concurrent projects, I include covariates that reflect these factors alongside the reliability projects of interest. My results withstand robustness tests, including those proposed by [Rambachan and Roth \(2023\)](#) to bound potential violations of post-treatment parallel trends<sup>6</sup>, as well as various other robustness checks.<sup>7</sup>

The second part of my analysis is a structural model that quantifies the value con-

---

They are often referred to as “distribution stations” or “transformer stations”. See Section 2 for more details on the setup of the grid network.

<sup>4</sup>Feeders are medium-voltage lines that connect locations from which electricity is sourced to distribution stations (or transformer stations) which serves directly end consumers. See Section 2 for more details on the setup of the grid network.

<sup>5</sup>In my setting, there are 47 districts in total.

<sup>6</sup>This includes testing both the relative magnitude of violations compared to pretreatment deviations and a linear extrapolation of the pretrend coefficient slope, providing robustness without assuming pretreatment parallel trends.

<sup>7</sup>These checks include alternative outcome variable specifications, the inclusion of control variables, different fixed effects and clustering approaches, and the use of alternative estimators.

sumers place on improved grid electricity availability. The quantitative model begins with an economic model linking electricity reliability to consumer welfare, where consumers derive utility from non-electric goods and services provided by electricity-powered appliances (e.g., lighting, refrigeration, heating). Though electricity itself is not directly consumed, it powers these services, with each consumer understanding the expected outage frequency at their distribution station. The model reveals two key channels through which reliability impacts electricity use: (i) a positive mechanical effect on electricity use (e.g. lights are turned on for longer periods when electricity reliability improves) provided that the intertemporal elasticity of substitution is sufficiently high, and (ii) a demand effect (i.e. appliance upgrades due to reliability changes), whose sign depends on changes in appliance usage when reliability changes. By disentangling these two effects, the model captures the direct consumer surplus from alleviating reliability constraints as well as the additional surplus arising from associated behavioral changes.

The economic model is used in two ways to estimate the monetary value of grid supply at baseline, and the change in this value coming from the reliability project. In a first estimation approach, I use the economic model to directly estimate the surplus associated with reliability projects by aggregating consumer surplus for firms and households before and after the reliability projects. I express consumer surplus as a function of three parameters: price elasticity of demand, technical efficiency of appliances and a parameter governing consumer preferences for electrical services over other goods and services. I estimate the price elasticity of demand with simulated price instruments from the industrial organization and public finance literature ([Saez et al., 2012](#); [Ito, 2014](#)), and calibrate the rest of the parameters using the bills data. The estimation and calibration leverage the richness of the billing data for the Senegalese electricity market and the spatial and temporal heterogeneity in price schedules across the universe of residential and commercial clients that these bills cover over a 12-year period. These elasticities are used to derive a demand curve under the assumption of quasi-linear utility.

In a second estimation approach, I turn this economic model into a discrete choice mixed logit model with random correlated coefficients of grid connection, via parametrization of the utility function and Taylor approximations. The discrete choice model incorporates the extensive margin of new connections to the grid. In the discrete choice model, households decide between having no grid connection at all, being connected to the grid, or relying solely on solar panels. Conditional on being connected to an energy source, consumers decide which appliances to use among low-energy-intensity appliances (light bulbs, TVs, fans, refrigerators, laptops, tablets, and phones) and high-energy-intensity appliances (air conditioners, freezers, irons, washing machines, and vacuums). The choice attributes comprise both the benefits associated with electricity availability during peak hours and off-peak hours and the costs incurred to consume energy services: electricity expenditures, connection costs, and appliance costs. I allow the coefficients to be random, to be correlated, and to interact with observ-

able consumer characteristics to account for heterogeneity in preferences for connection type and appliances across both observable and unobservable characteristics.

The discrete choice model is estimated using detailed geocoded household surveys via simulated maximum likelihood (SML) estimation. I use the estimated individual-level parameters to obtain a distribution of the WTP for an additional hour of electricity separately for peak hours and off-peak hours. To address concerns related to the endogeneity of electricity availability, I follow [Ito and Zhang \(2020\)](#) to instrument electricity availability with reliability projects, effectively considering only the change in reliability due to the grid upgrades to obtain the change in WTP caused by these projects.

My first set of results shows a positive impact of electricity reliability improvements on both households and firms. At baseline, the share of households connected to the grid was 61% and total duration of outages was 20 hours per month on average. First, I find that the reliability projects led to a 40% reduction in the duration of outages for treated stations, with strong spillovers of treated on untreated stations that are upstream of treated stations and smaller but significant spillover effects on untreated stations that are downstream of treated stations. This reduction in the duration of outages is accompanied by an 8% increase in electricity consumption in the 5 years following the projects, with effects that are stronger in later years than in earlier years post-treatment. This increase in consumption is not driven entirely by the mechanical effect of outages on electricity use. I document that the additional effect of outages on electricity use comes from customers increasing their ownership of common electric appliances such as phones, fans, and TVs. I also find that the reliability projects lowered the probability of households and firms closing their accounts and increased the likelihood of having new accounts. On the household side, I find that the reliability projects led to a reallocation of time away from households chores to labor market hours.

My second set of results, based on the direct estimation of consumer surplus, suggests that reliability projects increase consumer surplus per kilowatt-hour (kWh) by between 10% and 20%. For small households, surplus rose from approximately \$0.20 to \$0.22 per kWh, translating to an additional \$1.18 in value per household. Small firms experienced a similar improvement, with surplus increasing from \$0.22 to \$0.28 per kWh—equating to about \$2.19 per firm—while large firms saw an increase from \$0.44 to \$0.55 per kWh. In general, firms enjoyed higher consumer surplus than households both before and after reliability projects, and smaller units—both households and firms—tend to gain more from reliability projects relative to larger ones.

My third set of results from the estimation of the mixed logit model yields a baseline marginal WTP for an additional hour of grid electricity of \$0.40 for peak hours and \$0.10 for off-peak hours, on average. These estimates align with the few stated-preference estimates in the literature. My WTP estimates are similar in magnitude to retail electricity tariffs in the Senegalese electricity market, suggesting that when electricity is reliable, consumers in low-income countries are willing to pay market rates. I

also find that the change in WTP following reliability projects is 4 cents for peak hours and 1 cent for off-peak hours. This change in WTP induced by reliability projects corresponds, on average, to a 10% increase in WTP relative to its baseline levels before these projects. The change in WTP following reliability projects is higher for clients with higher baseline consumption and more appliances.

Considering the relatively low costs of the reliability projects that I study, the estimated surplus suggests a very high internal rate of return of 30–150%, depending on the accounting assumptions made. These high returns to investment suggest a short payback period of less than a few years for the benefits of reliability improvements to outweigh the costs incurred by the utility company, even under conservative assumptions absent significant adaptation and behavioral changes in the long run and accounting for installation and maintenance costs. Thus, the projects that I study differ from more common reliability projects deployed in low-income settings (e.g., line bifurcations, the undergrounding of lines) in terms of investment size and cost-effectiveness.

To the best of my knowledge, this paper is the first to estimate the causal effect of reliability on the WTP for grid electricity. Establishing this causal link has been challenging, as improved grid reliability influences the WTP by changing both consumers' usage and appliance adoption (intensive margin) and their choice of energy sources (extensive margin). Accounting for changes across both margins requires (i) variation in reliability that is exogenous to economic outcomes, (ii) consumer-level data on their behavior and characteristics, and (iii) observable choices regarding energy sources (grid versus alternatives). Satisfying these conditions allows this paper to contribute to three key areas of the economic literature.

First, existing studies on electricity reliability in low-income countries provide causal estimates but often focus on larger firms ([Allcott et al., 2016](#); [Fisher-Vanden et al., 2015](#)), smaller changes in reliability ([Mensah, 2024](#); [Berkouwer et al., 2024](#)), short durations ([Burlando, 2014](#)), or limited spatial areas ([Meeks et al., 2023](#)).<sup>8</sup> In contrast, this paper analyzes a large-scale natural experiment leading to significant, sustained reliability changes over the medium (2–3 years) and long term (4–6 years) across Senegal. The broader geographic scope enables me to go beyond local changes in grid quality and to directly test for spillover effects. My focus on large changes in reliability clarifies whether prior findings of modest impacts of reliability changes stem from truly low impacts or merely reflect limited reliability improvements.

Second, this paper quantifies the value of reliability for households and firms using a revealed-preference approach. Few studies measure the monetary value of reliability in low-income settings. Two related studies in areas with low baseline reliability are [Deutschmann et al. \(2021\)](#), who use a stated-preference choice experiment in Senegal,

---

<sup>8</sup>For further studies on the impact of reliability metrics on consumers, see [Hardy and McCasland \(2021\)](#), [Abeberese et al. \(2021\)](#), [Abeberese \(2020\)](#), [Grainger and Zhang \(2019\)](#), [Cole et al. \(2018\)](#), and [Scott et al. \(2014\)](#).

and [Khanna and Rowe \(2024\)](#), who estimate a revealed-preference measure in India by combining outage responses with price elasticity estimates. Departing from these methods, this study employs quasi-experimental reliability changes to measure welfare impacts for both households and firms, allowing for incentive-compatible consumer decisions without restrictive assumptions on external variables that affect welfare. More broadly, this paper contributes to the value of lost load (VoLL) literature, which generally estimates supply disruptions in high-reliability economies.<sup>9</sup> I extend VoLL research by using revealed preferences to estimate the value of reliability improvements (or “gained load”), providing a distribution of values of reliability changes instead of a single estimate.<sup>10</sup>

Third, this paper contributes to the literature on the WTP for essential goods in low-income countries by highlighting how service quality affects the low baseline WTP for a good that is essential to both households and firms. For electricity, the baseline WTP is often low and below the marginal selling price, which explains why some consumers remain unconnected despite available grid access ([Lee et al., 2020](#); [Burgess et al., 2022](#)). Recent macroeconomic studies also suggest that improving electricity-sector productivity yields only modest long-term gains, challenging the assumption that electricity is a primary constraint on development ([Colmer et al., 2023](#)). Leveraging a natural experiment, I demonstrate that the WTP increases with reliability improvements, a finding that is particularly relevant for low-income countries where consumers may be willing to pay more for public services if quality is improved. This suggests that last-mile electrification, now underway in many developing countries, can provide greater value to consumers if it is accompanied by enhanced reliability, which is achievable through cost-effective technological solutions.

The rest of this paper is organized as follows. Section 2 presents the context of the project. Section 3 presents the data used in the analysis. Section 4 discusses the reduced-form empirical strategy and results. Section 5 presents the economic model that links electricity reliability to consumer welfare, the direct estimation of consumer surplus and the mixed logit discrete choice model of energy types. Concluding remarks are made in Section 6.

---

<sup>9</sup>[Gorman \(2022\)](#) provides a comprehensive review of the VoLL literature, discussing challenges in VoLL measurement.

<sup>10</sup>According to [Gorman \(2022\)](#), only two out of more than 130 VoLL studies use revealed preferences; the notable exceptions are [Brown and Muehlenbachs \(2023\)](#) and [Harris \(2023\)](#), who use battery adoption and generator sales in the U.S. context.

## 2 Context

### 2.1 General Context

Electricity in Senegal is provided by a state-owned utility company, Société Nationale d'Electricité du Sénégal (SENELEC), and is regulated by the Commission de Régulation du Secteur de l'Electricité (CRSE). SENELEC is partially in charge of production and fully in charge of transmission and distribution. At the end of 2022, SENELEC had 2,198,624 customers and sold 4,863 GWh<sup>11</sup> to these clients on the basis of an increasing block pricing (IBP) tariff that the CRSE sets.<sup>12</sup>

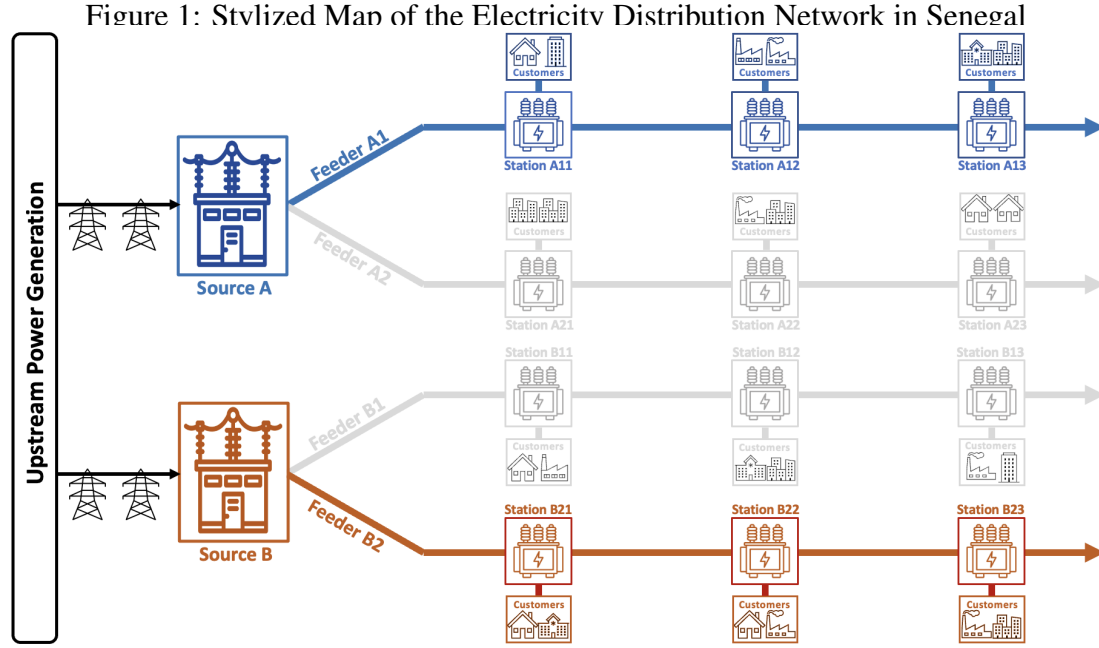
SENELEC's transmission and distribution network consists of five different parts: (i) transmission lines with high voltage levels of either 225 kV or 90 kV and transporting electricity from generation sources to substations, (ii) substations that transform high-voltage electricity received from the transmission lines into medium-voltage electricity of 30 kV, (iii) feeder lines midstream, which connect substations upstream to distribution stations downstream and have voltage levels of either 30 kV or 6.6 kV, and (iv) distribution stations downstream, which turn the medium-voltage electricity into low voltage levels of either 380 volts or 220 volts for (v) end users, who receive electricity directly from the distribution stations at the low voltage levels.<sup>13</sup>

---

<sup>11</sup>A kilowatt-hour (kWh) is a unit of energy commonly used to measure electricity consumption. It represents the amount of energy consumed by a 1,000-watt appliance running for one hour. Thus, it combines power (in kilowatts) and time (in hours) to quantify total energy consumption. 1 megawatt-hour (MWh) = 1,000 kWh, and 1 gigawatt-hour (GWh) = 1,000,000 kWh.

<sup>12</sup>See Appendix [Section A.3.2](#) for more details on the different tariff schedules and the different customer groups present in the market, as well as their evolution over time.

<sup>13</sup>A volt is a measure of how much energy can be created by the electric charges moving through an electrical circuit.



Notes: This figure shows a stylized map of the electricity distribution network in Senegal. Each line color corresponds to a feeder line. “Station” refers to a distribution station or transformer station.

Figure 1 shows an example of the structure of the electricity distribution network in Senegal. Electricity is generated upstream at power plants and is transported to major distribution stations (e.g., Source A and Source B in the map) across high-voltage transmission lines. From these source stations, electricity is transmitted to distribution stations (or transformer stations or simply stations) through feeder lines. Multiple distribution stations thus share the same feeder lines. Distribution stations directly serve firms and households connected to the grid of the utility company. I have data on 10,789 distribution stations, with each station serving approximately 200 customers, on average. Figure 8 shows the actual geocoded map of the electricity distribution network in Senegal as of March 2022.

## 2.2 Outages

Outages are characterized by two factors: (i) whether they are local or systemwide and (ii) whether they are planned or unplanned. Outages can be driven by either local conditions (such as the conditions of local line infrastructure) or nonlocal conditions (such as major incidents at source stations). In this context, local outages are those confined to a few stations only, while system-wide outages are those caused by nonlocal conditions and are shared by stations across feeder lines.

Planned outages occur due to scheduled maintenance work, load shedding, and tests

conducted on power lines to ensure system reliability and safety. They are typically announced in advance, allowing consumers to potentially prepare for the temporary loss of power. On the other hand, unplanned outages arise from unexpected events such as line incidents, weather-related disruptions, and unknown breakdowns. These unforeseen interruptions can cause significant inconvenience and are often more challenging to manage and rectify. Planned outages often occur due to damage caused by unplanned events; thus, the two are not unrelated. Table 6 in the appendix shows the different causes of outages in the data. The four most common causes are maintenance work on the network, incidents in the network, default or damage to line equipment, and production incidents.

Throughout the study period, customers experience 20 hours of outages per month (20 occurrences of 1-hour-long outages), on average. The source of outages can be multiple. Table 7 gives the official system average interruption frequency index (SAIFI) and the system average interruption duration index (SAIDI) from 2018 to 2022.<sup>14</sup> In general, I note a U-shape trend in the frequency and duration of outages during this time period: outages decreased between 2018 and 2020 and increased between 2020 and 2022. The nonmonotonic trend in outages suggests the continuous challenges that the utility company still faces to deliver reliable electricity. Figure 9 reports the SAIDI and SAIFI values for 145 countries in 2019 and shows that Senegal is around the middle of the distribution for African countries for which data are available.

## 2.3 Reliability Projects

One of SENELEC’s goal over the last few years has been to strengthen the transmission and distribution network through the *Projet d’Appui au Secteur de l’Electricité (PASE)* to reduce technical losses. To achieve this objective, SENELEC has focused on deploying many preventive maintenance technologies across the distribution network: infrared thermography, ultrasound and cable diagnostics to anticipate incidents, or remote monitoring and remote control equipment to reduce the duration of interruptions.

This paper focuses on the installation of RTUs, i.e., remote control switches on distribution lines. Remote control switches are devices used in electricity distribution networks to remotely open or close electrical circuits. They are used to control the flow of electricity, isolate faults, and manage the network more efficiently. Remote control switches may affect the network through remote operations, automation and improved response time.<sup>15</sup>

---

<sup>14</sup>The SAIDI is the total minutes of noninstantaneous electric interruptions per year that customers experience, on average. The SAIFI is the number of noninstantaneous electric interruptions per year that customers experience, on average. By definition, they exclude micro-outages, which are accounted for in the analysis data described in Section 3. Hence, the outage metrics are significantly different between the two reporting methods.

<sup>15</sup>Remote control switches allow operators to quickly isolate faulty sections of the network and to

The remote control switches come with smart meters, which are equipped with advanced monitoring and control systems that track voltage, current, and other electricity-related parameters in near real time. Smart stations may affect the network through real-time monitoring, predictive maintenance, and optimized operations.<sup>16</sup> With the installation of remote control switches and smart stations, the objective of the utility company is to reduce the pass-through of outages across the network and to turn systemwide outages into local outages.

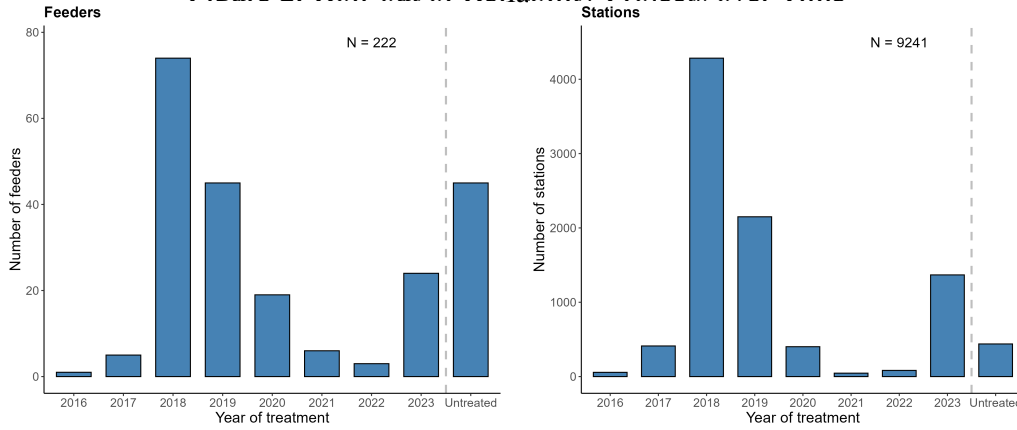
Figure 2 shows the roll-out of the reliability projects over time in terms of both the number of feeder lines with at least one treated station and the number of stations on feeder lines with at least one treated station. The figure shows that the reliability projects started in 2015 and are still ongoing. Few stations were treated in 2015, and by the end of 2023, most stations in the network were on feeder lines with a treated station. Figure 10 shows the spatial variation in the timing of the reliability projects between 2015 and 2023. In the empirical strategy in Section 4, I leverage the variation in the roll-out of these reliability projects to estimate the impact of outages on electricity consumption and on various firm-level and household-level economic outcomes. In 4, I also discuss the level at which I define the treatment status depending on the outcome variable.

---

restore power to unaffected areas remotely, reducing the time needed to locate and manually operate switches. Remote control switches can also automatically detect faults and take corrective actions, such as rerouting power or switching to alternate sources, without human intervention by integrating with monitoring and control systems. The switches can also reduce the impact of outages on customers by responding to network disturbances faster than manual switches.

<sup>16</sup>Smart meters provide operators with detailed information about the status of the network, allowing them to identify and address issues promptly. Smart stations can also reduce downtime and outages by predicting potential failures and schedule maintenance proactively by continuously monitoring equipment health and performance. The stations can also reduce the likelihood of outages and improve overall grid performance by optimizing power flow and voltage levels to improve efficiency and reliability.

Figure 2: Roll-out of Reliability Projects over Time



Notes: This figure shows the number of feeder lines that received reliability projects (left panel) and the number of stations on feeder lines with at least one reliability project (right panel) by year from 2016 to 2023. The year of treatment (i.e., the year of the reliability project) is sourced from annual digitized maps of the electricity grid. Stations created after the installation of the reliability projects are excluded from the analysis. “Untreated” corresponds to feeder lines with no reliability projects and to stations on feeder lines with no reliability projects.

Three criteria were considered by SENELEC to determine the timing of the installation of devices on distribution stations. First, the dispatch time for technicians was considered, with remote areas being prioritized over urban areas due to longer response times affecting downtime and customer convenience. Second, stations with a higher number of clients, typically in urban areas, were prioritized to reduce the broader impact of outages, highlighting a trade-off between remote, lower-density areas and high-demand urban centers. Third, the utility ranked stations based on their service importance. Urgent stations (e.g., hospitals and public offices) received top priority, followed by strategic (e.g., airports, communication centers), critical (e.g., cultural venues), important (e.g., schools, markets), and, finally, standard stations. Finally, stations with three or more sublines were prioritized, as they serve a larger customer base than those with fewer lines.

Table 8 shows summary statistics related to the different characteristics of households served by distribution stations covered by the reliability projects between 2015 and 2018 relative to those served by distribution stations that benefited from these projects between 2019 and 2023. In general, the two groups appear relatively similar across multiple dimensions, despite some differences in a few key variables. As highlighted above, the relative balance of the sample stems from the criteria used by the utility and the prioritization of both urban and rural areas. Table 9 also shows the correlates of the timing of the reliability projects in a regression.

### 3 Data

I use the following datasets for the empirical analyses: (i) digitized yearly maps of the distribution network from 2015 to 2023, (ii) the universe of station-level outages from 2016 to 2023 for all stations in Senegal, (iii) the universe of disaggregated post-paid billing and prepaid transactions data for SENELEC clients from 2012 to 2024, and (iv) a geocoded panel of household surveys.<sup>17</sup> Below, I summarize the key variables in each of these datasets. I provide more details on the collection and the matching of these different datasets in Appendix [Section A.3.1](#).

**Digitized Electricity Distribution Network.** I rely on a series of detailed maps of the entire electricity distribution network from 2015 to 2023. Each map represent a snapshot of the network in a given year. The maps show all components of the distribution network (source stations, feeder lines, poles, transformer stations, devices on lines, and distribution stations) and the links between the different components of the network (e.g., underground vs. normal lines). Importantly for the analysis, the maps show RTUs and smart devices installed on stations in each year, allowing me to construct the event study dummy variable for when a station is treated with these devices. In addition, the maps show the order in which the distribution stations are placed in a given feeder line, which allows me to obtain the relative upstream/downstream position of a station on a line. The maps also allow me to see the dynamic changes in the network over time in terms of the feeder lines to which the distribution stations belong, new distribution stations added to the network, and reorganizations of parts of the network. I use these maps to construct a time series of the implementation of the reliability projects across distribution stations over time.

**Universe of Outages.** I use the data on the universe of outages, which contain all official records of outage incidents from 2016 to 2023 made by the utility. Each row of the data is an outage incident. For each outage incident, the data contain information on the source station, the feeder station, the nature of the outage (planned or unplanned), the reason for the outage (e.g., weather-related accident, maintenance work, low supply available, voltage issues, unknown reasons), the devices damaged by the outage, the start and end times of the outage, the amount of electricity lost during the outage, the transformers that were hit by the outage, and the agent who recorded the incident. In total, there were 219,414 incidents between January 2016 and December 2023. Note that one incident often affects multiple stations or feeders. From this raw dataset, I construct a dataset aggregated at the station level to the month of the sample, which includes the

---

<sup>17</sup>All these datasets were obtained from a partnership agreement with SENELEC and Agence Nationale de la Statistique et de la Démographie (ANSD). I thank Marco Gonzalez-Navarro for helping to set up the partnership agreement with SENELEC.

number of planned and unplanned outages. For stations with no outages in a month, these variables are equal to 0. Table 6 gives the causes of outages in the context and their relative importance in terms of the number of occurrences and duration. Appendix Section A.3.1 discusses additional steps that I take to ensure that the data incorporate outage types often missed in feeder-level outages used in the literature.<sup>18</sup> In particular, I use high-frequency smart meter data with readings at an interval of 10 minutes of voltage levels, current, and different power flows for 2,644 smart stations from January 2021 to May 2023. These 10-minute high-frequency data on smart stations allow me to construct a panel of outages that is similar to the panel used in the analysis. These smart meter data have the advantage of being recorded directly from the stations (as opposed to being manually recorded by operators of the distribution network), relying on very granular data, and being less subject to potential misreporting and manipulation. Thus, this high-frequency dataset is mostly used to construct a time series of voltage fluctuations.

**Universe of Bills.** I use the universe of bills for all customers in Senegal from 2012 to 2024. The data have the following information on customers: their full name, address, phone number, subscription date, and electricity use for a post-paid bill or electricity purchased for prepaid purchases, as well as the customer size (small vs. medium vs. large), customer type (residential vs. firm), and tariff structure (block pricing vs. peak load pricing). In total, I have approximately 100 million customer-month pairs for post-paid customers and 25 million customer-month pairs for prepaid customers between 2012 and 2024 across Senegal, for final electricity consumption data amounting to 125,573,214 customer-month pairs. Please see Appendix Section A.3.1 for details on how I construct monthly and yearly electricity consumption for post-paid and prepaid clients.

**Representative Survey of Households.** I use longitudinal household expenditure surveys conducted in Senegal in 2018/2019 and in 2021/2022, the Enquête Harmonisée sur les Conditions de Vie des Ménages 2018–2019 (EHCVM) and the Enquête Harmonisée sur les Conditions de Vie des Ménages 2021–2022, to incorporate in the analysis socioeconomic variables beyond those related to electricity that the datasets above provide. These surveys are very similar in spirit to the Living Standard Measurement Survey (LSMS) commonly conducted in several low-income countries. The data were collected by Agence Nationale de la Statistique et de la Démographie (ANSD) through

---

<sup>18</sup>I discuss in detail the outages that are covered by the data. I also show evidence that the data pick up micro-outages that are often missed by feeder-level outage data typically used in the literature. In particular, I show strong correlations between outages reported by surveyed households and the data. I also show strong evidence between smart meter readings of outages and voltage fluctuations and the 2016–2023 data.

a 2-stage sampling methodology. A total of 598 enumeration areas (EAs) were selected in the first stage, and 12 households were randomly selected in each EA in the second stage. The total survey sample size is 7,156 households, with 3,941 from urban areas and 3,215 from rural areas. The EHVCVM is a rich dataset covering education, health, employment, nonemployment income, saving and credit, food consumption, food security, nonfood consumption, nonagricultural enterprises, housing, household assets, transfers, shocks and survival strategies, safety nets, agriculture, livestock, fishing, agricultural equipment, and relative poverty. Importantly, I have the exact centroids of the EAs where respondents live. I use the centroids to match the EHCVM with the rest of the datasets.

**Client-Level Characteristics.** Finally, I also rely on a client-level dataset that provides information on the socioeconomic characteristics of the customers of the utility company, their payment and theft behavior, their geographical reference, and technical information on their meters. The socioeconomic characteristics include their age, gender, occupation, type of commercial activity, nationality, address, phone number, and establishment type (house, apartment, store, office, factory, etc.), as well as the size category of the client (small vs. medium vs. large), whether the client is residential or commercial, whether the client is private or public, and whether the client is a special client. In addition to these socioeconomic variables, the dataset includes a detailed census of baseline appliances. For each client, I have the number of light bulbs, refrigerators, TV, freezers, heaters, fans, air conditioners, Hi-Fi, irons, washing machines, generators, and laptops, as well as the required power in watts associated with each of these appliances. The payment and theft behavior information in the dataset includes the number of disconnections, the number of moratoria, and the number of fraud incidents since subscription. It also provides the types of taxes that the client is subject to (e.g., stamps, VAT, TV tax, municipality tax) as well as the fixed maintenance fee paid for each billing cycle and the payment method (cash, bank withdraw, check, etc.). The dataset also includes a 10-digit geographical reference number that allows me to know the relative location of the client without necessarily having geocoded addresses. I mostly use the information in this last piece of data as a covariate in the analysis.

**Mapping the Datasets.** For my empirical analysis, I need to perform three matches: (i) matching network maps to outage incidents to link reliability projects to outage incidents, (ii) matching network maps to household surveys to link reliability projects to household outcomes, and (iii) matching electricity bills to network maps to link reliability projects to monthly electricity consumption. For (i), the matching is performed through an exact matching of station names in the outage incident datasets and station names in the network maps. For (ii), I rely on the geocoded information that I have for the network maps and household surveys. I match each EA in the household survey to

the closest station in the network map. To check whether this distance-based matching is accurate, I check the outage duration and frequency as recalled by survey respondents against the official outage data received by the utility company for the matched station. This correlation is shown in Figure 26 in the appendix. Similarly, for (iii), I first geocode the literal addresses that I have for the customers in the electricity billing data. I match each geocoded customer in the billing data to the closest station in the network map. Figure 11 shows the overlay of the EAs selected in the household surveys, along with the distribution stations in the network map and the geocoded customers from the billing data. Appendix Section A.3.1 provides details on how these matching exercises are implemented.

In Appendix Section A.3.3, I present a few stylized facts from the data to help motivate the reduced-form and model analyses that follow. These facts are as follows: (i) there is significant spatiotemporal variation in outages; (ii) most station-level outages occur during the day; (iii) unplanned outages are shorter in duration but are more frequent than planned outages; and (iv) consumers have very few appliances in general but upgrade their appliances over time.

## 4 Impacts of Reliability Projects on Consumers

In this section, I present my empirical strategy for evaluating the impact of changes in electricity reliability on a range of outcomes at both the station and client levels. The outcome of interest is the duration and frequency of power outages, in terms of both the direct effects on treated stations and the potential spillover effects of the treatment on untreated stations. At the client level, I examine a range of downstream economic outcomes, including electricity expenditures, the location choice of firms and households, bill payment behavior, and time use. Additionally, I analyze household-level appliance ownership and the time use of household members.

### 4.1 Empirical Strategy

Given the complexity of electricity networks and the multiple channels through which reliability projects may affect both treated and untreated locations, I address two primary challenges that arise when assessing the impact of reliability changes. The first challenge is the issue of simultaneity between changes in economic activity and reliability, making it difficult to determine the causal direction. For instance, areas with higher economic growth may demand and receive improvements in electricity reliability, which could bias the estimated effects if not accounted for appropriately. The second challenge involves potential spillovers—instances where improvements in reliability in treated locations may affect untreated stations, either directly through the network or indirectly through changes in regional economic conditions. If not properly controlled for, these

spillovers could confound the estimation of the true effects of reliability projects on both treated and untreated stations. These spillovers are particularly relevant in interconnected electricity networks where interventions in one area may lead to changes in reliability in other parts of the grid.

To overcome these challenges, I structure my empirical analysis as follows. I estimate the direct effects of reliability projects on outages at treated stations by comparing early-treated stations with late-treated stations and stations located on untreated feeders. Doing so allows me to isolate the direct effects of the reliability projects. Then, I estimate the potential spillover effects of reliability projects by comparing untreated stations based on their upstream or downstream locations relative to treated stations. Finally, I estimate the impact of reliability projects on the economic outcomes of clients served by treated feeders by comparing these clients to those served by late-treated or untreated feeders within the same district.

First, to evaluate the direct effects of reliability interventions on outages at treated stations, I estimate the following model for station  $s$  in feeder  $f$  at time  $t$ :

$$y_{sft} = \alpha_{sf} + \delta_t + \sum_{k \neq -1} \beta^k \cdot 1(t - T_{sf} = k) + \Gamma X_{ft} + \epsilon_{sft} \quad (1)$$

In regression (1),  $y_{sft}$  represents the total hours of outages experienced by station  $s$  in feeder  $f$  at time  $t$ . In additional specifications, we let  $y_{sft}$  represent the total number of outages or the duration/frequency of planned and unplanned outages separately.  $\alpha_{sf}$  captures station fixed effects and controls for time-invariant characteristics of stations.  $\delta_t$  represents time fixed effects, capturing any common time trends that affect all stations simultaneously, such as macroeconomic shocks or seasonal effects.  $T_{sf}$  denotes the year in which station  $s$  received its reliability upgrade. Here,  $t$  represents a month, and the times to project date  $l = t - T_{sf}$  included in the regression are  $\{-60, 48\}$ . The vector  $X_{ft}$  includes time-varying controls such as the number of stations on feeder  $f$  and the number of feeder lines originating from the same source station. These controls help account for general changes in the network structure and for other interventions that could affect the outcome variables of interest. Standard errors are clustered at the feeder level to account for potential correlation in outcomes across stations within the same feeder. The identifying assumption in regression (1) is that, in the absence of reliability projects, the outage trends at early-treated stations would have followed a trajectory similar to those at late-treated stations. The coefficients of interest are  $\beta^k$ , the leads and lags of the treatment effect, which capture the pretreatment and post-treatment dynamic changes in outages due to reliability projects. Figure 12 provides a visual intuition of the variation exploited to estimate  $\sum_{k \neq -1} \beta^k$  in regression (1).

Second, I examine whether reliability improvements at treated stations generate spillover effects on untreated stations located on the same feeder or located on other feeders. Specifically, I employ a DDD approach to compare untreated stations lo-

cated upstream and downstream of early-treated stations to untreated stations located upstream and downstream of late-treated feeders. The DDD specification for untreated station  $s$  on feeder  $f$  is as follows:

$$y_{sft} = \alpha_{sf} + \delta_t + \beta_1(\mathbf{P}_t \cdot \mathbf{T}_f) + \beta_2(\mathbf{P}_t \cdot \mathbf{U}_{sf}) + \beta_3(\mathbf{P}_t \cdot \mathbf{T}_f \cdot \mathbf{U}_{sf}) + \Gamma X_{sft} + \epsilon_{sft} \quad (2)$$

where  $\mathbf{T}_f$  is a dummy variable indicating whether feeder  $f$  is treated early. Unlike regression (1), I define the treatment to be at the feeder level.  $\mathbf{U}_{sf}$  is a dummy variable for whether station  $s$  is located upstream of a treated station.<sup>19</sup>  $\mathbf{P}_t$  is a dummy variable for whether the year is post-treatment. The coefficient  $\beta_1$  of the interaction term  $\mathbf{P}_t \cdot \mathbf{T}_f$  captures the effect of the reliability projects on outages of untreated stations that are downstream of treated stations on feeders treated early relative to outages of untreated stations that are downstream of treated stations in feeders treated late. The coefficient  $\beta_2$  of the interaction term  $\mathbf{P}_t \cdot \mathbf{U}_{sf}$  captures the difference in the post-treatment period between outages of untreated stations that are upstream of treated stations and outages of untreated stations that are downstream of treated stations. The coefficient  $\beta_3$  of the triple interaction term  $\mathbf{P}_t \cdot \mathbf{T}_f \cdot \mathbf{U}_{sf}$  captures the additional effect of the reliability projects on outages of untreated stations that are upstream of treated stations in an early-treated feeder.  $\alpha_{il}$  and  $\mu_t$  are customer fixed effects and year fixed effects, respectively. Standard errors are clustered at the feeder level. Figure 13 provides a visual intuition of the variation exploited to estimate  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  in regression (2).

In theory, upstream and downstream stations could be differently affected by reliability projects. In general, if an incident occurs on a feeder line with a distribution station that can be controlled remotely, then the utility company can turn the station on or off to quickly detect and address the incident remotely and to prevent the problem from spreading across the feeder lines to other distribution stations. When an issue occurs at a given station, the utility company determines the most downstream station that it can close and still handle the issue to minimize the number of stations affected by the incident. In this case, the stations upstream of the incident will receive electricity while the issue is resolved for the downstream stations. Hence, stations upstream of incidents are more likely to be immune to interruptions, while stations downstream are not, provided that the utility company knows the locations of the incidents on the line. As noted

<sup>19</sup>As noted in Section 3, the detailed digitized maps of the network allow me to obtain the exact rank of the stations within their feeder and the sequence in which electricity is distributed across the stations of the same feeders. I construct a dummy variable that takes the value of 1 if the station comes before the closest treated station in the sequence of stations on the feeder line and 0 otherwise. For feeder lines that are relatively straight and with only one treated station, the classification of untreated stations between upstream and downstream stations is straightforward. For feeder lines that are nonlinear (e.g., feeder lines in a tree format or circular feeder lines), the upstream status of an untreated station is defined based on whether electricity flows there before a treated station. For feeder lines with more than one treated station, the upstream status of an untreated station is defined based on its position relative to the closest treated station.

in Section 2, reliability devices allow the utility company to locate incidents on feeder lines more quickly. With the reliability projects, treatment of a given station allows the utility company to minimize interruptions for stations that are upstream of the treated stations. While downstream stations are less affected by the device, they could still benefit simply because they are on the same line as treated and upstream stations. Thus, if the utility company determines that a line incident occurred on the upstream part, it can take action on this part, and all stations on the line benefit from the issue being resolved more quickly. Upstream stations just have the added advantage of losing power for a shorter period of time (owing to the smart devices) when the utility company checks where on the feeder line the incident occurred.

Third, I examine the impact of reliability projects on the downstream economic outcomes of clients served by the treated stations. To estimate these effects, I use the following specification for client  $i$  living in district  $d$  and served by station  $s$  on feeder  $f$ :

$$y_{idsft} = \alpha_{sf} + \delta_{dt} + \sum_{k \neq -1} \beta^k \cdot 1(t - T_f = k) + \Gamma X_{sft} + \epsilon_{idft} \quad (3)$$

where  $y_{idsft}$  represents the outcome variables of interest for client  $i$  in district  $d$  served by station  $s$  on feeder  $f$  in year  $t$ .  $\alpha_{sf}$  captures station fixed effects, and  $\delta_{dt}$  represents district-time fixed effects, controlling for unobserved heterogeneity at the district level.  $T_f$  denotes the year in which feeder  $f$  received its first treatment. Unlike regression (1), I define the treatment status to be at the feeder level (i.e., a station is treated if it is on a feeder with at least one treated station), having established the presence of spillover effects across distribution stations on the same feeder.<sup>20</sup> The years to project date  $l = t - T_f$  included in the regression are  $\{-6, 4\}$ . The vector  $X_{sft}$  includes controls for station and feeder characteristics, and standard errors are clustered at the feeder level. The identifying assumption for this model is that, in the absence of reliability interventions, the economic outcomes of clients within the same district but served by stations that belong to different feeders would have followed similar trends. The variation that I exploit is the difference in reliability change for stations that are near each other but that belong to different feeders with different treatment statuses. Standard errors are clustered at the station level to account for cross-sectional correlation across customers served by the same stations. The coefficients of interest are  $\beta^k$ . Figure 14 provides a visual intuition of the variation exploited to estimate  $\beta^k$  in regression (3).

I now turn to the assumptions needed to causally estimate our parameters of interest in regressions (1), (2), and (3). Let  $Y_{i,t}(g)$  represent the potential outcomes of client  $i$  in period  $t$  if it received treatment in period  $g$ ,  $G_i$  be the treatment group of client  $i$ , and  $Y_{i,t}(\infty)$  be the potential outcomes in period  $t$  if it never received treatment. In addition

<sup>20</sup>See the discussion of spillover results in Appendix Section A.4.

to standard regularity conditions, the assumptions required for  $\beta^k$  to yield unbiased estimates are as follows:

1. **Conditional parallel trends:**  $\mathbb{E}[Y_{i,t}(\infty) - Y_{i,t'}(\infty)|G_i = g, X_i] = \mathbb{E}[Y_{i,t}(\infty) - Y_{i,t'}(\infty)|G_i = g', x_i] \quad \forall t \neq t' \text{ and } g \neq g'$
2. **No anticipation:**  $Y_{i,t}(g) = Y_{i,t}(\infty) \quad \forall i \text{ and } t < g$

Assumption 1 states that the outcomes for all treatment groups would have evolved in parallel under a counterfactual world with no reliability project, conditional on observed prespecified covariates. Assumption 1 thus generalizes the standard 2x2 parallel trends assumption under uniform treatment timing to all possible combinations of  $t$  and  $g$ . Assumption 2 imposes that the outcome for a not-yet-treated distribution station for time period  $t$  prior to its treatment year does not depend on the year that it receives reliability projects in the future.

In practice, the average treatment effect in period  $t$  of the reliability projects received at time  $g$  that I estimate is as follows:

$$ATT(g, t) = \mathbb{E}[Y_{i,t}(g) - Y_{i,t-1}(g)|G_i = g] - \mathbb{E}[Y_{i,t}(g') - Y_{i,t-1}(g')|G_i = g'] \quad \forall g' > t$$

## 4.2 Results

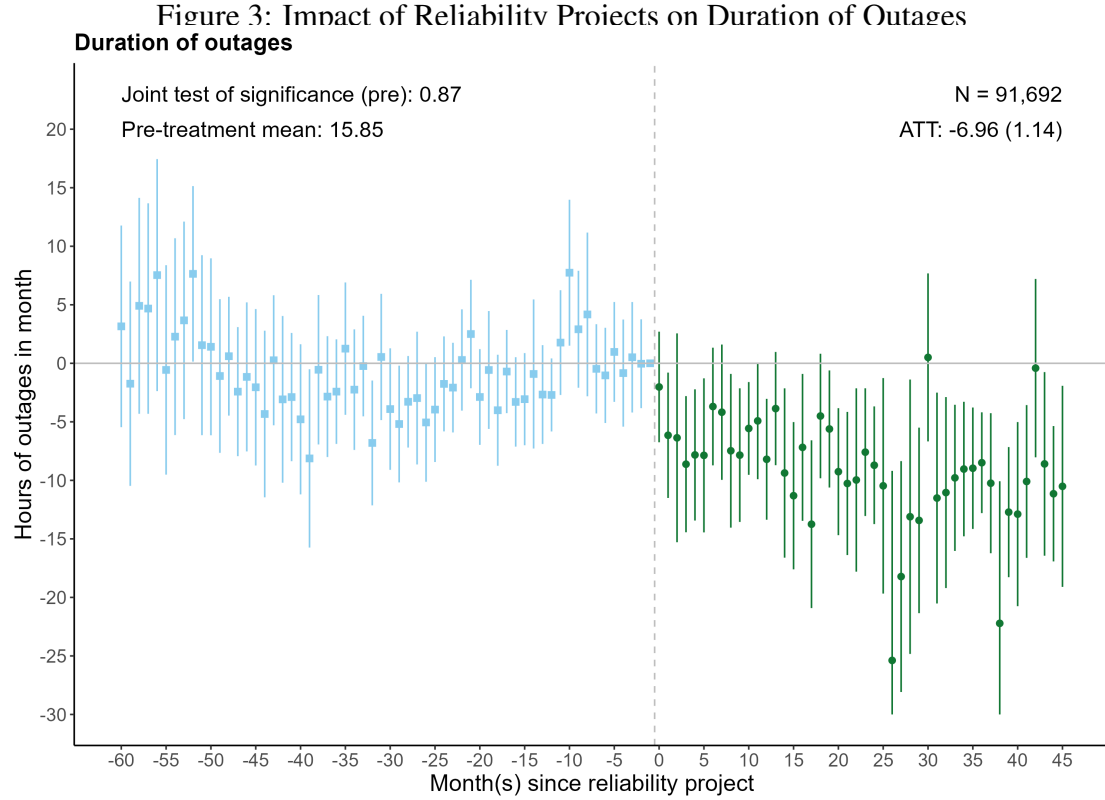
In this subsection, I present the results of the empirical analysis on the impact of reliability projects on outages, electricity use, and firm and household outcomes. <sup>21</sup>

### Outages and Electricity Use

Figure 3 shows the event study coefficients corresponding to regression (1), with monthly total hours of outages at the level of the distribution station. Reassuringly, I do not see any significant pretrends before project completion between stations that were selected earlier and stations that were selected later for reliability projects. In addition, the post-event coefficients in green suggest a significant decrease in the duration of outages. The average treatment effect of reliability projects on the treated is a reduction in the duration of outages of 6.96 hours (standard error = 1.14) from a baseline total hours of outages of 15.85 hours, which corresponds to a 44% reduction in outages at stations treated early relative to stations treated late and to untreated stations.

---

<sup>21</sup>All regressions are estimated via the `did` package developed by Callaway and Sant'Anna (2021).



Notes: This figure shows the coefficients  $\lambda_t$  in regression (1) and their corresponding confidence intervals, estimated following Callaway and Sant'Anna (2021). All coefficients are estimates relative to a baseline value for the month preceding the project date. The mean monthly number of outages in the baseline period is 15.85. The y-axis shows the frequency of outages in terms of the number of outages in the month. The x-axis shows the number of months since the completion of the reliability project. The vertical dashed gray line corresponds to the month of the project date. The data used for estimation are a balanced sample of treated stations with outage observations for all time periods of the analysis, from 5 years before the project date to 4 years after the project date. The estimation is performed at the distribution station level. "ATT" refers to the overall average treatment effect on the treated stations, with the associated standard errors in brackets. "N" corresponds to the number of station-month pairs used as observations in the regression.

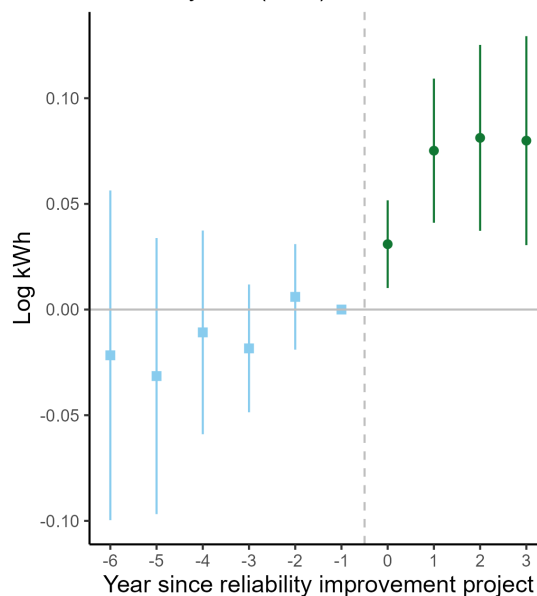
Figure 15 in the appendix shows the event study coefficients corresponding to regression (1), with the total number of outages as the outcome variable. The coefficients of the post-project completion time periods for this figure show a significant decrease after the completion of reliability projects, but this decrease is not as strong as the decrease in the duration of outages in Figure 3. This finding aligns with the discussion in Section 2 regarding the theory of change underlying the reliability projects and the margin of reliability that they are expected to affect. Figure 16 shows the equivalent of Figures 3 and 15 with data aggregated at the year level.

Figures 37 and 38 in the appendix show the results of regression (1) separately for planned outages and unplanned outages. These figures show a decrease in the duration and frequency of planned outages but no major change in the duration and frequency of unplanned outages. Figure 39 shows the results of regression (1) with several other re-

gressors used in the literature. Figure 41 includes control variables in the specifications.

Figure 4 shows the results of estimating regression (3) when the log of electricity consumption by clients (including both households and firms) is used as the outcome variable. Since I use billing data for this purpose, I have a longer time horizon and can expand the event study results to 6 years before treatment. In line with the previous event study graph, the data do not reveal any pretrends in electricity use before the project date. After a reliability project, the data exhibit an increase in electricity use by a factor of 1.02–1.15, which corresponds to an average increase of 8%. Figure 17 shows the results of estimating regression (3) when the log of nominal electricity expenditures is used as the outcome variable instead. The results for expenditures are similar to those for electricity use, as expected.

Figure 4: Impact of Reliability Projects on Electricity Use  
Electricity use (kWh)



Notes: This figure shows the effects of reliability projects on electricity use, with the coefficients  $\lambda_t$  in regression (3) and their corresponding confidence intervals. All coefficients are estimates relative to a baseline value for the year preceding the project date. The y-axis shows electricity use in terms of the mean log electricity use (kWh) in the year. The x-axis shows the number of years that have passed since the completion of the reliability project. The vertical dashed gray line corresponds to the year of the project date. The data used for estimation are geocoded billing data, where each observation is a customer of the utility company. The estimation is performed at the client level.

Figures 3 and 4 jointly show that a 44% reduction in total hours of outages leads to a 8% increase in total electricity expenditures, which corresponds to an elasticity of consumption with respect to electricity reliability of 0.18 or a reduction in consumption of 2.6% for each hour of outages. Despite their significance, the estimates are lower in magnitude than the few estimates in the literature. For example, for households in India,

Khanna and Rowe (2024) estimate that an additional hour per month of power outages reduces electricity consumption by 4.8%.

### **Mechanical vs. Demand Effect in Electricity Use**

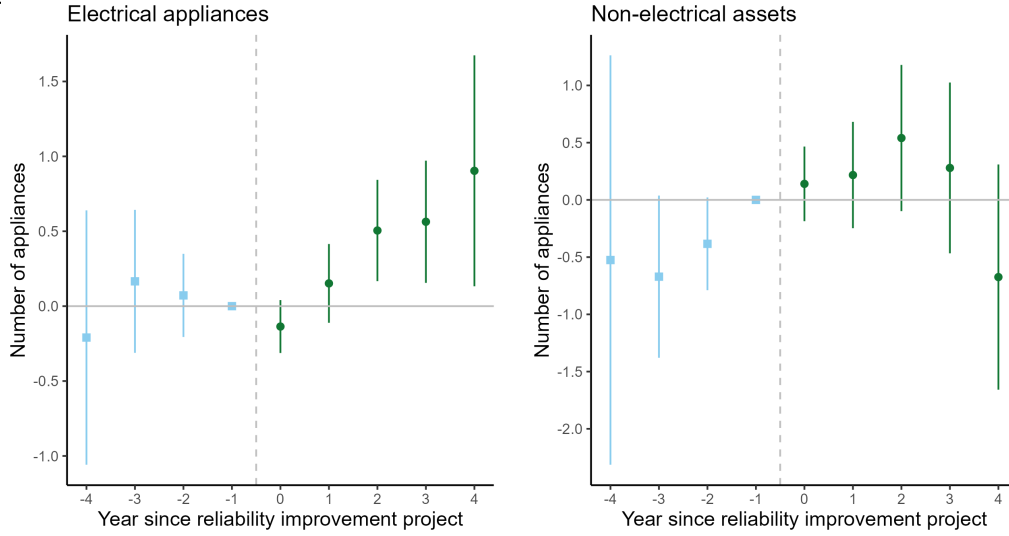
The change in electricity use as a result of changes in electricity reliability could be driven by three different channels: (i) a mechanical effect, that is, a passive increase in electricity use coming solely from having electricity for longer hours with no change in consumer behavior, (ii) a higher utilization of the same appliances, or (iii) a demand effect through an upgrade in appliances. Channel (i) is relevant for appliances that are typically on when electricity is on (e.g., refrigerators and light bulbs). For these appliances, simply having more hours of electricity available could turn into higher usage and higher electricity bills without any action by the consumers. This channel rests on the assumption that consumers have inelastic demand that does not vary much over time or that the marginal utility from these appliances for the additional hours largely outweighs the marginal price of electricity provision. Channel (ii) is relevant insofar as a reduction in power outages (and, subsequently, a reduction in voltage fluctuations) increases the efficiency and usability of existing appliances, leading consumers to want to use them more. For example, consumers could be more inclined to use their freezer and may start putting more perishable goods in it if power is supplied more reliably. This channel is different from channel (i) since it hinges on consumers actively making the decision to consume more. Finally, channel (iii) captures the long-term adaptation of consumers once they realize that grid supply has been more reliable for a long enough period of time that it could be worthwhile for them to invest in more appliances.<sup>22</sup>

Below, we show strong evidence that the demand effect drives most of the change in electricity use and that the role of the mechanical effect is relatively minimal.

---

<sup>22</sup>This channel rests on the assumption that consumers are aware of and have internalized the reliability changes. Figure 18 provides a visual intuition for the change in demand coming solely from channel (i). Figure 19 provides a visual intuition for the change in demand when both channel (i) and channel (ii) apply.

Figure 5: Impact of reliability projects on Electric Appliances and Nonelectric Assets



Notes: This figure shows the coefficients  $\lambda_t$  in regression (3) and their corresponding confidence intervals. All coefficients are estimates relative to a baseline value for the year preceding the project date. The left-hand side graph shows the total number of electric appliances (TVs, phones, laptops, fans, etc.), while the right-hand side graph shows the total number of nonelectric assets (bicycles, cars, radios, beds, tables, sofas, etc.). In both graphs, the x-axis shows the number of years that have passed since the reliability project. The vertical dashed gray line corresponds to the year of project completion. The data used for estimation are a balanced sample of treated households with nonmissing values for all time periods of the analysis, from 6 years before the project date to 4 years after the project date. The estimation is performed at the household level.

The left-hand side of Figure 5 shows the results of estimating regression (3) with the number of electric appliances as the outcome variable. Electric appliances include phones, fans, fridges, laptops, TVs, microwaves, tablets, washing machines, air conditioners, printers, vacuums, generators, electrical irons, and freezers. Reassuringly, I do not see any significant pretrends in appliance adoption before project completion dates between customers served by stations that were selected early and customers served by stations that were not selected early for reliability projects. After project implementation, the number of electric appliances increases by 0.58 units from a baseline total appliances of 5.37, corresponding to a 9% increase in the total number of appliances. Importantly, the increase in common appliances does not come right after project completion and is significant only two years post-project completion. These results suggest that the change in appliances is more of a longer-term response to the reliability projects, which is in line with the idea that consumers change their appliance roster only when they see a sustained change in electricity reliability. The 9% increase in appliances is in line with the 8% increase in electricity use documented above, despite the two estimates being based on two different datasets. Figure 44 shows the results of regression (3) with a sample of households that have the same exposure to treatment (i.e., where the length of exposure is the same across the analysis sample).

The right-hand side of Figure 5 shows the results of estimating regression (3) with

the number of nonelectric assets as the outcome variable. Nonelectric assets includes bicycles, tables, carpets, mattresses, radios, and nonelectrical irons, among many others. Unlike the number of electric appliances, which is meaningfully affected by reliability changes, the number of nonelectric assets does not increase. This result reinforces the argument that reliability projects most directly affect electric appliances and that any income effect would likely channel through the adoption and use of these appliances. Figure 20 shows the event study coefficients 4 years after project implementation by appliance type. In line with the stylized facts presented in Appendix Section A.4 on the landscape of appliances in the study context, there does not appear to be any upgrade in more sophisticated appliances: freezers, microwaves, printers, vacuums, generators, etc.<sup>23</sup>

However, for the more prevalent appliances, including fans, refrigerators, and laptops, Figure 20 shows a clear uptick in the number of appliances post-project completion. These changes are not solely driven by an intensification of existing appliances, as the questions on the EHCVM surveys were specifically for new appliances purchased.<sup>24</sup> While these significant coefficient estimates might seem small, they need to be put in perspective based on Table 16. As noted in that table, the study context is characterized by a very low level of appliance adoption. For example, the mean number of fans in the sample is 0.68; thus, an upgrade of 0.7 fans is approximately equivalent to a 100% increase in fan ownership, on average, which is considerable.

Figure 21 shows the same estimates as those in Figure 20 separately by home ownership (renters vs. homeowners) and by income group (above vs. below median income). Homeownership is used to compare changes in appliances following reliability projects for homeowners and renters. In particular, the electric appliances that are affected by reliability changes in Figure 20 are quite portable and easily movable by renters. They are not the heavy or large appliances that renters may leave behind in case they leave a place. Figure 21 shows that homeowners and renters behave relatively similarly when reliability improves. Similarly, the right-hand side of Figure 21 shows that households whose income is above the median do not differentially change their appliances relative to households whose income is below the median.

Taken together, the results in Figures 5, 20, and 21 indicate that appliance upgrades (channel iii) are the primary way reliability improvements impact customers. These upgrades align with changes in electricity use observed in billing data, suggesting a long-term effect where customers adjust their appliances in response to more reliable electricity, rather than a short-term mechanical effect of outages on consumption.

To stress this point even further, I note that a change in total hours of outages of 7

---

<sup>23</sup>Given the very low share of the sample who own any of these appliances, I cannot rule out the possibility that the null effects for these appliances are due to the sample being underpowered for such analysis.

<sup>24</sup>As discussed at the beginning of this section, there could also be a more intense use of existing appliances that used to be off frequently. My data do not allow me to test this hypothesis.

hours from a baseline of 16 hours corresponds to a increase in total hours of electricity availability from 704 hours to 711 hours out of a total of 744 per month. Assuming that appliances that are more likely to be affected by the mechanical effect discussed in channel (i) are on most of the time during the month, this increase in electricity availability will lead to an increase in the usage of these appliances of less than 1%. This result suggests that channel (i) is unlikely to be the main driver of the change in electricity use that I document.

In Appendix [Section A.4](#), I run robustness checks and address potential SUTVA-related concerns, the endogenous selection criteria, and the presence of other interventions at the same time as the reliability projects. I also conduct formal tests of violations of parallel trends using the techniques proposed by [Rambachan and Roth \(2023\)](#) to bound potential violations of post-treatment parallel trends relative to the worst pre-treatment parallel trends violation. I also follow the recommendations from [Roth et al. \(2023\)](#) and conduct additional checks to ensure that the results are robust to different formulations of the outcome variables, different levels of fixed effects, and the use of different estimators. I also show the robustness of the results to individuals and stations with the same number of years of treatment.

In Appendix [Section A.5](#), I present results on the impact of improved reliability on household time use. I find that households reallocate their time away from household chores towards work hours. I also present results on the impact of improved reliability on the location of residential and commercial clients. I find that clients are more likely to enter areas with improved reliability and less likely to exit areas with higher reliability.

## 5 Value of Electricity Grid Reliability

While the reduced-form estimates are in themselves interesting, they are only suggestive of the value that consumers place on reliability of grid electricity. The reason is that electricity reliability is critical to many aspects of the economic activities of households and firms and its impact may be too diffuse to be fully captured by a restricted set of outcome variables. In this section, I estimate the monetary value of grid supply and the change in this value due to reliability projects.

The section is organized as follows. First, I present an economic model that links electricity reliability to consumer welfare and describes the change in optimal electricity use as reliability changes. Second, I use this economic model to derive the consumer surplus associated with electricity reliability for the full sample of residential and commercial clients present in the billing data and calibrate parameters needed for its estimation. To do this, I leverage the richness of the individual-level billing data to estimate demand elasticities separately for firms and households, both before and after the reliability projects. Third, I link the economic model to a discrete choice model of energy choice to estimate the marginal willingness to pay (MWTP) for one additional

hour of electricity available. The discrete choice model has the advantage of including both consumers already connected to the grid and consumers unconnected to the grid. Finally, I close this section with a cost–benefit analysis of the reliability projects.

## 5.1 Economic Model

Consumers derive utility from the consumption of physical nonelectric goods and from services generated by electricity through the use of appliances. These services include lighting, refrigeration, leisure, cooling, and heating. Each consumer  $i$  is served by a distribution station  $l$ . The distribution station  $l$  is subject to outages and voltage fluctuations, which are passed through uniformly to the consumers whom they serve. Thus, all consumers served by the same station face essentially the same supply of grid electricity.

Let  $t$  represent months (billing cycles). Each month  $t$  is divided into a continuum of measure 1 of micro-instants  $m$ . Let  $\iota_{m|t}$  be an indicator variable taking the value of 1 if grid electricity is available at station  $l$  at instant  $m$  of month  $t$  and 0 otherwise. Consumers are assumed to know the expected value of outages:  $E(\iota_{m|t}) = \rho_{lt}$ . Despite these signals in the distribution of outages, consumers do not know which instants in the time period electricity will be available.<sup>25</sup>

Let  $\tau_{ilt}$  be the coefficient that turns electricity use  $x_{im|t}$  into electrical services  $z_{im|t}$  through the following linear relationship:  $z_{im|t} = \tau_{ilt} \cdot x_{im|t}$ .<sup>26</sup> The function  $\tau_{ilt}$  reflects both the energy efficiency of appliances and the importance of the service generated at time  $t$ .

All the instants within the period have the same weight for the consumer. There is a perfect substitution between all instants when grid power is on and electricity use for these instants is always  $x_{i1|t}$ . Likewise, there is a perfect substitution between all instants when grid power is off and electricity use for these instants is always  $x_{i0|t} = 0$ .

Let  $\mu_i(\cdot)$  be the utility function from electricity services.  $\mu_i(\cdot)$  is a concave twice-differentiable utility function (e.g.,  $\mu'(\cdot) > 0$  and  $\mu''(\cdot) < 0$ ) and has the subscript  $i$  to account for the different levels of risk aversion across consumers with respect to the availability of electricity. The consumer's total utility from electrical services in period  $t$  equals the utility from the sum of electrical services across the instants when grid power is on during  $t$ . Total electrical services for period  $t$  across all instants in which electricity is on is therefore given by  $z_{ilt} = \tau_{ilt}\rho_{lt}x_{i1|t}$ . Assuming a quasi-linear utility function gives the following:

$$u_{ilt} = \rho_{lt}\mu_i(\tau_{ilt}\rho_{lt}x_{i1|t}) + (1 - \rho_{lt})\mu_i(0) + y_{ilt} \quad (4)$$

<sup>25</sup>For example, they may know that electricity is available 90% of the time in the year, but they will not know which month, day, or hour the 10% of the time without electricity will be realized in that year.

<sup>26</sup>Our baseline derivation assumes that  $\tau$  is exogenous. See Appendix [Section A.6](#) for a version of the proposition 1 with endogenous energy efficiency that depends on electricity reliability.

where  $x_{i1lt}$  is the electricity use during any instant when grid electricity is on,  $u_{ilt}$  is the utility of consumer  $i$  served by station  $l$  at time  $t$ , and  $y_{ilt}$  is the consumption of numéraire physical goods. Let  $w_{ilt}$  be income, and  $e_t$  be the unit price of electricity. A maximization of the utility in (4) subject to the constraint  $e_t \cdot \rho_{lt} \cdot x_{i1lt} + y_{ilt} = w_{ilt}$  gives the following first-order condition:

$$\mu'_i(\tau_{ilt}\rho_{lt}x_{i1lt})\tau_{ilt}\rho_{lt} = e_t \quad (5)$$

I now turn to disentangling the mechanical effect from the demand effect discussed and documented in Subsection 4.2.

**Change in Electricity Demand for One Instant: A Demand Effect.** Dropping the subscripts for ease of notation, a total differentiation of (5) w.r.t  $x$  and  $\rho$  expressed as elasticities yields the following:

$$\mu''\tau^2\rho^2dx + (\mu''\tau^2\rho x + \mu'\tau)d\rho = 0$$

$$\frac{dx}{d\rho} = -\frac{\mu''\tau^2\rho x + \mu'\tau}{\mu''\tau^2\rho^2}$$

$$\frac{dx}{d\rho} \frac{\rho}{x} = -1 - \frac{\mu'}{\mu''z}$$

$$\epsilon_{x,\rho} = -\left(1 + \frac{1}{\epsilon_{\mu',z}}\right)$$

where  $\epsilon_{x,\rho}$  measures the elasticity of electricity use with respect to electricity reliability.  $\epsilon_{\mu',z} = \mu''\frac{z}{\mu'}$  measures the responsiveness of the marginal utility  $\mu'(z)$  with respect to changes in  $z$ . It reflects the concavity of the marginal utility function or how quickly marginal utility decreases with consumption.

The derivations above yield the following proposition:

**Proposition 1** *For any instant  $m$ ,  $\epsilon_{x,\rho} > 0$  iff  $\epsilon_{\mu',z} > -1$ .*

Proposition 1 states that the sign of  $\epsilon_{x,\rho}$  will depend on the magnitude of  $\epsilon_{\mu',z}$ . Intuitively, the demand effect in electricity use that a consumer experiences as a result of a change in electricity reliability depends on the change in the marginal utility of electrical services. In particular, a value of  $\epsilon_{\mu',z}$  between -1 and 0 indicates that marginal utility decreases with  $z$  but not too rapidly. A marginal utility function that is not too concave means consumers still gain significant additional utility from consuming more electricity, even at higher levels of energy efficiency or electricity use. In this case, marginal utility is not too concave (i.e.,  $-1 < \epsilon_{\mu',z} < 0$ ), and an increase in  $\rho$  positively affects

electricity usage  $x$ . In contrast, if the marginal change in utility decreases, then consumers might prefer to decrease consumption despite efficiency gains from reliability. An improvement in reliability may lead to a more intense use of appliances since the marginal utility of electrical services is always positive. The positive marginal utility would induce consumers to use more electricity through a more intense use of existing appliances or through the purchase of new appliances.

Critically, if  $\mu_i(\cdot)$  exhibits constant relative risk aversion (CRRA) and is of the form  $\mu_i(x) = \frac{x^{1-\gamma}-1}{1-\gamma}$ , then  $\varepsilon_{\mu',z}$  is the negative of the coefficient of relative risk aversion  $\gamma$ , and  $-\frac{1}{\varepsilon_{\mu',z}} = \frac{1}{\gamma}$  is the intertemporal elasticity of substitution (IES). Having IES as a sufficient statistic for the effect of reliability changes on electricity usage reflects the importance of consumers' ability to adjust the timing of appliance usage consumption. This ability to adjust usage timing (flexibility) allows consumers to smooth consumption when reliability improves, especially for flexible appliances. For inflexible and essential appliances, improved reliability directly translates to increased usage because consumers can finally operate these appliances effectively.

For essential appliances that require continuous power—such as refrigerators—improved reliability allows these consumers to utilize these appliances more effectively. As reliability improves, they are more inclined to increase their consumption per instant  $x$  because the utility gained from consuming more is substantial. Therefore, for consumers with  $\frac{1}{\gamma} > 1$ , an increase in reliability leads to an increase in consumption per instant, resulting in a positive elasticity  $\varepsilon_{x,\rho} > 0$ . For these essential appliances, as long as the IES is higher than 1, then usage will increase when reliability increases.

For flexible appliances—such as mobile phone chargers, lighting, or entertainment devices—consumers can easily adjust the timing of their usage. When reliability improves, these consumers can spread their consumption over more periods rather than consuming a large amount in a short time due to uncertain availability. Because the marginal utility of additional consumption per instant is low, they may prefer to reduce their consumption per instant  $x$  and benefit from the ability to use electricity more reliably over time. Therefore, for consumers with  $\gamma > 1$ , an increase in reliability leads to a decrease in consumption per instant, resulting in a negative elasticity  $\varepsilon_{x,\rho} < 0$ .

**Change in Electricity Demand for One Period: Mechanical Effect vs. Demand Effect.** By construction, electricity use at time  $t$  is the sum of the electricity use for all instants when electricity is available during time  $t$ . Since all instants  $m$  with power are symmetric,  $x_{ilt} = \rho_{lt} \cdot x_{i1lt}$ . It follows from this equality that the change in the aggregate consumption over all instants of period  $t$  will be as follows:  $\frac{\partial x_{ilt}}{\partial \rho_{lt}} = x_{i1lt} + \rho_{lt} \cdot \frac{\partial x_{i1lt}}{\partial \rho_{lt}}$ . The aggregate change in electricity use is composed of two parts: a mechanical direct effect and a behavioral change (or indirect effect). The mechanical direct effect corresponds to  $x_{i1lt}$ , the change in aggregate use because there are now more instants with electricity. By construction, this mechanical effect is always positive: when there is more power,

consumers have more moments to use electricity. The indirect effect corresponds to  $\rho_{lt} \cdot \frac{\partial x_{i1lt}}{\partial \rho_{lt}}$ , the change in aggregate use due to the change in the level of electricity use when electricity is on (conditional on electricity being on). This indirect effect is exactly the change characterized in proposition 1. Thus, the total change in electricity use for time  $t$  depends on both the mechanical direct effect and the indirect effect. In particular, it will be negative only if the indirect effect is negative and its magnitude outweigh the positive direct effect.

## 5.2 Consumer Surplus

In this subsection, I derive the consumer surplus associated with electricity reliability for the full sample of clients present in the billing data. I leverage the richness of the individual-level billing data to estimate consumer surplus separately for firms and households and separately for the short run and the medium run. This approach also has the advantage of allowing me to disentangle the mechanical effects from the demand effects.

This subsection proceeds as follows. First, I derive the formula for consumer surplus associated with electricity consumption. Second, I calibrate the parameters needed for this estimation by estimating demand elasticities and leveraging additional details of the Senegalese electricity market. Finally, I use the results from my calibration to estimate consumer surplus for different groups before and after the reliability projects.

### 5.2.1 From Economic Model to Consumer Surplus

I start with the economic model presented in subsection 5.1 above.

$$u_{ilt} = y_{ilt} + \alpha_{ilt} \cdot \frac{z_{ilt}^{1-\gamma_{ilt}} - 1}{1 - \gamma_{ilt}} \quad (6)$$

where  $y_{ilt}$ ,  $z_{ilt}$  and  $\gamma_{ilt}$  are as defined previously. I make three changes to the general quasi-linear utility in equation (4). First, I express utility in terms of electrical services, defined as  $z_{i1lt} = \tau_{ilt}x_{i1lt}$ . With this substitution, the utility function simplifies to  $u_{ilt} = \rho_{lt}\mu_i(z_{i1lt}) + y_{ilt}$ . Second, I assume that  $\mu_i$  exhibits constant relative risk aversion (CRRA) with parameter  $\gamma_{ilt}$ , which modifies the utility function to  $u_{ilt} = y_{ilt} + \rho_{lt} \cdot \frac{(z_{ilt})^{1-\gamma_{ilt}} - 1}{1-\gamma_{ilt}}$ . Third, I introduce  $\alpha_{lt}$ , a preference weight for electrical services relative to non-electrical goods and services. This parameter depends directly on electricity reliability.

Dropping the subscripts for ease of notation, the consumer's optimization problem becomes:

$$\max_{x,y} \left\{ U = y + \alpha \cdot \frac{(\tau x)^{1-\gamma} - 1}{1 - \gamma} \right\} \quad \text{subject to} \quad ex + y = w \quad (7)$$

The first-order conditions yield the following demand function:

$$x = \left( \frac{\alpha \tau^{1-\gamma}}{e} \right)^{\frac{1}{\gamma}} \quad (8)$$

The inverse demand function derived from the first-order condition is:

$$e(x) = \alpha \tau^{1-\gamma} x^{-\gamma} \quad (9)$$

Consumer surplus (CS) is calculated by integrating the demand curve from zero to the observed consumption level  $x$  and subtracting total expenditure. Evaluating this integral yields the following expression for CS:<sup>27</sup>

$$CS = \alpha \tau^{1-\gamma} x^{1-\gamma} \cdot \frac{\gamma}{1-\gamma} \quad (10)$$

To estimate consumer surplus using equation (10), the following parameters need to be calibrated:  $\gamma, \alpha, \tau$ .

## 5.2.2 Calibration of Parameters

**Calibration of  $\gamma$ .**  $\gamma$  can be obtained once I estimate the price elasticity of demand  $\epsilon$ . Since  $\frac{1}{\gamma}$  represents the IES, I use the demand function in equation (8) to express  $\epsilon$  as follows:

$$\epsilon = \frac{dx}{de} \cdot \frac{e}{x} = -\frac{1}{\gamma} \quad (11)$$

I estimate the short-run elasticities of electricity use by exploiting the spatiotemporal variation in the price of electricity. For the short run, I use simulated price instruments (for examples, see [Ito \(2014\)](#) or [Sears \(2022\)](#)) and follow the approaches typically used in the industrial organization literature. Appendix [Section A.6.5](#) presents the details of the methodology that I use to estimate these elasticities.

Table 19 reports the results of regressions of residualized electricity at time  $t$  and the simulated price instrument. Table 20 reports the first-stage estimates of the regression of  $\Delta \ln(P_{i,t})$  on  $\Delta \ln(\widehat{P}_{i,t})$ .

Table 1 reports short-run price elasticities separately for the sample of bills before and after reliability projects and separately for households and firms with low and high levels of electricity use. Each coefficient and its associated standard error correspond to a different regression. I note three patterns from table 1: (i) consumers are generally more price sensitive after electricity reliability improves, (ii) firms are more price sensitive than households, and (iii) consumers with a high level of consumption are more price sensitive than consumers with a low level of consumption.

<sup>27</sup>See Appendix [Section A.6](#) for a full step-by-step derivation of equation (10).

Table 1: Short-run price elasticity of demand before and after reliability projects

|                            | $\Delta \log$ electricity use (kWh) |                        |
|----------------------------|-------------------------------------|------------------------|
|                            | Before<br>(1)                       | After<br>(2)           |
| <i>Panel A: Households</i> |                                     |                        |
| All                        | -0.6318***<br>(0.0174)              | -0.6647***<br>(0.0203) |
| Low Consumption            | -0.5623***<br>(0.0190)              | -0.4732***<br>(0.0215) |
| High Consumption           | -0.9424***<br>(0.0430)              | -1.229***<br>(0.0531)  |
| <i>Panel B: Firms</i>      |                                     |                        |
| All                        | -0.2020***<br>(0.0476)              | -0.7602***<br>(0.0429) |
| Low Consumption            | -0.6471***<br>(0.0648)              | -1.718***<br>(0.0452)  |
| High Consumption           | -0.0125<br>(0.0571)                 | -0.1470**<br>(0.0616)  |
| District-Year-Month FEs    | Yes                                 | Yes                    |
| Meter Type FEs             | Yes                                 | Yes                    |
| Past Cons. Controls        | Yes                                 | Yes                    |
| Consumer FEs               | Yes                                 | Yes                    |

Notes: This table reports the price elasticity of electricity use for residential and commercial clients between 2012 and 2023, similar to the the elasticity reported in column (6) of Table 23. Each coefficient and its associated standard error correspond to the second-stage estimates of equation (87) based on the sample described in the first column of the table. In each regression, the independent variable is  $\Delta \log(\hat{MP})$ , the simulated marginal price instrument constructed using consumption levels from six months earlier, as described in equation (86) and visualized in Figures 46 and 47. Low-consumption households are households with mean monthly consumption below 170 kWh, and high-consumption households are households with mean monthly consumption above 170 kWh. Low-consumption firms are firms with mean monthly consumption below 60 kWh, and high-consumption firms are firms with mean monthly consumption above 60 kWh. Columns (1) reports the price elasticity before reliability projects, while column (2) report the price elasticity after reliability projects. All regressions include district by month by year fixed effects interacted with dummies for consumption deciles at  $t - 6$ , along with consumer fixed effects. The regressions are run using 94,475,034 bills between 2012 and 2024. Standard errors are clustered at the consumer level. \* :  $p < 0.1$ , \*\* :  $p < 0.05$ , and \*\*\* :  $p < 0.01$ .

Figures ?? and ?? show the change in electricity price and electricity use 4 years before forced reclassification events and 5 years after such events for households and firms, respectively. The figures show a price elasticity of -0.53 and -0.98 5 years after the price change for households and firms, respectively.

The estimated short-run and long-run price elasticities are generally low in magnitude and less than unity, which means that consumers' demand for electricity is price inelastic. These estimates are broadly in line with those documented by [Zhu et al. \(2018\)](#) in their meta-analysis of 103 papers on the price and income elasticities of electricity demand. They find that the short-term price elasticity estimates range from -0.948 to 0.61, with a mean of -0.228, while the long-term price elasticity estimates range from -4.2 to 0.6, with a mean is -0.577. The estimates in the medium term a few years after the events are between the mean short-term and long-term estimates.

Using the relationship from equation (11):

$$\gamma = -\frac{1}{\epsilon} \quad (12)$$

**Calibration of  $\tau$ .** To accurately model the conversion of electricity consumption  $x$  into electrical services  $z$  in the context of Senegal, a low-income country, I calibrate the technical efficiency parameter  $\tau$  focusing on the most commonly used appliances: lighting, refrigerators, and televisions. These appliances constitute the primary contributors to household electricity consumption in Senegal.

Specifically, I retrieve per-household usage weights for lighting, refrigerators, and televisions directly from the baseline appliance roster dataset for the universe of clients in the billing data.<sup>28</sup> I use this dataset so that the usage weights reflects each consumer's actual electricity consumption profile. Next, I assume that household-level adoption rates for energy-efficient appliances follow a uniform distribution between 0 and 1, thereby capturing variability across consumers. I then calculate the adjusted efficiency (AE) and adjusted contribution (AC) for each appliance based on recognized efficiency improvement factors from external commercial reports—for instance, lumens per watt for modern lighting or average annual kWh for refrigerators. Finally, by integrating these usage weights and adoption rates, I obtain a per-household value of  $\tau$ , which offers a comprehensive measure of technical efficiency that accounts for both consumption patterns and technology diffusion. Appendix [Section A.6.4](#) walks through the calculations and assumptions to calibrate  $\tau$  for a given consumer.

---

<sup>28</sup>As mentioned in [Section 3](#), this is the roster of appliance that the utility company collects at baseline to decide the right usage category of the customer.

| Table 2: Results of Calibration of Energy Efficiency Factor $\tau$ |                 |        |                |        |          |         |
|--------------------------------------------------------------------|-----------------|--------|----------------|--------|----------|---------|
|                                                                    | Before projects |        | After projects |        |          |         |
|                                                                    | Mean            | SE     | Mean           | SE     | Obs      | P-Value |
|                                                                    | (1)             | (2)    | (3)            | (4)    | (5)      | (6)     |
| <b>Households</b>                                                  |                 |        |                |        |          |         |
| Small HHs                                                          | 2.44            | 0.0002 | 2.66           | 0.0002 | 28536647 | 0.00    |
| Large HHs                                                          | 2.38            | 0.0002 | 2.65           | 0.0002 | 39845058 | 0.00    |
| <b>Firms</b>                                                       |                 |        |                |        |          |         |
| Small firms                                                        | 2.56            | 0.0005 | 2.68           | 0.0005 | 6103888  | 0.00    |
| Large firms                                                        | 2.51            | 0.0004 | 2.66           | 0.0003 | 9908432  | 0.00    |

Notes: This table reports the mean and standard errors of the energy efficiency factor, separately for small and large residential and commercial clients, both before and after reliability improvement projects. Columns (1) and (2) correspond to values estimated using the billing data before the reliability improvement projects. Columns (3) and (4) correspond to values estimated using the billing data after the reliability improvement projects. Higher values correspond to higher efficiency levels. The calibration is done using the baseline appliance roster in the billing data, along with external data on efficiency levels of appliances. An example of the calibration exercise done to derive  $\tau$  for each consumer is shown in Appendix [Section A.6.4](#). The categorization of households and firms between small and large follows the official categorization done by the utility company. The “Obs” column gives the total number of bills used in the calibration of  $\tau$  for each group of customers. “P-Value” is the p-value of the difference in mean factor value before and after reliability changes.

Table ?? above reports the mean and standard deviations of the energy efficiency factor, separately for small and large residential and commercial clients, both before and after reliability improvement projects. The calibrated values of  $\tau$  whose moments are shown in table reflects the average technical efficiency of appliances used by consumers in the billing data, accounting for both the efficiency improvements of modern appliances and their adoption rates. Higher values reflect higher energy efficiency factors, or more electrical services generated from electricity usage and appliances. We note two patterns in the data. First, firms have a higher energy efficiency factor than households. Second, larger clients have a higher energy efficiency factor than small clients. Third, energy efficiency improves after the reliability projects. These improvements reflect a higher adoption of fancier appliances which are more likely to be energy-efficient than more common appliances, through these improvements cannot be entirely attributed to the reliability projects.

**Calibration of  $\alpha$ .** The parametrization that I did to derive a closed-form formula of the consumer surplus allows me to the preference parameter  $\alpha$  to be equal to  $\rho$ , the share of moments with electricity on. To calibrate this share using the outage data, I need to obtain two inputs: (A) the total hours for which the consumer *would* use electricity, and

(B) the total hours for which electricity is unavailable. B is directly observed in the outage data. To obtain A, I impose that total hours follows a uniform distribution from 150 hours to 450 hours (5 hours to 15 hours a day). Once I have inputs (i) and (ii),  $\alpha$  is calculated as :  $1 - B/A$ .

| Table 3: Results of Calibration of Preference Parameter $\alpha$ |                 |         |                |         |          |         |
|------------------------------------------------------------------|-----------------|---------|----------------|---------|----------|---------|
|                                                                  | Before projects |         | After projects |         | Obs      | P-Value |
|                                                                  | Mean            | SE      | Mean           | SE      |          |         |
|                                                                  | (1)             | (2)     | (3)            | (4)     | (5)      | (6)     |
| <b>Households</b>                                                |                 |         |                |         |          |         |
| Small HHs                                                        | 2.87            | 0.00005 | 2.89           | 0.00006 | 32327337 | 0.00    |
| Large HHs                                                        | 2.88            | 0.00005 | 2.92           | 0.00004 | 43125523 | 0.00    |
| <b>Firms</b>                                                     |                 |         |                |         |          |         |
| Small firms                                                      | 2.87            | 0.00011 | 2.90           | 0.00013 | 6838460  | 0.00    |
| Large firms                                                      | 2.88            | 0.00010 | 2.92           | 0.00008 | 10989508 | 0.00    |

Notes: This table reports the mean and standard errors of the parameter governing the preference for electrical services over non-electrical goods/services, separately for small and large residential and commercial clients, both before and after reliability improvement projects. Columns (1) and (2) correspond to values estimated using the outage data before the reliability improvement projects. Columns (3) and (4) correspond to values estimated using the outage data after the reliability improvement projects. Higher values correspond to higher efficiency levels. The categorization of households and firms between small and large follows the official categorization done by the utility company. The “Obs” column gives the total number of bills used in the calibration of  $\tau$  for each group of customers. “P-Value” is the p-value of the difference in mean factor value before and after reliability changes.

Table 3 above reports the mean and standard deviation of the parameter governing preferences for electrical services over non-electrical goods/services, separately for small and large residential and commercial clients, both before and after reliability improvement projects. Overall, the value of  $\alpha$  increases post reliability projects since these projects led to more hours of grid electricity available to consumers.

### 5.2.3 Estimates of Consumer Surplus

Table 4 reports the average consumer surplus per kWh before and after reliability projects, separately for large and small residential and commercial clients. The table presents group-level mean consumer surplus per kWh estimates and their associated standard errors for different customer groups before and after reliability projects using the closed-formula for consumer surplus in equation (10) and the calibrated parameters in the previous subsections.

**Table 4: Consumer surplus per kWh before and after reliability projects**

|                   | Before projects |         | After projects |         | Obs    | P-Value |
|-------------------|-----------------|---------|----------------|---------|--------|---------|
|                   | Mean            | SE      | Mean           | SE      |        |         |
|                   | (1)             | (2)     | (3)            | (4)     | (5)    | (6)     |
| <b>Households</b> |                 |         |                |         |        |         |
| Small HHs         | 0.20            | 0.00001 | 0.22           | 0.00001 | 649035 | 0.00    |
| Large HHs         | 0.36            | 0.00001 | 0.36           | 0.00001 | 574149 | 0.00    |
| <b>Firms</b>      |                 |         |                |         |        |         |
| Small firms       | 0.22            | 0.00017 | 0.28           | 0.00009 | 166903 | 0.00    |
| Large firms       | 0.44            | 0.00012 | 0.55           | 0.00023 | 130331 | 0.00    |

Notes: This table reports the consumer surplus derived from electricity consumption before and after reliability projects, separately for large and small residential and commercial clients. The table presents group-level mean consumer surplus per kWh estimates and their associated standard errors for different customer groups before and after reliability projects. To do this, the energy efficiency factor ( $\tau$ ), preference weight ( $\alpha$ ), and the coefficient of relative risk aversion ( $\gamma$ ) were calibrated using observed consumption and outage data. With these calibrated values, individual consumer surplus per kWh was calculated based on the closed-formula for consumer surplus in equation (10). To obtain standard errors at the individual level, the delta method was applied, yielding standard errors for each individual's consumer surplus estimate that incorporate uncertainty in the estimated parameters ( $\tau$ ,  $\alpha$ , and  $\gamma$ ). For each group of customers, the mean consumer surplus was computed as the arithmetic average of the individual estimates. The standard error of the group mean was obtained by aggregating the individual standard errors: specifically, squaring each individual's standard error, summing these squares within the group, dividing by the square of the group's sample size, and taking the square root. Columns (1) and (2) correspond to values estimated using the billing data before the reliability improvement projects. Columns (3) and (4) correspond to values estimated using the billing data after the reliability improvement projects. "Obs." shows the number of clients in each customer group. "P-Value" is the p-value of the difference in mean consumer surplus before and after reliability changes. The PPP exchange rate of 1 US dollar = 219.13 CFA franc (according to the World Bank's conversion factor) is used for the conversion of monetary values from CFA franc to US dollars.

Table 4 shows that consumer surplus per kWh increased after the reliability projects across all four customer groups. For both households and firms, the mean surplus in the post-project period is higher than in the pre-project period, with all reported changes being statistically significant (p-value = 0.00). Furthermore, firms generally experience higher consumer surplus compared to households, as evidenced by larger mean values for both small and large firms relative to their household counterparts. Additionally, the change in surplus from before to after the projects is generally greater for smaller units (both small households and small firms) than for larger units, indicating that smaller firms and households may derive relatively more benefit from improved reliability. The larger relative gains for smaller units suggest that reliability improvements disproportionately benefit smaller consumers, who may have been more adversely affected by

unreliable service and who are less likely to have defensive investments to deal with unreliable electricity.

After reliability projects, the average consumer surplus per kWh for small households increased from approximately \$0.20 to \$0.22. Given that small households consume an average of about 59.19 kWh, this improvement translates into an increase of roughly \$1.18 in consumer surplus per household. This increase reflects the added value derived from enhanced reliability of electricity, as customers can utilize electricity more effectively and with fewer interruptions. For small firms, the mean surplus per kWh rose from \$0.22 to \$0.28. With an average consumption of around 36.59 kWh, this change leads to an approximate gain of \$2.19 in total consumer surplus per small firm. In contrast, large firms experienced a substantial rise in surplus from \$0.44 to \$0.55 per kWh. These variations among groups highlight that while all customer segments benefit from reliability improvements, the magnitude of the increase in consumer surplus depends on both the change in surplus per kWh and the scale of electricity consumption.

These findings can also be interpreted in the economic context of Senegal. The utility company currently sets the marginal price of electricity at approximately 0.20 US dollars per kWh for the first consumption block and slightly higher for subsequent blocks. Given this price level, the observed increases in consumer surplus—albeit numerically modest—reflect substantial improvements in value received by consumers due to enhanced reliability. The monetary values of the changes in consumer surplus can also be analyzed with local income. For small households for example, the average consumer surplus per kWh increased from about 20 cents before reliability projects to approximately 22 cents after these projects. This improvement in surplus translates into an increase of roughly \$1.18 per household using their average consumption of 59.19 kWh. This gain represents about 0.07% of GDP per capita, based on Senegal’s 2022 GDP PPP of \$1,598.73.

### 5.3 Discrete Choice Model

One limitation of the consumer surplus estimates in subsection 5.2 is that it focuses on consumers that are already clients of the utility company and are thus already connected to the grid. In this subsection, I turn to the discrete choice model of the joint choice of how to access electricity (none, grid, solar) and the wattage level of appliances to have as a function of the number of hours that electricity is available, the cost of using electricity, the cost of appliances, and other characteristics of the households. This discrete choice model is estimated using the EHCVM data presented in Section 3, and which include both consumers already connected to the grid and consumers unconnected to the grid.

In the model, hours are separated between peak and off-peak hours and are total time less hours of blackouts. The model thus provides the marginal value of an hour of availability. It can be easily extended to account for other dimensions of reliability for which I have data, such as the nature of outages (planned vs. unplanned) and the

frequency or intermittency of outages (number of outages).<sup>29</sup> The discrete choice model allows for random correlated coefficients to account for random preference variation over choice attributes, flexible substitution patterns between choices, and correlations in unobserved factors over time. I estimate the model via SML estimation.

Throughout this section, I divide appliances into low-intensity appliances and high-intensity appliances, depending on their general wattage levels. Low-intensity appliances are those with average wattage levels below 200 watts and include light bulbs, TVs, fans, refrigerators, laptops, and tablets. High-intensity appliances are those with average wattage levels above 200 watts and include air conditioners, freezers, irons, generators, microwaves, mixers, printers, vacuums, and washing machines.

This subsection proceeds as follows: I start by transforming the economic model presented in Subsection 5.1 above into a random utility model. Then, I estimate the standard discrete choice model by instrumenting grid supply with the reliability projects. I then use my estimates to derive the baseline MWTP for grid supply and the change in this MWTP following reliability improvements. Finally, I employ a mixed logit model to flexibly estimate changes in the MWTP for grid supply across multiple covariates, including the appliance composition. I estimate this mixed logit model via SML estimation.

### 5.3.1 From Economic Model to Discrete Choice Model

To go from the economic model presented above to the discrete choice model, I start by deriving the optimal electricity and service consumption (based on the assumed efficiency of appliances). I then parametrize the utility function and use Taylor approximations to go from the transformed economic model to a discrete choice model, and I finish with some economic adjustments to make the model a more standard discrete choice model. Appendix Section A.6 shows the steps taken in more detail.

Taking the first-order condition of equation(4) subject to budget constraints yields the following:

$$\mu'_i(\tau_{ilt}(\rho_{lt}) \cdot x_{i1lt}) \cdot \tau_{ilt}(\rho_{lt}) = e_t \quad (13)$$

Solving for  $x_{i1lt}$  yields the following:

$$\mu'_i(s_i) = \frac{e_t}{\tau_{ilt}(\rho_{lt})}, \quad \text{where } z_i = \tau_{ilt}(\rho_{lt}) \cdot x_{i1lt} \quad (14)$$

As a reminder,  $z_i$  denote the total electrical services consumed. The inverse of the marginal utility function gives the following:

---

<sup>29</sup>However, the model abstracts away from dimensions of reliability that I do not observe directly, such as voltage fluctuations. See Berkouwer et al. (2024) for a paper on the economic consequences of voltage fluctuations.

$$z_i = (\mu'_i)^{-1} \left( \frac{e_t}{\tau_{ilt}(\rho_{lt})} \right) \quad (15)$$

Thus, the optimal electricity consumption is as follows:

$$x_{i1lt} = \frac{z_i}{\tau_{ilt}(\rho_{lt})} = \frac{1}{\tau_{ilt}(\rho_{lt})} \cdot (\mu'_i)^{-1} \left( \frac{e_t}{\tau_{ilt}(\rho_{lt})} \right) \quad (16)$$

For analytical tractability, I assume that the utility function  $\mu_i(\cdot)$  is logarithmic:

$$\mu_i(z) = \ln(z) \quad (17)$$

Using equation (15), I have

$$z_i = (\mu'_i)^{-1} \left( \frac{e_t}{\tau_{ilt}(\rho_{lt})} \right) = \frac{\tau_{ilt}(\rho_{lt})}{e_t} \quad (18)$$

and the optimal electricity consumption simplifies to

$$x_{i1lt} = \frac{z_i}{\tau_{ilt}(\rho_{lt})} = \frac{1}{e_t} \quad (19)$$

Substituting  $z_i$  and  $x_{i1lt}$  back into equation (4) gives the indirect utility function:

$$u_{ilt} = \rho_{lt} \cdot \ln \left( \frac{\tau_{ilt}(\rho_{lt})}{e_t} \right) + w_{ilt} - \rho_{lt} \quad (20)$$

I now expand the model to include different options  $j$  for accessing electricity service. The options  $j$  are as follows: 0 for no connection to the grid with no appliances, 1 for grid connection with only low-intensity appliances, 2 for grid connection with both low- and high-intensity appliances, and 3 for the use of solar panels only. Each option  $j$  is defined by characteristics including reliability  $\rho_{ljt}$  (availability of electricity in location  $l$  at time  $t$ ), the electricity price  $e_{jt}$ , appliance efficiency  $\tau_{iljt}$  (efficiency of appliances under option  $j$ ), and the appliance cost  $p_{ljt}$ .

The utility function for comparing these options omits any term that is constant across choices, such as  $w_{ilt}$ , because it does not affect the utility difference. This simplification allows the utility to be expressed as follows:

$$u_{ilt} = \rho_{lt} \cdot (\ln(\tau_{ilt}(\rho_{lt})) - \ln(e_t) - 1)$$

Adjusting for the characteristics of each service option  $j$ , the utility function becomes

$$u_{iljt} = \rho_{ljt} \cdot (\ln(\tau_{iljt}) - \ln(e_{jt}) - 1) - p_{ljt}$$

To further capture differences in reliability and appliance efficiency between peak and off-peak hours, I decompose these elements. Let  $h_{1,ljt}$  and  $h_{2,ljt}$  represent the hours of electricity availability during peak and off-peak hours, respectively. Thus, total reliability is given by  $\rho_{ljt} = h_{1,ljt} + h_{2,ljt}$ . Similarly, I allow appliance efficiency to vary by time, as denoted by  $\tau_{iljt}^{\text{peak}}$  for peak hours and  $\tau_{iljt}^{\text{off}}$  for off-peak hours.

Generalizing equation (20) yields the following:

$$u_{iljt} = h_{1,ljt} \cdot \left( \ln \left( \tau_{iljt}^{\text{peak}} \right) - \ln(e_{jt}) - 1 \right) + h_{2,ljt} \cdot \left( \ln \left( \tau_{iljt}^{\text{off}} \right) - \ln(e_{jt}) - 1 \right) - p_{ljt}$$

Using a first-order Taylor approximation around a reference price  $\bar{e}$ , I have the following:

$$\ln(e_{jt}) \approx \ln(\bar{e}) + \frac{1}{\bar{e}}(e_{jt} - \bar{e})$$

and the utility function simplifies to

$$u_{iljt} = \beta_{3,ilt} h_{1,ljt} + \beta_{4,ilt} h_{2,ljt} + \beta_5 e_{jt} (h_{1,ljt} + h_{2,ljt}) - p_{ljt} \quad (21)$$

where  $\beta_{1,ilt} = \ln \left( \tau_{iljt}^{\text{peak}} \right) - \ln(\bar{e}) - 1 + \gamma \bar{e}$ ,  $\beta_{2,ilt} = \ln \left( \tau_{iljt}^{\text{off}} \right) - \ln(\bar{e}) - 1 + \gamma \bar{e}$  and  $\beta_3 = -\gamma$ . To account for unobserved factors, I include fixed effects and an error term:

$$u_{iljt} = \beta_{1,ilt} h_{1,ljt} + \beta_{2,ilt} h_{2,ljt} + \beta_3 e_{jt} + \beta_4 p_{ljt} + \xi_{jt} + \xi_l + \xi_{ij} + \epsilon_{iljt} \quad (22)$$

I also make two additional adjustments. First, I unrestrict  $\beta_4$ , the coefficient on the appliance cost, to allow it to be estimated from the data.<sup>30</sup> Second, I drop the interaction term on electricity costs to make the coefficient easier to interpret.

Our final random utility for household  $i$  living in EA (or market)  $l$  from option  $j$  at time  $t$  is as follows:

$$u_{iljt} = \beta_{1,ilt} h_{1,ljt} + \beta_{2,ilt} h_{2,ljt} + \beta_3 e_{jt} + \beta_7 p_{ljt} + \xi_{jt} + \xi_l + \epsilon_{iljt} \quad (23)$$

where  $h_{1,ljt}$  represents the hours of electricity available in market  $l$  under plan  $j$  during peak hours in year  $t$ , defined as the time window between 7 pm and 11 pm, in line with how the utility company defines it.  $h_{2,ljt}$  represents the hours of electricity available in market  $l$  under option  $j$  during off-peak hours, defined as the time window between 11 pm and 7 pm. For grid connection,  $h_{1,ljt}$  and  $h_{2,ljt}$  correspond to the total number of hours for which the distribution station serving market  $l$  is not subject to outages. For clients not connected to the grid or with no solar panels,  $h_{1,ljt}$  and  $h_{2,ljt}$  are 0 by construction. For clients connected to solar panels, this reflects the availability of peak

<sup>30</sup>In the derivation, the coefficient on the appliance cost was set to  $-1$  for simplicity. However, there is no theoretical basis for this simplified assumption.

and off-peak solar energy at time  $t$ .  $e_{jt}$  represents the marginal price of electricity under option  $j$  at time  $t$ .<sup>31</sup>  $p_{iljt}$  is the unit price of appliances under option  $j$  in market  $l$ .<sup>32</sup>

$\xi_{jt}$  are product fixed effects, which account for unobserved time-invariant choice characteristics that are homogeneous across locations.  $\xi_{lt}$  are location-time fixed effects, which account for unobserved time-varying location characteristics that are homogeneous across products.  $\xi_{lt}$  thus accounts for general changes in grid reliability over time.  $\epsilon_{iljt}$  is the error term, which includes unobservable quality characteristics.  $\epsilon_{iljt}$  follows a type I extreme value distribution. Customer  $i$  chooses option  $j$  if  $U_{iljt} > U_{ilk} \forall k \neq j$ .  $U_{i0t}$ , the utility corresponding to no grid connection and no appliances, is normalized to  $\epsilon_{i0t}$ .

Each coefficient corresponds to a specific component from the economic model.  $\beta_{3,ilt}$  is the marginal utility of electricity availability during peak hours, incorporating appliance efficiency and base electricity price.  $\beta_{4,ilt}$  is the marginal utility of electricity availability during off-peak hours.  $\beta_5$  is the marginal disutility of electricity price  $e_{jt}$  per hour of availability.  $\beta_6$  is the coefficient on connection costs  $c_{jt}$ , and  $\beta_7$  is the coefficient on appliance costs  $p_{ljt}$ .

The fixed effects  $\xi_{jt}$  and  $\xi_l$  capture unobserved attributes that are specific to products and locations. The error term  $\epsilon_{iljt}$  accounts for random shocks and unobserved factors.

### 5.3.2 Mixed Logit Model

**Model.** Appendix Section A.7 presents the results of a standard logit model, with instruments for electricity reliability and electricity prices. While the standard logit model allows me to obtain the baseline MWTP for grid supply and to use reliability projects to estimate the change in this MWTP, it does not allow me to flexibly estimate the change in MWTP across covariates. In this subsection, I turn to a mixed logit model for this purpose.

I return to the same random utility model in equation (23):

$$u_{iljt} = \beta_{1,ilt}h_{1,ljt} + \beta_{2,ilt}h_{2,ljt} + \beta_{3,ilt}e_{jt} + \beta_{4,ilt}p_{ljt} + \xi_j + \xi_{lt} + \epsilon_{iljt} \quad (24)$$

Now,  $\beta_{1,ilt}$ ,  $\beta_{2,ilt}$ ,  $\beta_{3,ilt}$  and  $\beta_{4,ilt}$  are individual-specific random coefficient, allowing for heterogeneity in preferences. In particular, each of these random coefficients takes

<sup>31</sup>Since the tariff schedule follows an increasing block pricing, the marginal price changes as usage levels change. In addition, consumers with low-intensity appliances could in practice use them enough to end up with as high a marginal price or even higher than consumers with high-intensity appliances not used often. Hence, I calculate  $e_{jt}$  as the weighted-average marginal price across consumers observed to choose option  $j$  at time  $t$  in the data. This approach accounts for the nondegenerate marginal price across consumers of a given choice  $j$ . See Appendix Section A.3.2 for a discussion of the different electricity plans and how they are set in the Senegalese market.

<sup>32</sup>I calculate  $p_{iljt}$  as the weighted-average purchase price across all appliances used by consumer  $i$  in market  $l$  observed to choose option  $j$  at time  $t$  in the data.

the following form:

$$\beta_{k,ilt} = \gamma_1^{(k)} + \gamma_2^{(k)} \mathbf{LI}_{ilt} + \gamma_3^{(k)} \mathbf{HI}_{ilt} + \gamma_4^{(k)} \mathbf{X} + \nu_{ilt}^{(k)} \quad (25)$$

where  $\gamma_1^{(k)}$  is a homogeneous term common to all consumers and represents the mean utility derived from the variable associated with  $\beta_{k,ilt}$ .  $\mathbf{LI}_{ilt}$  is the total number of low-energy-intensity appliances (TVs, fans, refrigerators, light bulbs, laptops, and tablets) owned by consumer  $i$  at time  $t$ .  $\mathbf{HI}_{ilt}$  is the total number of high-intensity appliances (air conditioners, freezers, washing machines, vacuums, electric irons, or printers) owned by consumer  $i$  at time  $t$ .  $\mathbf{X}$  is a vector of covariates such as income, household size, and the gender and employment status of the household head. The structural errors  $\nu_{ilt}^{(k)}$  are random preference shifters that follow a log-normal distribution.<sup>33</sup>

Each of the  $\beta$  parameters depends on the mean coefficient, consumer-level characteristics, and random unobserved heterogeneity. The mean utility level, which is the part of the utility that does not depend on  $i$ , can be expressed as follows:

$$\delta_{jlt} = \gamma_1^{(1)} h_{1,ljt} + \gamma_1^{(2)} h_{2,ljt} + \gamma_1^{(3)} (e_{jt} + c_{jt}) + \gamma_1^{(4)} p_{jt} + \xi_{jt} + \xi_l + \epsilon_{ijlt} \quad (26)$$

Let  $\mu_{ijlt}^{(k)} = \gamma_2^{(k)} \mathbf{LI}_{ilt} + \gamma_3^{(k)} \mathbf{HI}_{ilt} + \gamma_4^{(k)} \mathbf{X} + \nu_{ilt}^{(k)}$  for  $k = \{1, 2, 3, 4\}$ . The part of the utility that depends on  $i$  is as follows:

$$\mu_{ijlt} = \mu_{ijlt}^{(1)} h_{1,ljt} + \mu_{ijlt}^{(2)} h_{2,ljt} + \mu_{ijlt}^{(3)} (e_{jt} + c_{jt}) + \mu_{ijlt}^{(4)} p_{jt} \quad (27)$$

Assuming a number of consumers  $N_{lt}$  for location  $l$ , the market share for option  $j$  in location  $l$  can then be evaluated using Monte Carlo integration as follows:

$$s_{jlt} = \frac{1}{N_{lt}} \sum_{i=1}^{N_{lt}} \frac{\exp(\delta_{jlt} + \mu_{ijlt})}{\sum_{k=1}^J \exp(\delta_{klt} + \mu_{iklt})} \quad (28)$$

**Estimation.** Let  $\beta = \{\beta_1, \beta_2, \beta_3, \beta_4\}$  be the random coefficients. The probability that individual  $i$  chooses alternative  $j$  at time  $t$ , conditional on the random coefficients  $\beta$  taking a certain realization  $\mathbf{b}$ , is given as follows:

$$P_{iljt} \mid \mathbf{b} = \frac{\exp(V_{ijlt}(\mathbf{b}))}{\sum_{k \in J} \exp(V_{ilklt}(\mathbf{b}))} \quad (29)$$

Since I have random coefficients, the unconditional choice probability is obtained by integrating over the distribution of  $\beta$ :

$$P_{iljt} = E(P_{iljt} \mid \beta) = \int_{\beta} (P_{iljt} \mid \beta) f(\beta, \theta) d\beta \quad (30)$$

---

<sup>33</sup>I choose these distributions so that the model parameters are correctly bounded to have the right sign.

where  $f(\beta, \theta)$  is the joint density function of the random coefficients and  $\theta$  is the vector of the parameters of the distribution of  $\beta$ .

One of the challenges in estimating the mixed logit model is that the choice probabilities, which represent the likelihood that an individual will choose a particular option, involve integrating over the distribution of the random coefficients. This integral cannot be solved analytically and must be approximated using numerical methods. I use SML estimation to approximate this integral. In particular, I simulate the random coefficients by taking a finite number of draws from their assumed distribution and then averaging the choice probabilities over these draws. I use low-discrepancy Halton sequences for the draws to minimize potential inaccuracies in the approximation due to draws not covering the distribution space uniformly.<sup>34</sup> This simulation-based approach allows me to approximate the integral with a high degree of accuracy.

The simulated choice probability for consumer  $i$  is computed as follows:

$$\hat{P}_{iljt} = \frac{1}{R} \sum_{r=1}^R P_{iljt}(\beta_r), \quad (31)$$

where  $R$  is the number of draws,  $\beta_r$  is the distribution of random coefficients for draw  $r$ , and  $P_{iljt}(\beta_r) = P_{iljt} \mid \beta_r = \frac{\exp(V_{iljt}(\beta_r))}{\sum_{k \in J} \exp(V_{ilkt}(\beta_r))}$ .

The simulated log-likelihood function for the entire sample is given as follows:

$$\hat{\mathcal{L}}(\theta) = \sum_{i=1}^N \sum_{t=1}^T \sum_{j=1}^J \sum_{l=1}^L \ln(\hat{P}_{iljt}) \quad (32)$$

where  $N$  is the number of individuals in the EHCVM dataset,  $T$  is the number of time periods in the EHCVM dataset, and  $J$  is the number of choices.

The parameters to be estimated are  $\theta = \{\gamma_1^{(k)}, \gamma_2^{(k)}, \gamma_3^{(k)}\}$  for each  $k$  in the model. These  $\theta$  parameters are estimated by maximizing the simulated log-likelihood function, which is computed by summing the logarithms of the simulated choice probabilities across all individuals and choice options in equation (32).

I estimate the mixed logit model by pooling the two rounds of household surveys. I face the common endogeneity issue that electricity reliability is driven by factors unobserved by the researcher and potentially correlated with other choice characteristics (tariff schedules, price of appliances, etc.). Hence, naively estimating the discrete choice model based on the EHCVM data might lead to biased coefficients and erroneous consumer surplus estimates for reliability improvements. I include product-by-year fixed effects, location fixed effects, and individual fixed effects in the estimation to absorb the

---

<sup>34</sup>A Halton sequence is a type of quasi-random sequence that provides more uniform coverage of the distribution space than purely random sequences provide. They are constructed using prime number bases and are designed to avoid clustering and gaps, which are common in random sequences.

unobserved time-invariant product or individual characteristics, as well as general time trends in reliability changes, tariff schedules, connection costs, and appliance prices.

Once the model is estimated, the MWTP for each attribute is directly obtained from the estimated coefficients. The MWTP for one additional hour of supply of electricity from the grid during peak hours is as follows:

$$\text{MWTP}_i = \frac{\beta_{1,it}}{\beta_{3,it} + \beta_{4,it}} \quad (33)$$

The MWTP for one additional hour of supply of electricity from the grid during off-peak hours is as follows:

$$\text{MWTP}_i = \frac{\beta_{2,it}}{\beta_{3,it} + \beta_{4,it}} \quad (34)$$

To estimate the effect of reliability projects, I predict the random parameters in a counterfactual scenario in which electricity reliability changes according to the results in the empirical analysis in Section 4. To do so, I add the change in peak and off-peak electricity availability coming from reliability projects to the baseline level of reliability prior to the projects. In particular, I assign to each customer the average change in electricity use experienced by clients served by the same distribution station that they rely on. As shown in Section 4 above, this average change is causally identified by comparing customers from the billing data who are served by treated stations earlier in the sample to customers served by not-yet-treated stations. The estimation thus rests on the assumption that the event study estimates reported in Section 4 give us change in electricity availability driven solely by the reliability projects and not by other factors.

My estimation also partials out any individual-specific variation that does not vary between connection types, such as income, household characteristics, employment status, age, gender, and reliability metrics. Similarly, time-varying factors such as reliability factors that do not vary between choices are also partialled out by the estimation.

**Results.** Table 5 reports the results of the mixed logit model presented. The last three columns of the table show the set of model parameters that are specified in equation (25) and that minimize the log-likelihood function specified in equation (32).

For the variables related to electricity reliability, Table 5 shows positive  $\gamma_1$ s,  $\gamma_2$ s, and  $\gamma_3$ s estimated for  $\beta_1$  and  $\beta_2$ . These positive coefficients indicate a positive mean utility among customers from both peak and off-peak electricity availability, as well positive additional utility derived from the basic and sophisticated appliances that customers own. In general, the mean utility derived from peak hours of electricity and the additional utility derived from the appliance composition during these peak hours is higher than the utility derived from off-peak hours of electricity and the additional utility derived from the appliance composition during these off-peak hours; however,

the difference does not appear to be substantial. For the cost variables, Table 5 shows negative  $\gamma_1$ s,  $\gamma_2$ s, and  $\gamma_3$ s estimated for  $\beta_3$  and  $\beta_4$ . These negative parameters indicate a positive mean disutility among customers from both electricity usage costs and appliance costs, as well positive additional disutility derived from the usage and costs of the basic and sophisticated appliances that customers own. In general, Table 5 shows a marginal positive utility for electricity availability in peak and off-peak hours and a marginal disutility for the costs associated with such consumption (electricity prices, appliance costs, connection costs, etc.).

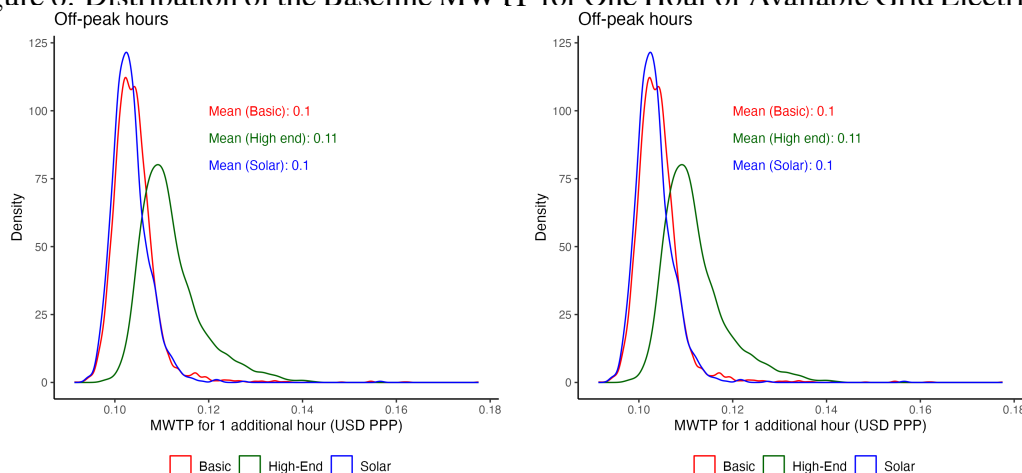
**Table 5: Results of Model Estimation via Simulated Maximum Likelihood**

| Variable        | Parameter | Distribution of MXL parameters |      |       |        |       | SML parameters    |                   |                   |                   |
|-----------------|-----------|--------------------------------|------|-------|--------|-------|-------------------|-------------------|-------------------|-------------------|
|                 |           | Mean                           | SD   | Min   | Median | Max   | Constant          | Basic             | High-End          | Income            |
|                 |           | (1)                            | (2)  | (3)   | (4)    | (5)   | $\gamma_1$<br>(6) | $\gamma_2$<br>(7) | $\gamma_3$<br>(8) | $\gamma_4$<br>(9) |
| Peak hours      | $\beta_1$ | 3.96                           | 1.02 | 1.06  | 3.75   | 11.56 | 0.51              | 0.49              | 0.52              | 0.47              |
| Off-peak hours  | $\beta_2$ | 0.94                           | 0.18 | 0.51  | 0.90   | 2.22  | 0.48              | 0.10              | 0.08              | 0.06              |
| Usage costs     | $\beta_3$ | -0.90                          | 0.20 | -2.18 | -0.86  | -0.23 | -0.07             | -0.11             | -0.07             | -0.11             |
| Appliance costs | $\beta_4$ | -1.15                          | 0.41 | -3.91 | -1.10  | -0.12 | -0.11             | -0.33             | -0.06             | -0.14             |

Notes: This table reports the results of the mixed logit (MXL) model presented in Subsections 5.3.2 and 5.3. Columns (1)–(5) present the distribution of individual-level parameters ( $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ , and  $\beta_4$ ) estimated using an MXL model. The utility coefficients  $\beta_1$ – $\beta_4$  are random and modeled as functions of observable household characteristics (ownership of low-intensity and high-intensity appliances) and other covariates and as depending on structural errors that follow a log-normal distribution. The MXL model is run using simulated maximum likelihood (SML) estimation. Columns (6)–(9) show the parameters  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ , and  $\gamma_4$ , which correspond to the mean and interaction terms of the random coefficients in equation (25). These parameters are derived by maximizing the simulated log-likelihood of observed choices based on the average choice probabilities across 75 draws.

Figure 6 below shows the distribution of the customer-level MWTP for electricity availability separately for peak hours and off-peak hours. The individual MWTP for peak and off-peak electricity availability is estimated according to equations (33) and (34). The left-hand side figure shows the estimated MWTP for 1 hour of electricity during peak hours, while the right-hand side figure shows the estimated MWTP for 1 hour of electricity during off-peak hours. The distribution of the MWTP is shown separately for clients connected to the grid with only low-intensity appliances (in red), clients connected to the grid with both low-intensity and high-intensity appliances (in green), and clients with solar panels (in blue).

Figure 6: Distribution of the Baseline MWTP for One Hour of Available Grid Electricity



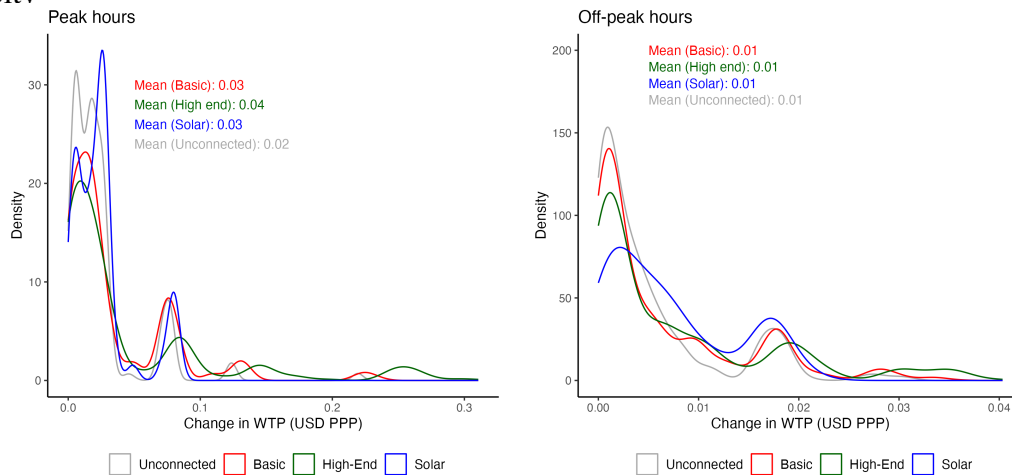
Notes: This figure shows the distribution of the household-level MWTP for grid electricity separately for peak hours and off-peak hours. The individual MWTP for one additional hour of grid electricity is estimated according to equations (33) and (34). These coefficients are obtained through simulated maximum likelihood estimation based on the procedure presented in Subsection 5.3.2 above. The left-hand side figure shows the estimated MWTP for 1 hour of grid electricity during peak hours, while the right-hand side figure shows the estimated MWTP for 1 hour of grid electricity during off-peak hours. The distribution of the MWTP is shown separately for clients connected to the grid with only low-intensity appliances (in red), clients connected to the grid with both low-intensity and high-intensity appliances (in green), and clients with solar panels (in blue). Low-intensity appliances include light bulbs, mobile phones, TVs, refrigerators, fans, laptops, and tablets. High-intensity appliances include air conditioners, freezers, irons, generators, microwaves, mixers, printers, vacuums, and washing machines. The PPP exchange rate of 1 US dollar = 219.13 CFA franc (according to the World Bank's conversion factor) is used for the conversion of monetary values from CFA franc to US dollars.

I note two patterns from Figure 6. First, the MWTP for one hour of grid electricity in peak hours is 3–4 times larger than that in off-peak hours. Second, consumers with high-end appliances have a higher MWTP than consumers with only basic appliances, who have a higher WTP than consumers with solar panels. In terms of magnitude, the MWTP estimates are generally low. However, although these values are very low in absolute terms, they are comparable in relative terms to revealed-preference estimates from developing countries. In the U.S., for example, [Harris \(2023\)](#) find that a household is willing to pay \$1.57 per kWh of outage avoided using a revealed-preference approach based on generator sales around hurricanes seasons in the U.S. [Khanna and Rowe \(2024\)](#) find that households are willing to pay \$1.50 per kWh of lost consumption. Both these estimates and mine represent roughly 1% of monthly electricity bills. In addition to being comparable to estimates from other settings in relative terms, the estimates are broadly in line with the residential market price of electricity in Senegal, which is approximately 0.20 cents/kWh, with my MWTP for peak hours being higher than the unit price of electricity and my MWTP for off-peak hours being lower than the tariffs.<sup>35</sup>

<sup>35</sup>See Figure 22 for a visual representation of the latest tariff schedule and the level of increasing block pricing.

Figure 7 shows the change in MWTP due to reliability projects. The trends in this figure follow those in Figure 6: (i) the change in MWTP for one hour of grid electricity in peak hours is 3–4 times larger than the change in MWTP for one hour of grid electricity in off-peak hours following reliability projects, and (ii) consumers with high-end appliances experience a higher change in MWTP than consumers with only basic appliances and those with solar panels.

Figure 7: Distribution of the Change in MWTP for One Hour of Available Grid Electricity



Notes: This figure shows the distribution of the change in the household-level MWTP for one hour of grid electricity separately for peak hours and off-peak hours. The individual MWTP for one additional hour of grid electricity is estimated according to equations (33) and (34). The change in the MWTP coming from reliability projects is estimated via the model with the counterfactual electricity availability post-reliability projects. These coefficients are obtained through simulated maximum likelihood estimation based on the procedure presented in Subsection 5.3.2 above. The left-hand side figure shows the estimated change in the estimated MWTP for 1 hour of grid electricity during peak hours, while the right-hand side figure shows the change in the estimated MWTP for 1 hour of grid electricity during off-peak hours. The distribution of the change in MWTP is shown separately for clients connected to the grid with only low-intensity appliances (in red), clients connected to the grid with both low-intensity and high-intensity appliances (in green), and clients with solar panels (in blue). Low-intensity appliances include light bulbs, mobile phones, TVs, refrigerators, and fans. High-intensity appliances include air conditioners, freezers, irons, generators, laptops, microwaves, mixers, printers, tablets, vacuums, and washing machines. The PPP exchange rate of 1 US dollar = 219.13 CFA franc (according to the World Bank's conversion factor) is used for the conversion of monetary values from CFA franc to US dollars.

These results relate to the broader literature on the impact of electricity in developing countries. This literature documents relatively underwhelming gains of electricity provision across a few important dimensions, including labor market outcomes and the educational outcomes of children, and a very low WTP for new connections.<sup>36</sup> As highlighted by Lee et al. (2020), providing poor customers with access to electricity without complementary investments might not be enough to improve economic and

<sup>36</sup>While studies such as Dinkelman (2011), Grogan and Sadanand (2013), and Lipscomb et al. (2013) find positive impacts, more recent studies such as Burlig and Preonas (2024a) and Lee et al. (2020) find little to no effect of electricity consumption on many important economic dimensions.

noneconomic outcomes if this provision is not accompanied by the necessary complementary investments in reliable service quality. Figure 7 highlights the important role of electricity reliability in explaining the low WTP for electricity connections and the potential positive impact of reliability improvements in increasing this low WTP. For the last-mile electrification that many developing countries are currently undertaking, the findings suggest that it could be worthwhile to customers if electricity access were provided with a better level of reliability.

## 5.4 Cost–Benefit Analysis

Performing a cost-benefit analysis of reliability projects would require estimates of both the costs and benefits of these projects. The estimates of the WTP for reliability projects provide an estimate of benefits in the short to medium term. On the cost side, the RTUs purchased by the utility company and studied in this paper cost between 6,000 US dollars and 8,000 US dollars. Installing each device costs an additional 300 US dollars, with an additional 40 US dollars per maintenance activity undertaken.

I can measure the cost-effectiveness of reliability projects by estimating their return on investment. Each reliability project costs approximately \$8,000. The annualized benefit per customer is the product of the benefit per kWh of consumption and the number of kWh consumed. The representative consumer in our sample uses 136.8 kWh of electricity per month over the study period.

$$\text{Annual Benefit} = 12 \times 136.8 \times 0.042 = 68.95$$

There are 2,198,624 clients served by 10,789 stations. Thus, each station serves, on average, 204 consumers, which implies a total station-level benefit of \$14,050.90. The internal rate of return can be computed as the solution to the following:

$$8,000 = \sum_{t=1}^T \frac{14,050.90}{(1+r)^t}$$

where  $T$  is the useful life of the reliability improvement device in years. I assume that each device lasts only 5 years, which is a conservative assumption, as reliability devices installed in 2018 or earlier are still mostly functional. This implies that the internal rate of return is approximately 174.5%. This IRR implies that the \$8,000 investment is recouped in less than 1 year. Even under a more conservative assumption of \$0.01 of average realized surplus per consumer (in line with the MWTP for off-peak hours), the IRR amounts to 31.01%. Therefore, the reliability projects that we study have an exceptionally high IRR, unlike other common reliability projects that are costly to implement and yield only modest gains in electricity reliability.<sup>37</sup>

---

<sup>37</sup>For a fuller cost–benefit analysis, I would need additional information. First, I would need to account

## 6 Conclusion

This paper measures the value of energy reliability by using large-scale electricity reliability improvement projects in Senegal as a natural experiment. These projects, which involved the installation of remote devices, enabled the utility company to reduce outage durations by accelerating the resolution process. These interventions are critically important, as electricity reliability poses a significant challenge for both businesses and households in low-income countries. A substantial percentage of firms and households in these settings suffers from persistent electricity shortages, highlighting the urgent need to address this issue. By leveraging data from Senegal, I estimate the change in consumer surplus resulting from these reliability projects, offering valuable insights for policymakers and utility providers.

The findings of this paper provide several key takeaways regarding the distributional impacts of electricity reliability improvements in low-income countries. First, the reliability projects led to a 40% reduction in the duration of outages for treated stations, an 8% increase in electricity consumption for consumers, a noticeable increase in the use of common appliances, and lower exit rates from the grid. Despite requiring substantial investments, these reliability improvements generate a sufficient consumer surplus for households and businesses to justify their cost within a relatively short period. Notably, poorer and smaller consumers, who may lack the flexibility that larger clients have to mitigate the effects of shortages, show a stronger preference for reliability enhancements. The estimated gains in surplus for these groups are small in absolute terms but large in relative terms. This is particularly noteworthy in the context of studies that find a low baseline WTP for electricity access in poor settings. I estimate that the average MWTP for an additional hour of electricity is 40 cents during peak hours and 10 cents during off-peak hours, with the change in WTP due to reliability improvements being approximately 5 cents. These results suggest that reliability improvements can increase the perceived value of grid electricity, potentially enhancing welfare when paired with broader access.

One clear takeaway of the reliability improvement projects studied here is their relatively light-touch nature and cost-effectiveness compared to other capital-intensive interventions often undertaken in low-income countries, such as the undergrounding of power lines, line bifurcations, or grid expansions with additional substations. For instance, although undergrounding is often considered a solution that enhances reliability, it is notoriously expensive, and its effectiveness is dependent on local factors such as the soil structure, weather exposure, and flood vulnerability. Similarly, large-scale

---

for dynamic changes in the benefits expected from the reliability projects. The long-term effects, for example, might be different and arguably more important than the short-term and medium-term benefits. Second, the positive impact of reliability improvement on electricity use means that I may need to account for additional economic damage per ton of emissions, especially in my context where thermal power currently accounts for 70% of electricity supply.

grid expansions involve significant upfront costs and long timelines, requiring major investments in substations, transformers, and transmission lines. Moreover, the maintenance costs for underground infrastructure can be prohibitively high, limiting the cost-effectiveness of such projects. In contrast, the remote device installations analyzed in this study offer a more affordable and flexible alternative, allowing utilities to quickly detect and resolve outages without the need for extensive infrastructure upgrades. This aspect makes the installation of these devices a particularly viable option in countries such as Senegal, where budget constraints and environmental conditions make large-scale interventions such as under-grounding less feasible.

However, it is crucial to approach the notion that reliability improvements are uniformly beneficial with caution. A non-negligible proportion of consumers experience losses in surplus as a result of reliability projects due to the structure of the nonlinear tariff schedule. Specifically, as consumers' increased electricity consumption pushes them into higher tariff blocks, the dis-utility from the higher prices outweighs the benefits of increased electricity availability. For these consumers, limited awareness of the tariff schedule likely compounds the issue, as they are unable to make informed decisions regarding their electricity usage. This point highlights the importance of considering how nonlinear pricing schemes interact with reliability improvements, especially for consumers near tariff block cutoffs or those experiencing significant increases in consumption.

This paper focuses on the partial equilibrium effects of reliability improvements, even though the intervention is part of a networked system where treated and untreated components are interdependent. While this paper estimates the spillover effects of treatment on untreated stations, future research could benefit from exploring the general equilibrium effects of such large-scale interventions with a spatial equilibrium model, since the treated components of the electricity network directly and indirectly affect the untreated components through load sharing, outages, and other incidents. Understanding these broader interactions in highly interconnected infrastructure networks is left for future research.

## References

- ABEBERESE, A. B. (2020): "The Effect of Electricity Shortages on Firm Investment: Evidence from Ghana," *Journal of African Economies*, 26, 46–62.
- ABEBERESE, A. B., C. G. ACKAH, AND P. O. ASUMING (2021): "Productivity Losses and Firm Responses to Electricity Shortages: Evidence from Ghana," *World Bank Economic Review*, 35, 1–18.
- AFROBAROMETER (2019): "Afrobarometer Data, Senegal, Rounds 1-7," <http://www.afrobarometer.org>.
- ALLCOTT, H., A. COLLARD-WEXLER, AND S. D. O'CONNELL (2016): "How Do Electricity Shortages

- Affect Industry? Evidence from India,” *American Economic Review*, 106, 587–624, [10.1257/aer.20140389](https://doi.org/10.1257/aer.20140389).
- BERKOUWER, S., P. BISCAÏE, M. MIKDASH, S. PULLER, AND C. WOLFRAM (2024): “Voltage quality and economic activity,” *Working Paper*, [https://papers.pierrebiscaie.com/BBMPW\\_Voltage\\_Ghana.pdf](https://papers.pierrebiscaie.com/BBMPW_Voltage_Ghana.pdf).
- BORUSYAK, K., X. JARAVEL, AND J. SPIESS (2024): “Revisiting Event-Study Designs: Robust and Efficient Estimation,” *The Review of Economic Studies*, <https://doi.org/10.1093/restud/rdae007>.
- BROWN, D. P. B., AND L. MUEHLENBACHS (2023): “The Value of Electricity Reliability: Evidence from Battery Adoption,” *Working Paper*, [https://media.rff.org/documents/WP\\_23-10.pdf](https://media.rff.org/documents/WP_23-10.pdf).
- BURGESS, R., M. GREENSTONE, N. RYAN, AND A. SUDARSHAN (2022): “Smart Meters and Utility Payments,” *Working Paper*.
- BURLANDO, A. (2014): “Transitory Shocks and Birth Weights: Evidence from a Blackout in Zanzibar,” *Journal of Development Economics*, 108, 154–6.
- BURLIG, F., AND L. PREONAS (2024a): “Out of the Darkness and into the Light? Development Effects of Rural Electrification,” *Journal of Political Economy* (2024).
- (2024b): “Out of the Darkness and into the Light? Development Effects of Electrification in India,” *Journal of Political Economy*, 132.
- CALLAWAY, B., AND P. H. SANT’ANNA (2021): “Difference-in-Differences with multiple time periods,” *Journal of Econometrics*, 225, 200–230, <https://www.sciencedirect.com/science/article/pii/S0304407620303948>.
- CHARMES, J. (2021): “Measuring Time Use: An assessment of issues and challenges in conducting time-use surveys with special emphasis on developing countries,” Produced by The Global Centre of Excellence on Gender Statistics (CEGS), UN Women, <https://www.unwomen.org>, United Nations Entity for Gender Equality and the Empowerment of Women, UN Women.
- COLE, M. A., R. J. ELLIOTT, G. OCCHIALI, AND E. STROBL (2018): “Power outages and firm performance in Sub-Saharan Africa,” *Journal of Development Economics*, 134, 150–159, <https://www.sciencedirect.com/science/article/pii/S0304387818305054>.
- COLMER, J., D. LAGAKOS, AND M. SHU (2023): “Is the Electricity Sector a Weak Link in Development,” *Working Paper*.
- DEUTSCHMANN, J. W., A. POSTEPSKA, AND L. SARR (2021): “Measuring willingness to pay for reliable electricity: Evidence from Senegal,” *World Development*, 138, 105209, <https://www.sciencedirect.com/science/article/pii/S0305750X20303363>.
- DINKELMAN, T. (2011): “The Effects of Rural Electrification on Employment: New Evidence from South Africa,” *American Economic Review*, 101(7), 3078–3108.
- FISHER-VANDEN, K., E. T. MANSUR, AND Q. J. WANG (2015): “Electricity shortages and firm productivity: Evidence from China’s industrial firms,” *Journal of Development Economics*, 114, 172–188, <https://www.sciencedirect.com/science/article/pii/S0304387815000048>.

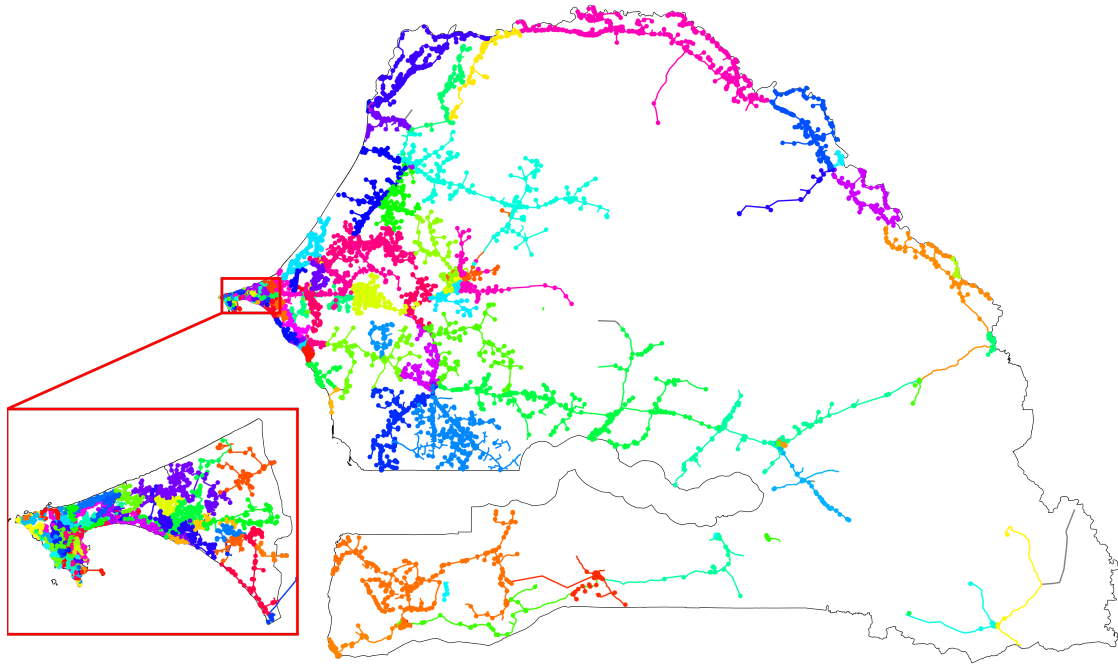
- GORMAN, W. (2022): “The quest to quantify the value of lost load: A critical review of the economics of power outages,” *The Electricity Journal*, 35(8), 107187, <https://doi.org/10.1016/j.tej.2022.107187>.
- GRAINGER, C. A., AND F. ZHANG (2019): “Electricity Shortages and Manufacturing Productivity in Pakistan,” *Energy Policy*, 132, 1000–8.
- GROGAN, L., AND A. SADANAND (2013): “Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua,” *World Development*, 43, 252–65.
- HARDY, M., AND J. MCCASLAND (2021): “Lights Off, Lights On: The Effects of Electricity Shortages on Small Firms,” *World Bank Economic Review*, 35, 19–33.
- HARRIS, R. I. (2023): “Willingness to Pay for Electricity Reliability: Evidence from U.S. Generator Sales,” *Working Paper*.
- IEA, IRENA, UNSD, WORLD BANK, AND WHO (2023): “Tracking SDG 7: The Energy Progress Report,” <https://www.worldbank.org/>, License: Creative Commons Attribution—NonCommercial 3.0 IGO (CC BY-NC 3.0 IGO).
- ITO, K. (2014): “Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing,” *American Economic Review*, 104 (2), 537–563.
- ITO, K., AND S. ZHANG (2020): “Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China,” *Journal of Political Economy*, 128 (5), 1627–1672.
- KAHN-LANG, A., AND K. LANG (2020): “The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications,” *Journal of Business & Economic Statistics*, 38, 613–620, [10.1080/07350015.2018.1546591](https://doi.org/10.1080/07350015.2018.1546591).
- KHANNA, S., AND K. ROWE (2024): “The long-run value of electricity reliability in India,” *Resource and Energy Economics*, 77, 101425, <https://doi.org/10.1016/j.reseneeco.2024.101425>.
- LEE, K., E. MIGUEL, AND C. WOLFRAM (2020): “Does Household Electrification Supercharge Economic Development?” *Journal of Economic Perspectives*, 34(1), 122–44.
- LIPSCOMB, M., A. M. MOBARAK, AND T. 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.
- MASSELUS, L., J. PETERS, G. GONZALEZ SUTIL, V. MODI, J. MUGYENYI, A. MUNYEHIRWE, N. WILLIAMS, AND M. SIEVERT (2024): “10 years after: Long-term adoption of electricity in rural Rwanda,” Ruhr Economic Papers 1086, RWI - Leibniz-Institut für Wirtschaftsforschung, Ruhr-University Bochum, TU Dortmund University, University of Duisburg-Essen, <https://EconPapers.repec.org/RePEc:zbw:rwirep:300563>.
- MEEKS, R., R. ISAEV, A. OMURALIEV, AND Z. WANG (2023): “Impacts of Electricity Quality Improvements: Experimental Evidence on Infrastructure Investments,” *Journal of Environmental Economics and Management*, 120, <https://doi.org/10.1016/j.jeem.2023.10283>.
- MENSAH, J. T. (2024): “Jobs! Electricity shortages and unemployment in Africa,” *Journal of Development Economics*, 167, 103231, <https://www.sciencedirect.com/science/article/pii/S0304387823001876>.

- RAMBACHAN, A., AND J. ROTH (2023): “A More Credible Approach to Parallel Trends,” *The Review of Economic Studies*, 90, 2555–2591, <https://doi.org/10.1093/restud/rdad018>.
- ROTH, J., P. H. SANT’ANNA, A. BILINSKI, AND J. POE (2023): “What’s trending in difference-in-differences? A synthesis of the recent econometrics literature,” *Journal of Econometrics*, 235, 2218–2244, <https://doi.org/10.1016/j.jeconom.2023.03.008>.
- SAEZ, E., J. SLEMROD, AND S. H. GIERT (2012): “The elasticity of taxable income with respect to marginal tax rates: a critical review,” *Journal of Economic Literature*, 50(1), 3–50.
- SCOTT, A., E. DARKO, A. LEMMA, AND J.-P. RUD (2014): “How Does Electricity Insecurity Affect Businesses in Low- and Middle-Income Countries?” *London: Overseas Development Institute*.
- SEARS, J. (2022): “Culpable Consumption: Residential Response to Price and Non-Price Water Conservation Measures,” *Working Paper*.
- SUN, L., AND S. ABRAHAM (2021): “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects,” *Journal of Econometrics*, 225, 175–199, <https://doi.org/10.1016/j.jeconom.2020.09.006>.
- THE WORLD BANK (2022): “World Bank Enterprise Surveys,” <http://www.enterprisesurveys.org>.
- WICHMAN, C. (2014): “Perceived price in residential water demand: evidence from a natural experiment,” *Journal of Economic Behavior & Organization*, 107, 308–323.
- ZHU, X., L. LI, K. ZHOU, X. ZHANG, AND S. YANG (2018): “A meta-analysis on the price elasticity and income elasticity of residential electricity demand,” *Journal of Cleaner Production*, 201, 169–177, <https://doi.org/10.1016/j.jclepro.2018.08.027>.

## 7 Additional Figures

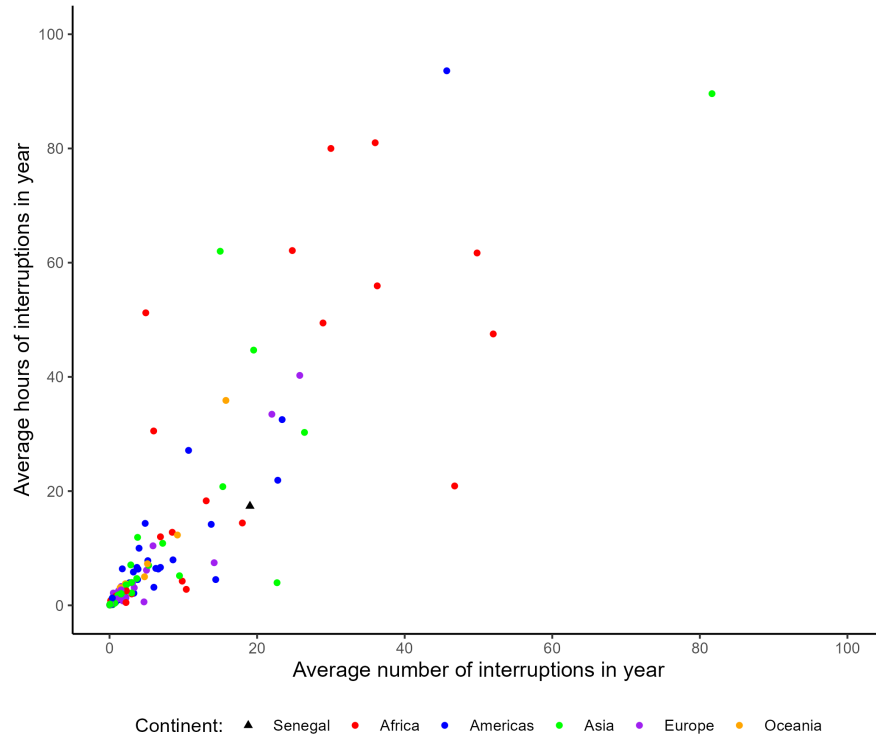
In this section, I present additional figures related to the analysis.

Figure 8: Map of Electricity Distribution Network in Senegal



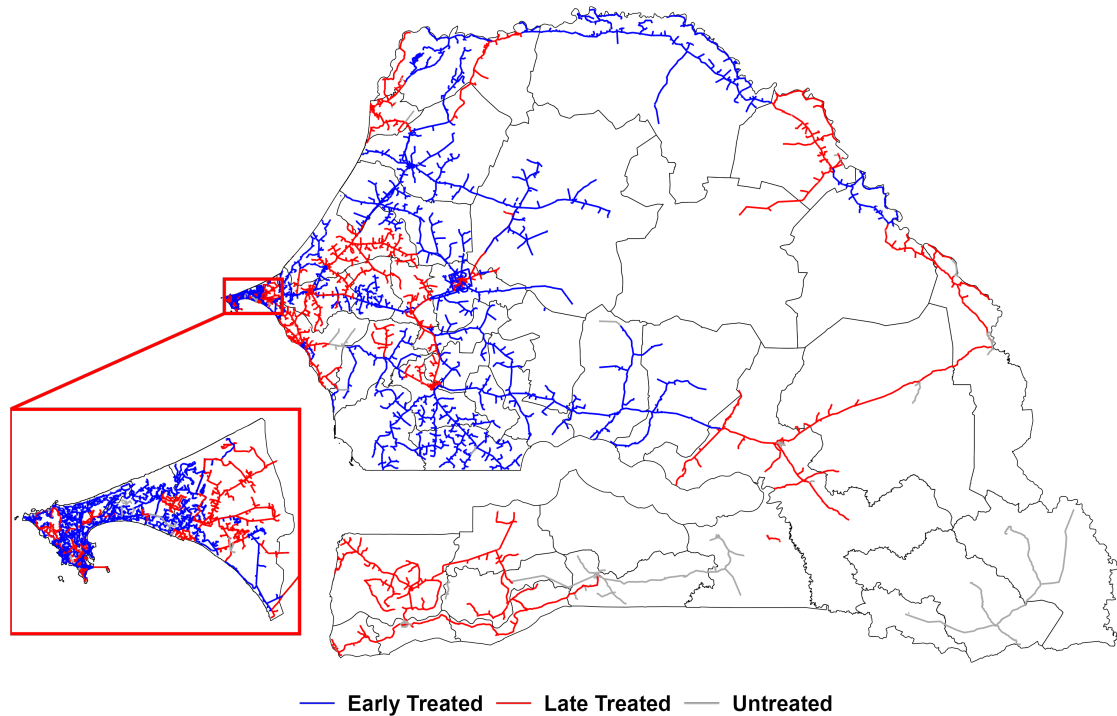
Notes: This map shows the distribution lines across Senegal along with the transformer stations located in those lines and serving end customers. Distribution stations are represented by dots and feeder lines are represented by lines. Each color corresponds to a different feeder line. The inset map in the bottom left corner shows the electricity distribution network of Dakar. Data corresponds to the geocoded map of network infrastructure done as of March 2022.

Figure 9: Duration and Frequency of Outages across Countries in 2019



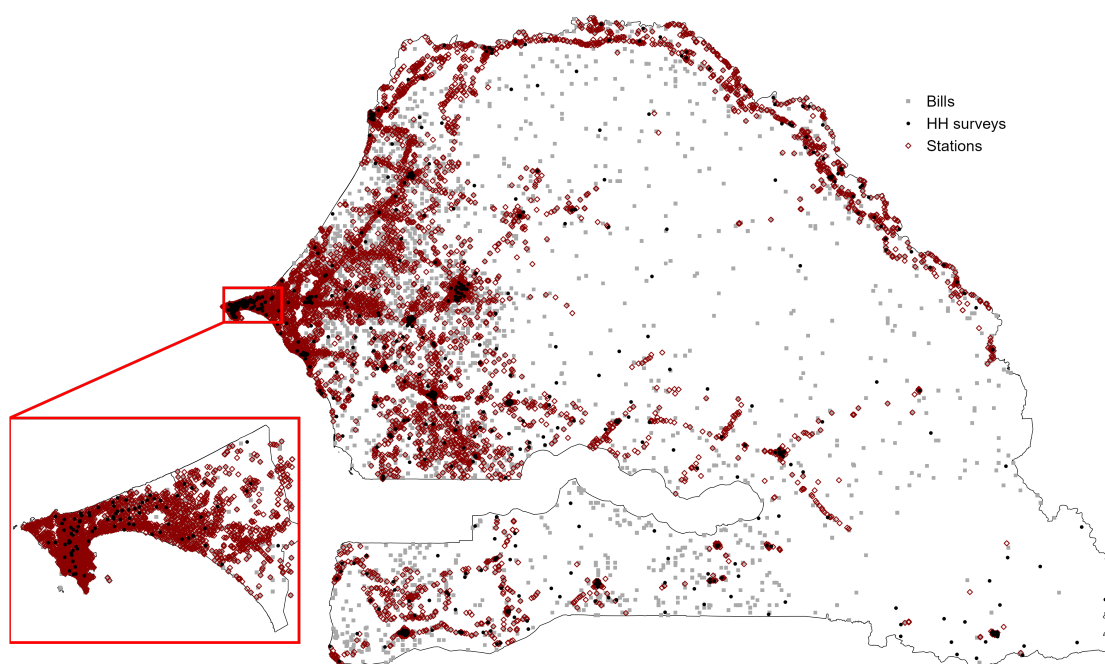
Notes: This figure shows duration and frequency of outages by country in 2019. The x-axis shows the System Average Interruption Duration Index (SAIDI), which is the total minutes of non-instantaneous electric interruptions, per year, customers experienced on average in 2019. The y-axis shows the System Average Interruption Frequency Index (SAIFI), which is the number of non-instantaneous electric interruptions, per year, customers experienced on average in 2019. Each point corresponds to a country for which the data is available. Each color corresponds to a continent (see legend). The dot for Senegal is shown in black. Data are sourced from World Bank's Doing Business Report. Each country uses a different minimum thresholds below which outages are not accounted for in the calculation of the indices. Index values above 100 are not reported in the raw data used for the construction of this figure.

Figure 10: Map of Feeder Lines by Timing of Reliability Project in Senegal



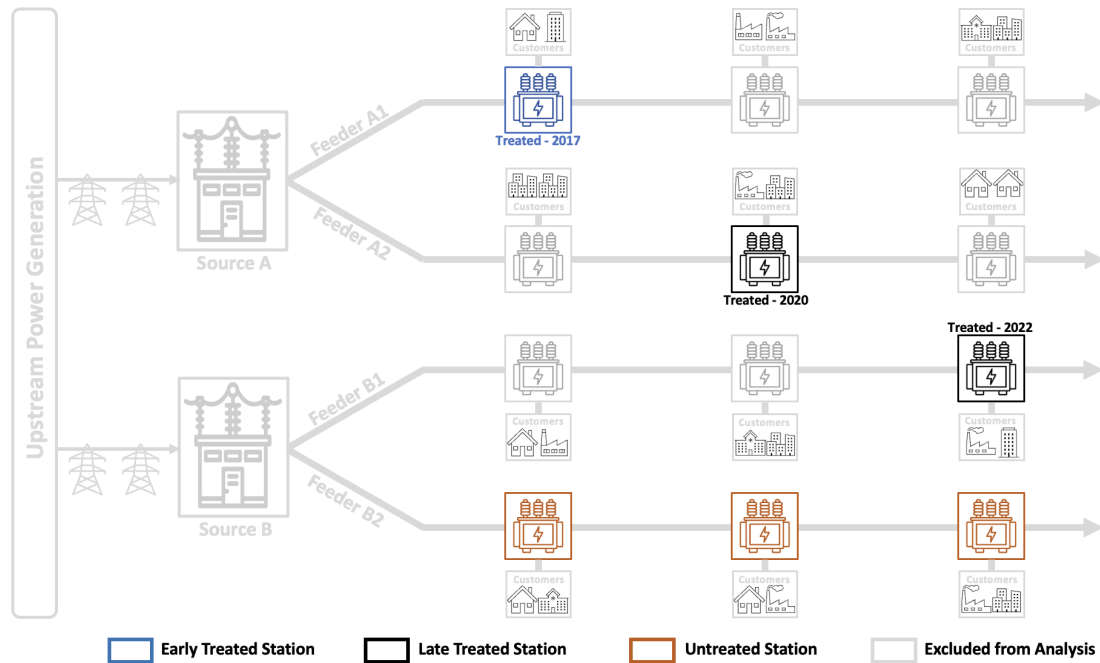
Notes: This map shows the distribution lines across Senegal by timing of the reliability project. Blue lines are treated early (2015) while red lines are treated late (2023). Lines in blue are feeder lines treated before or in 2018 while lines in red are feeder lines treated after 2018. Lines in grey are untreated feeder lines. The inset map in the bottom left corner shows the electricity distribution network of Dakar. Data corresponds to the geocoded map of network infrastructure done as of March 2022.

Figure 11: Geolocation of Observations in Datasets



Notes: This map shows the enumeration areas of households selected in the EHCVM household surveys (in black), the distribution stations in the electricity distribution network (in red), the customer locations in the clients billing data (in grey). All objects are shown as dots on the map. The background map is a map of Senegal. The inset map zooms in on the capital city of Dakar.

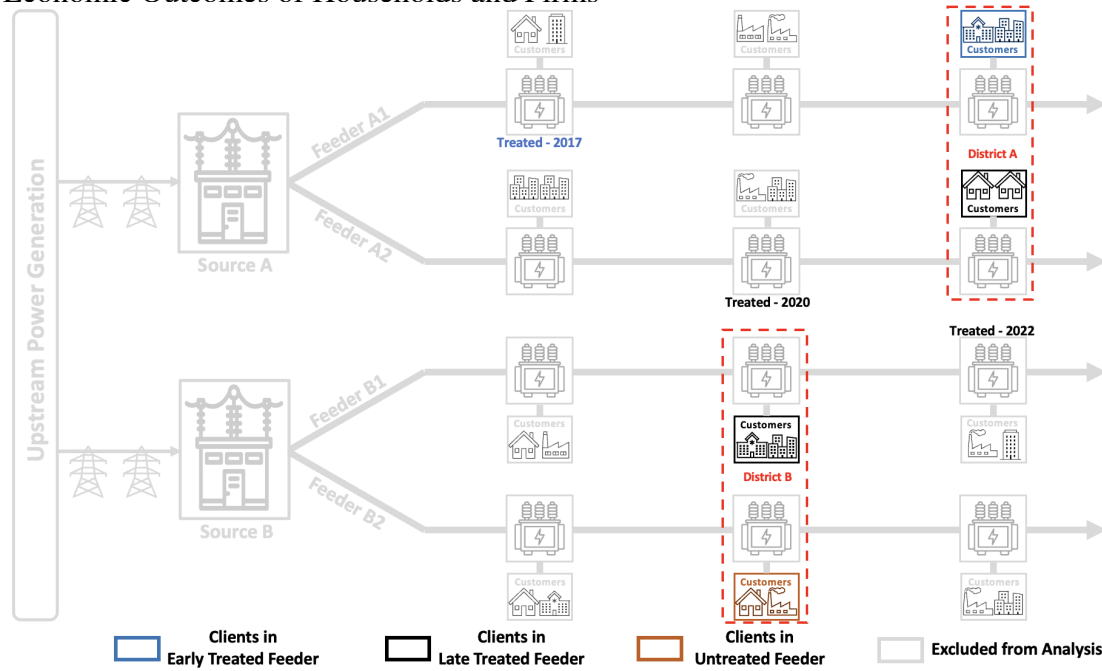
Figure 12: Illustration of Variation Used to Estimate Direct Effects of Reliability Projects



Notes: This figure shows a stylized map of the electricity distribution network in Senegal. Each line corresponds to a feeder line and each box to a distribution station (or transformer station). Stations in blue are stations treated early (2017 in this case). Stations in black are stations treated late (2020 and 2023 in this case). Stations in red are stations on feeder lines with no treated stations. Stations in grey are untreated stations on feeder lines with at least one treated station. These stations are excluded from the estimation of regression (1).

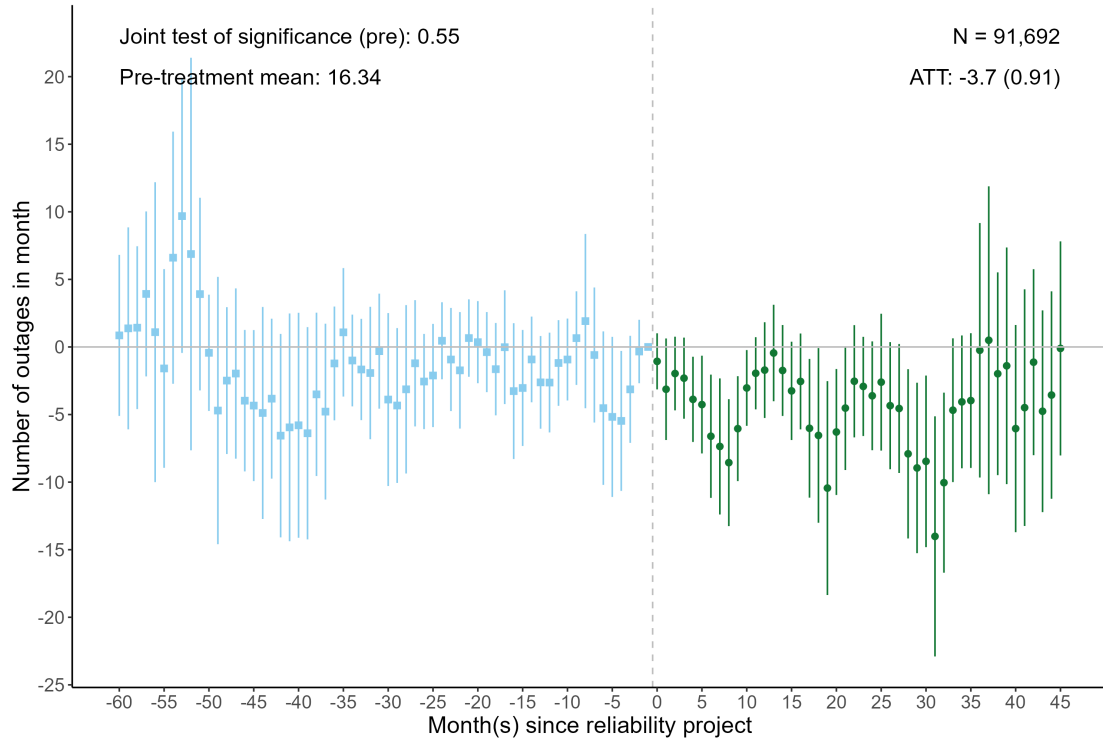


Figure 14: Illustration of Variation Used to Estimate Impact of Reliability Projects on Economic Outcomes of Households and Firms



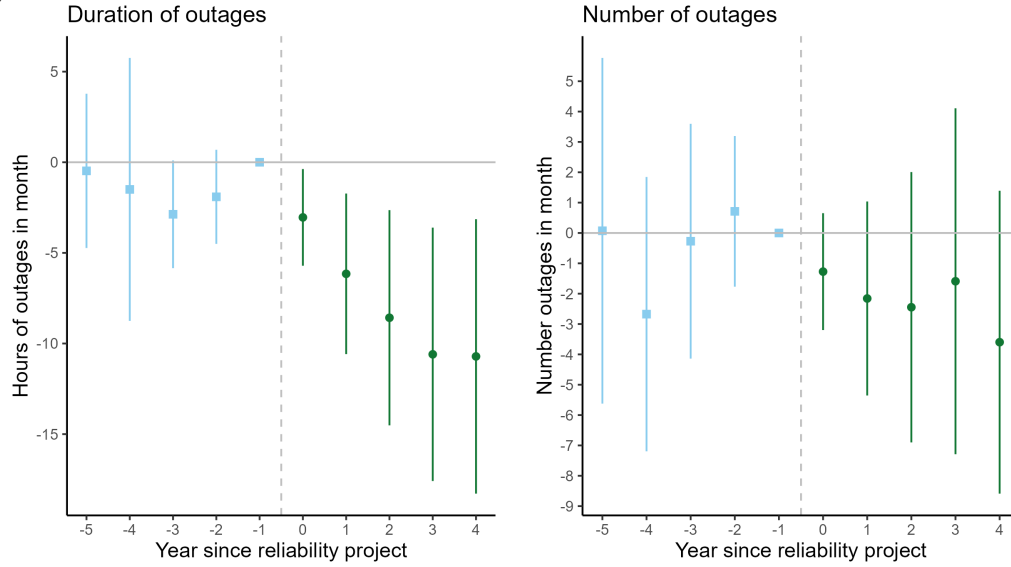
Notes: This figure shows a stylized map of the electricity distribution network in Senegal. Each line corresponds to a feeder line and each box to a distribution station (or transformer station). Stations in blue are untreated stations (i.e., they did not receive the reliability devices) on lines with a station treated early. Stations in black are untreated stations on lines with a station treated late. Stations in red are stations on feeder lines with no treated stations.

**Figure 15: Impact of Reliability Projects on Frequency of Outages**  
**Frequency of outages**



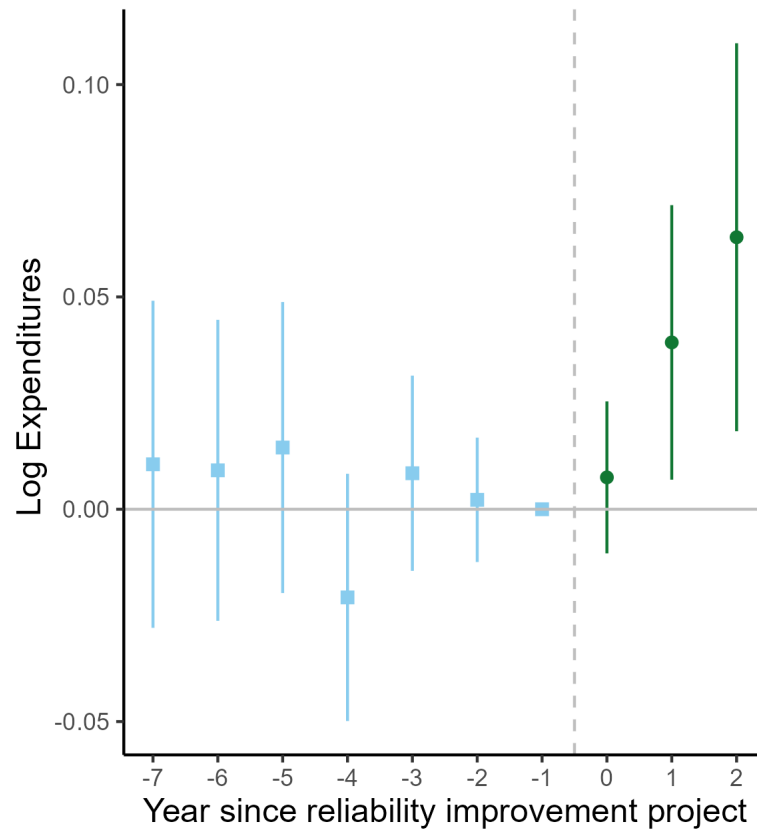
Notes: This figure shows the coefficients  $\lambda_t$  in regression (1) and their corresponding confidence intervals, estimated following Callaway and Sant'Anna (2021). All coefficients are estimates relative to a baseline value for the month preceding project completion date. The mean monthly number of outages in the baseline period is 16.34. The y-axis shows the frequency of outages in terms of number of outages in month. The x-axis shows the number of months since the completion of the reliability project. The vertical dashed grey line corresponds to the month of project date. Data used for estimation is a balanced sample of treated stations with outage observations for all time periods of the analysis, from 5 years before project date to 4 years after project date. The estimation is done at the distribution station level.

Figure 16: Impact of Reliability Projects on Yearly Duration and Frequency of Outages



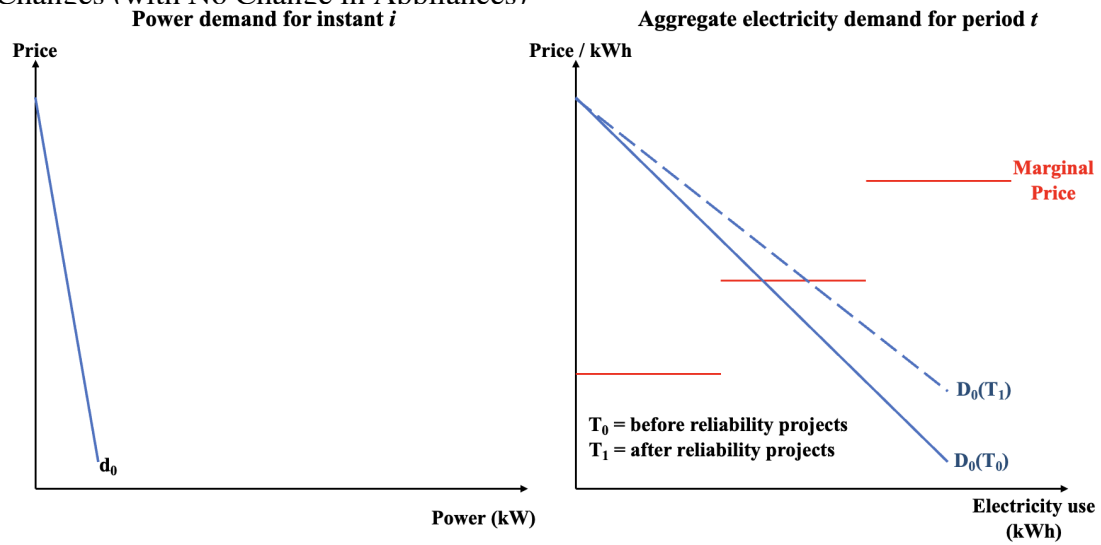
Notes: This figure shows the coefficients  $\lambda_t$  in regression (1) and their corresponding confidence intervals, estimated using Callaway and Sant'Anna (2021). All coefficients are estimates relative to a baseline value for the year preceding project date. The mean monthly hours of outages in the baseline period is 21.11 hours. The mean monthly number of outages in the baseline period is 19.64. The left-hand side graph shows the duration of outages in terms of total hours of outages in month. The right-hand side graph shows the frequency of outages in terms of number of outages in month. For both graphs, the x-axis shows the number of years since the completion of the reliability project. The vertical dashed grey line corresponds to the year of project date. Data used for estimation is a balanced sample of treated stations with outage observations for all time periods of the analysis, from 5 years before project date to 4 years after project date. The estimation is done at the distribution station level.

Figure 17: Impact of Reliability Projects on Electricity Expenditures  
Electricity expenditures



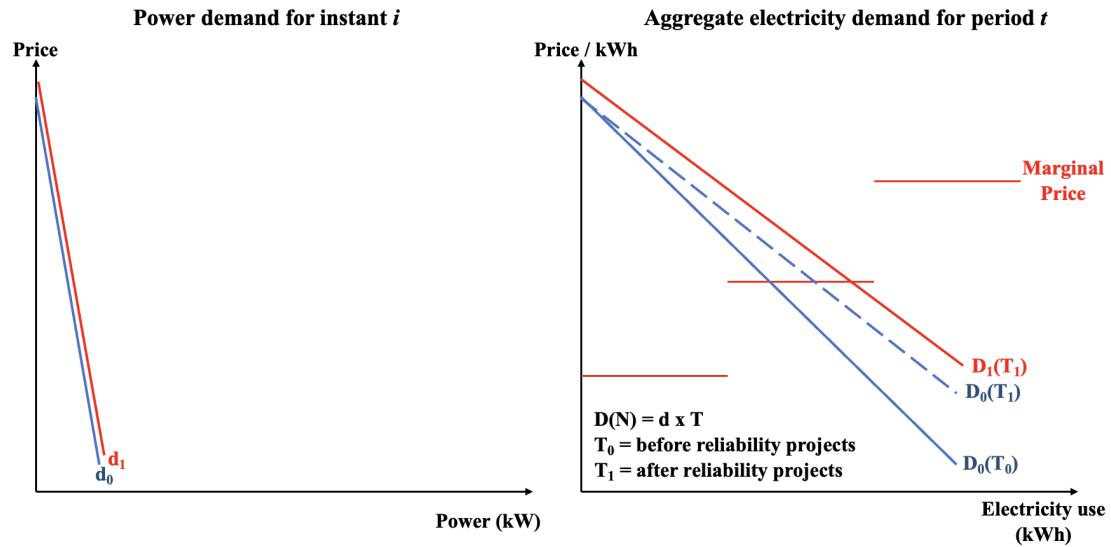
Notes: This figure shows the effects of the reliability projects on electricity expenditures, with the coefficients  $\lambda_l$  in regression (3) and their corresponding confidence intervals. All coefficients are estimates relative to a baseline value for the year preceding project date. The y-axis shows the mean log of electricity expenditures (in CFA Franc). The x-axis shows the number of years that has passed since the completion of the reliability project. The vertical dashed grey line corresponds to the year of project date. Data used for estimation is the geocoded billing data, where each observation is a customer of the utility company. The estimation is done at the client level.

Figure 18: Visual Representation of the Change in Electricity Use following Reliability Changes (with No Change in Appliances)



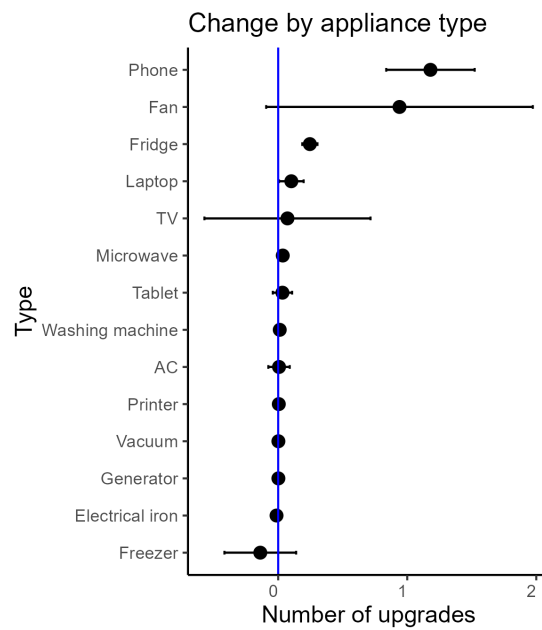
Notes: This figure illustrates the different channels through which electricity use could increase as a result of a change in reliability. The left panel shows the demand curve for the consumer for instant  $i$ . Instants can be considered to represent minutes, hours or days. For each instant, the consumer is initially facing the blue demand curve  $d_0$  before the reliability projects. Assuming that usage is the same across all instants of a given time period  $t$  (e.g. month or year), the horizontal aggregation of instant demand at each price level across all instant yields the initial demand curve before the reliability project. This initial demand curve is shown in solid blue as  $D_0(T_0)$  on the right panel, and is simply the product of the number of instants in the time period ( $T$ ) and the individual quantity demanded per instant, for each price point. Once the reliability projects occur, the number of instant with grid electricity increases from  $T_0$  to  $T_1$ . Assuming that instant demand stays the same at  $d_0$  post the reliability changes, this increase in reliability leads to an higher demand curve for the period, represented by the dashed blue line  $D_0(T_1)$ , coming solely from the increase in the number of instants when electricity is on.

Figure 19: Visual Representation of the Change in Electricity Use following Reliability Changes (with a Change in Appliances)



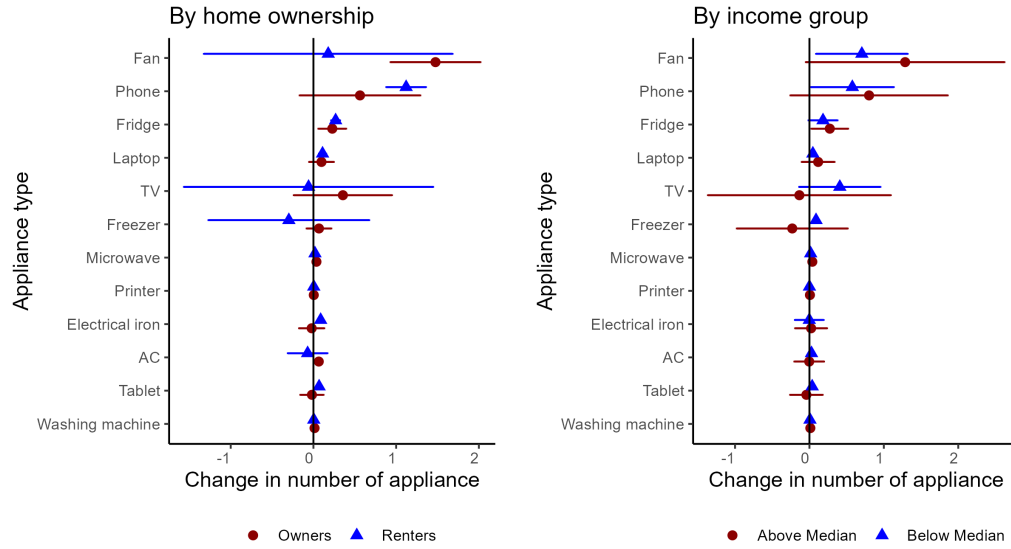
Notes: This figure illustrates the different channels through which electricity use could increase as a result of a change in reliability. The left panel shows the demand curve for the consumer for instant  $i$ . Instants can be considered to represent minutes, hours or days. For each instant, the consumer is initially facing the blue demand curve  $d_0$  before the reliability projects. Assuming that usage is the same across all instants of a given time period  $t$  (e.g. month or year), the horizontal aggregation of instant demand at each price level across all instant yields the initial demand curve before the reliability project. This initial demand curve is shown in solid blue as  $D_0(T_0)$  on the right panel, and is simply the product of the number of instants in the time period ( $T$ ) and the individual quantity demanded per instant, for each price point. Once the reliability projects occur, the number of instant with grid electricity increases from  $T_0$  to  $T_1$ . Assuming that instant demand shifts outwards from  $d_0$  to  $d_1$  post the reliability changes (due to an increase in the number of appliances for example), this increase in reliability leads to an outward shift in aggregate demand from  $D_0(T_1)$  (the aggregate demand with no change in instant demand) to  $D_1(T_1)$  (the aggregate demand with a change in instant demand).

Figure 20: Heterogeneity in the Impact of Reliability Projects on Appliance Adoption - By Appliance Type



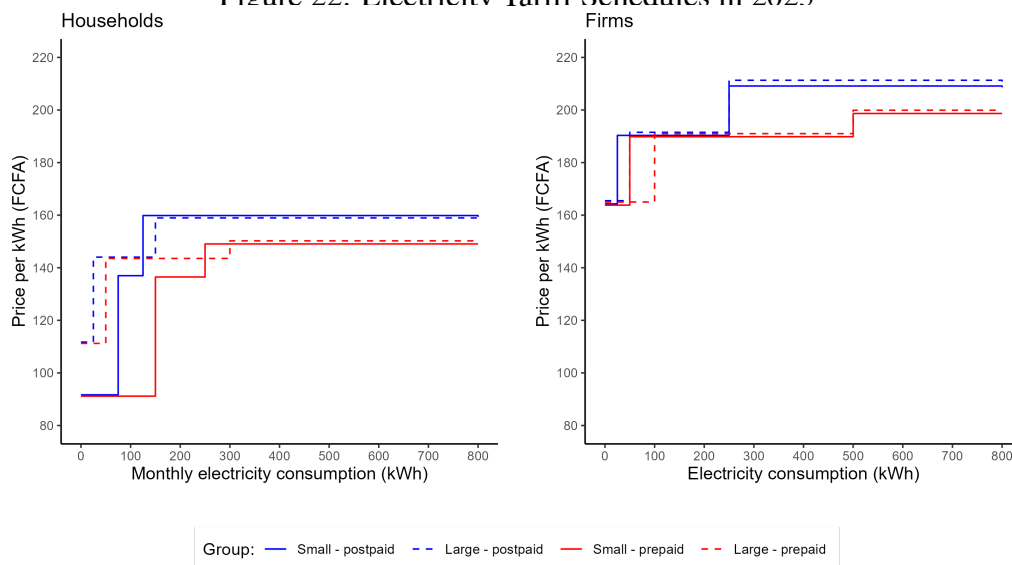
Notes: This figure shows the coefficients  $\lambda_l$  in regression (3) 4 years after project date. All coefficients are estimates relative to a baseline value for the year preceding project date. The x-axis shows the number of appliances. The vertical blue line corresponds to 0. Data used for estimation is a balanced sample of treated households with non-missing values for all time periods of the analysis, from 6 years before project date to 4 years after project date. The estimation is done at the household level.

Figure 21: Heterogeneity in the Impact of Reliability Projects on Electric Appliance Adoption



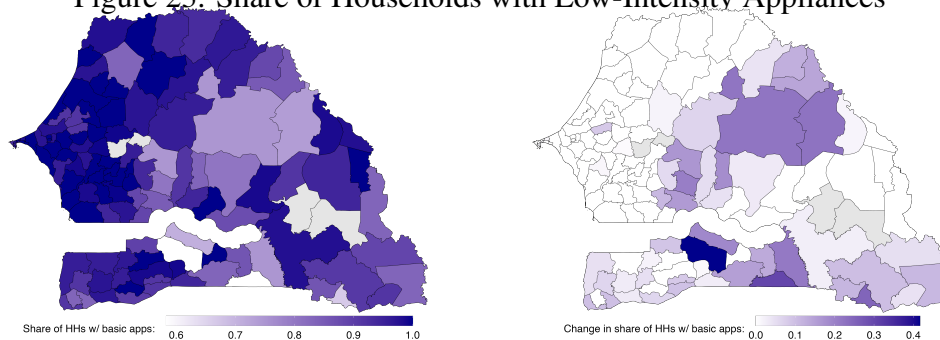
Notes: This figure shows the coefficients  $\lambda_l$  in regression (3) 4 years after project date, separately by home ownership status (left-hand side) and by income brackets (right-hand side). All coefficients are estimates relative to a baseline value for the year preceding project date. The left-hand side graph shows appliance adoption for home-owners and renters while the right hand side show appliance adoption for households above and below median income. For both graphs, the x-axis shows the number of appliances. The vertical black line corresponds to 0. Data used for estimation is a balanced sample of treated households with non-missing values for all time periods of the analysis, from 6 years before project date to 4 years after project date. The estimation is done at the household level.

Figure 22: Electricity Tariff Schedules in 2023



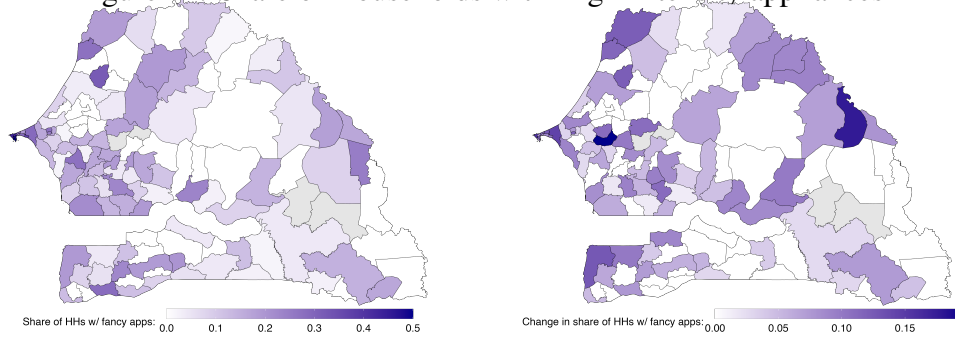
Notes: This figure displays the tariff schedules for different groups of clients in 2023. The left hand side graph shows the tariffs for residential clients and the right hand side graph shows the tariffs for commercial clients. The x-axis shows the monthly electricity usage (in kWh) while the y-axis shows the marginal price per kWh. Line colors are as follows: solid blue for small residential clients using postpaid meters, dashed blue for large residential clients using postpaid meters, solid red for small residential clients using prepaid meters, dashed red for large residential clients using prepaid meters. Small clients are those who receive a voltage level less than 1000 volts (based on their baseline number of appliances). Large clients are those that receive voltage levels between 1000 and 60 000 volts. Residential clients are clients whose main use of electricity is in a residential unit and firms are clients whose main use of electricity is in a commercial establishment. The unit prices are given in the local currency of CFA Franc, which is pegged to the Euro. 1 US dollar is approximately 550 CFA Franc.

Figure 23: Share of Households with Low-Intensity Appliances



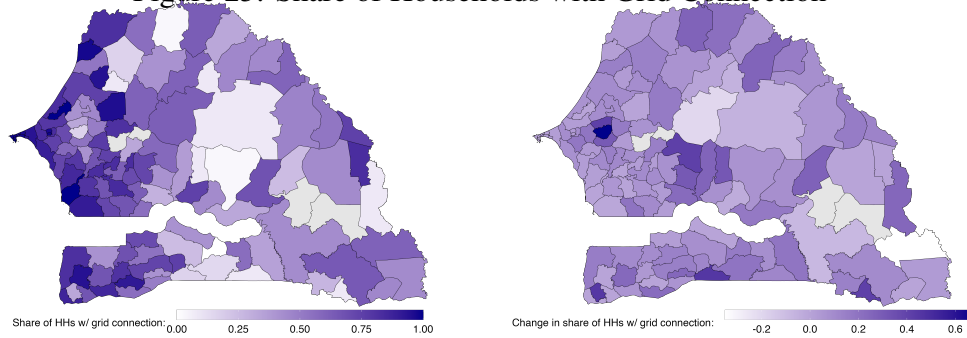
Notes: This map displays the share of households with high-intensity appliances in each administrative unit (level 3) of Senegal. low-intensity appliances include: light bulbs, mobile phones, TVs, refrigerators, and fans. The left hand side map shows the share of households with high-intensity appliances in 2018 and the right hand side map shows the change in the share of households with high-intensity appliances between 2018 and 2022. The color white corresponds to low values of the variable and the color dark blue corresponds to high values of the variable. Only households surveyed in both rounds of the EHCVM dataset are included in the analysis (N = 5, 824 households). See Section 3 for a description of the dataset. Grey areas are areas with no data available.

**Figure 24: Share of Households with High-Intensity appliances**



Notes: This map displays the share of households with high-intensity appliances in each administrative unit (level 3) of Senegal. high-intensity appliances include: air conditioners, freezers, irons, generators, laptops, microwaves, mixers, printers, tablets, vacuums, washing machines. The left hand side map shows the share of households with high-intensity appliances in 2018 and the right hand side map shows the change in the share of households with high-intensity appliances between 2018 and 2022. The color white corresponds to low values of the variable and the color dark blue corresponds to high values of the variable. Only households surveyed in both rounds of the EHCVM dataset are included in the analysis (N = 5, 824 households). See Section 3 for a description of the dataset. Grey areas are areas with no data available.

**Figure 25: Share of Households with Grid Connection**



Notes: This map displays the share of households with high-intensity appliances in each administrative unit (level 3) of Senegal. The left hand side map shows the share of households connected to the grid in 2018 and the right hand side map shows the change in the share of households connected to the grid between 2018 and 2022. The color white corresponds to low values of the variable and the color dark blue corresponds to high values of the variable. Only households surveyed in both rounds of the EHCVM dataset are included in the analysis (N = 5, 824 households). See Section 3 for a description of the dataset. Grey areas are areas with no data available.

## 8 Additional Tables

In this section, I present additional tables related to the analysis.

Table 6: Types of Outages

|                                           | Share of:          |                 |
|-------------------------------------------|--------------------|-----------------|
|                                           | Occurrences<br>(1) | Duration<br>(2) |
| <b>Planned outages</b>                    |                    |                 |
| Maintenance work on network               | 17.81              | 66.12           |
| Test on default                           | 4.09               | 1.52            |
| Shortage of production                    | 1.53               | 0.50            |
| Circumstantial                            | 0.39               | 0.30            |
| Maintenance work due to wildfire          | 0.36               | 0.31            |
| Passage of convoy of cars                 | 0.35               | 1.06            |
| <b>Unplanned outages</b>                  |                    |                 |
| Network incident                          | 8.83               | 9.30            |
| Default or damages on line equipment      | 8.65               | 7.50            |
| Production incident                       | 7.51               | 0.68            |
| Unknown / not investigated                | 5.53               | 2.18            |
| Failure of protection equipment or system | 3.05               | 0.54            |
| Low voltage of network                    | 2.55               | 1.46            |
| Accidental shock on lines                 | 0.69               | 0.21            |
| Foreign body on line                      | 0.59               | 0.36            |
| Rain or flood                             | 0.59               | 0.94            |
| Wind, storm or lighting                   | 0.51               | 0.77            |
| Other reasons                             | 0.15               | 0.13            |
| Pollution, fog or humidity                | 0.13               | 0.09            |
| Network overload                          | 0.09               | 0.11            |
| Malicious act on line                     | 0.03               | 0.01            |

Notes: This table shows the causes of interruptions in the data on outage logs from 2016 to 2023. The first column lists the different causes of outages present in the data. Column (1) gives the share of total number of outages that are due to the cause listed in the first column. Column (2) gives the share of total duration of outages that are due to the cause listed in the first column.

Table 7: Frequency and Duration of Outages

| Year | SAIFI | SAIDI |
|------|-------|-------|
| 2022 | 15.80 | 860   |
| 2021 | 11.69 | 455   |
| 2020 | 6.34  | 290   |
| 2019 | 11.11 | 605   |
| 2018 | 19.03 | 1042  |

Notes: This table shows the System Average Interruption Duration Index (SAIDI), which is the total minutes of non-instantary electric interruptions, per year, the average customer experienced, and the System Average Interruption Frequency Index (SAIFI), which is the number of non-instantary electric interruptions, per year, the average customer experienced. Data are sourced from SENELEC's annual reports. The indices only include outages that last more than 5 minutes in duration.

Table 8: Comparison of Distribution Stations By Treatment Timing

| Table 8: Comparison of Distribution Stations By Treatment Timing |               |         |              |         |         |
|------------------------------------------------------------------|---------------|---------|--------------|---------|---------|
|                                                                  | Early Treated |         | Late Treated |         | P-Value |
|                                                                  | Mean          | SD      | Mean         | SD      |         |
|                                                                  | (1)           | (2)     | (3)          | (4)     | (5)     |
| <b>Electricity</b>                                               |               |         |              |         |         |
| Connected to the grid (0/1)                                      | 0.72          | 0.45    | 0.71         | 0.45    | 0.42    |
| Monthly electricity expenditure ('000s FCFA)                     | 11.47         | 10.85   | 10.83        | 10.49   | 0.20    |
| Number of outages in the last 7 days                             | 4.53          | 5.26    | 5.06         | 5.69    | 0.12    |
| <b>Expenditures and assets</b>                                   |               |         |              |         |         |
| Total number of appliances                                       | 0.38          | 0.91    | 0.38         | 1.01    | 0.82    |
| Consumption expenditures (000s FCFA)                             | 1814.99       | 1082.73 | 1810.69      | 1189.63 | 0.92    |
| <b>HH Characteristics</b>                                        |               |         |              |         |         |
| Household size                                                   | 9.37          | 5.78    | 9.55         | 6.21    | 0.43    |
| Has a family enterprise (0/1)                                    | 0.16          | 0.37    | 0.17         | 0.38    | 0.71    |
| Owens establishment (0/1)                                        | 0.74          | 0.44    | 0.78         | 0.41    | 0.01    |
| Establishment is a house (0/1)                                   | 0.90          | 0.30    | 0.90         | 0.31    | 0.62    |
| Obs                                                              | 2224          | –       | 1169         | –       | –       |

Notes: This table shows summary statistics for households served by distribution stations included in the reliability projects between 2015 and 2018 (columns 1 and 2) relative to households served by distribution stations that benefitted from the reliability projects between 2019 and 2023 (columns 3 and 4). The “P-Value” gives the p-value associated with the coefficient on the regression of each variable on the leftmost column on a dummy for being selected early for the projects. Data are the households included in the EHCVM described in Section ?? of the main text.

| <u>Table 9: Correlates of Timing of Reliability Projects</u> |                      |                     |
|--------------------------------------------------------------|----------------------|---------------------|
|                                                              | Year of project      |                     |
|                                                              | (1)                  | (2)                 |
| Number of sub-lines                                          | 0.16***<br>(0.0517)  | 0.34***<br>(0.0882) |
| Total outages (hours)                                        | -0.02***<br>(0.0019) | -0.01<br>(0.013)    |
| Station is priority (0/1)                                    | -0.3***<br>(0.0615)  | -0.1<br>(0.2359)    |
| Distance to utility offices (km)                             | 0.02***<br>(0.0017)  | -0.01<br>(0.0052)   |
| Population in 2015 ('000s)                                   | 0.004<br>(0.0052)    | 0.04<br>(0.0277)    |
| Station is public (0/1)                                      | 0.22***<br>(0.0628)  | 0.05<br>(0.1815)    |
| Voltage level (kV)                                           | -0.07***<br>(0.0067) | -0.05<br>(0.0284)   |
| District FEs                                                 | N                    | Y                   |
| DV Mean                                                      | 2019.95              | 2019.95             |
| Number of stations                                           | 2315                 | 2315                |
| $R^2$                                                        | 0.13                 | 0.292               |

Notes: This table shows the results of a regression of the year in which stations received reliability improvement devices on characteristics of the distribution stations. Column (1) is run without district fixed-effects and column (2) is run with district fixed effects. Each observation is a distribution station.

| Table 10: Ownership of Electrical Appliances over Time |      |      |      |      |      |      |      |      |
|--------------------------------------------------------|------|------|------|------|------|------|------|------|
|                                                        | Year |      |      |      |      |      |      |      |
|                                                        | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 |
| <i>Panel A: Low-intensity Appliances</i>               |      |      |      |      |      |      |      |      |
| TV                                                     | 22.5 | 31.0 | 39.7 | 48.2 | 54.3 | 57.0 | 58.6 | 59.2 |
| Fridge                                                 | 9.7  | 13.2 | 16.5 | 19.6 | 22.6 | 24.8 | 26.5 | 26.9 |
| Mobile phone                                           | 68.0 | 85.8 | 89.1 | 91.4 | 94.4 | 95.0 | 96.2 | 97.4 |
| Fan                                                    | 8.0  | 14.2 | 22.3 | 31.2 | 36.9 | 42.0 | 46.7 | 48.0 |
| Laptop                                                 | 2.9  | 4.5  | 6.2  | 7.8  | 9.1  | 9.8  | 10.3 | 10.4 |
| Tablet                                                 | 0.6  | 1.4  | 2.5  | 3.8  | 4.5  | 5.1  | 5.8  | 6.1  |
| <i>Panel B: High-Intensity Appliances</i>              |      |      |      |      |      |      |      |      |
| Air conditioner                                        | 0.7  | 0.9  | 1.2  | 1.4  | 1.8  | 2.2  | 2.4  | 2.4  |
| Freezer                                                | 3.3  | 4.5  | 5.8  | 6.9  | 8.4  | 9.9  | 10.8 | 11.2 |
| Iron                                                   | 0.9  | 1.3  | 1.8  | 2.4  | 2.8  | 3.6  | 3.8  | 3.9  |
| Generator                                              | 0.1  | 0.1  | 0.1  | 0.1  | 0.2  | 0.2  | 0.2  | 0.2  |
| Microwave                                              | 0.2  | 0.3  | 0.5  | 0.6  | 0.8  | 1.0  | 1.2  | 1.3  |
| Mixer                                                  | 0.1  | 0.3  | 0.3  | 0.4  | 0.6  | 0.9  | 1.2  | 1.2  |
| Printer                                                | 0.2  | 0.3  | 0.3  | 0.4  | 0.5  | 0.5  | 0.6  | 0.6  |
| Vacuum                                                 | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  |
| Washing machine                                        | 0.1  | 0.2  | 0.2  | 0.3  | 0.3  | 0.4  | 0.5  | 0.6  |

Note: This table reports the share of households with a given electrical appliance type between 2015 and 2023, for the sample of respondents in the EHCVM dataset described in section 3. The time series of ownership of a given appliance X is constructed based on two survey questions in the EHCVM: (1) Do you own appliance X? if so, since when do you own appliance X? For appliances with more than one unit of ownership, only the last year of purchase is reported, so the values reported on the table represent the share of households who owns a given appliance with no replacement or upgrade of that asset for the stated year. Only households included in both rounds of the EHCVM surveys (2018/2019 and 2021/2022) are used to compute the statistics (N = 5,864 households).

| Table 11: Ownership of Non-Electrical Assets over Time |      |      |      |      |      |      |      |      |
|--------------------------------------------------------|------|------|------|------|------|------|------|------|
|                                                        | Year |      |      |      |      |      |      |      |
|                                                        | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 |
| Bed                                                    | 51.6 | 65.3 | 76.6 | 85.4 | 90.1 | 91.1 | 91.5 | 91.7 |
| Bicycle                                                | 13.3 | 15.6 | 17.6 | 19.7 | 20.6 | 21.4 | 22.2 | 23.0 |
| Car                                                    | 1.6  | 2.0  | 2.5  | 3.1  | 3.5  | 3.9  | 4.3  | 4.4  |
| Carpet                                                 | 6.6  | 11.7 | 20.2 | 31.4 | 38.4 | 42.2 | 45.8 | 47.1 |
| Closet                                                 | 28.6 | 35.1 | 40.4 | 45.0 | 48.0 | 49.8 | 51.1 | 51.3 |
| Gaz                                                    | 35.0 | 40.8 | 45.5 | 48.2 | 50.8 | 52.8 | 53.7 | 54.1 |
| House                                                  | 5.7  | 5.9  | 6.1  | 6.2  | 6.3  | 6.4  | 6.5  | 6.5  |
| Land                                                   | 5.8  | 6.7  | 7.9  | 8.6  | 9.7  | 10.5 | 11.0 | 11.1 |
| Mattress                                               | 37.9 | 54.5 | 72.7 | 88.0 | 96.2 | 96.9 | 97.3 | 97.4 |
| Non-electrical iron                                    | 25.5 | 32.6 | 39.2 | 42.8 | 46.0 | 49.0 | 50.4 | 51.0 |
| Non-electrical mixer                                   | 0.3  | 0.4  | 0.7  | 0.9  | 1.3  | 2.1  | 2.8  | 3.1  |
| Radio                                                  | 11.8 | 18.7 | 28.2 | 37.8 | 43.9 | 47.5 | 50.1 | 51.0 |
| Sofa                                                   | 13.1 | 15.0 | 17.0 | 19.0 | 20.6 | 21.9 | 22.8 | 23.1 |
| Stove                                                  | 5.4  | 8.6  | 12.1 | 15.4 | 18.6 | 22.7 | 25.1 | 26.0 |
| Table                                                  | 4.8  | 6.8  | 8.8  | 10.4 | 11.0 | 11.9 | 12.3 | 12.5 |

Notes: This table reports the share of households with a given non-electrical asset between 2015 and 2023, for the sample of respondents in the EHCVM dataset described in section 3. The time series of ownership of a given asset X is constructed based on two survey questions in the EHCVM: (1) Do you own asset X? if so, since when do you own asset X? For appliances with more than one unit of ownership, only the last year of purchase is reported, so the values reported on the table represent the share of households who owns a given asset with no replacement or upgrade of that asset for the stated year. Only households included in both rounds of the EHCVM surveys (2018/2019 and 2021/2022) are used to compute the statistics (N = 5,864 households).

Table 12: ~~Purchase Price of Electrical Appliances Between 2018 and 2022~~

|                                           | <b>Mean</b> | <b>SD</b> | <b>Min</b> | <b>Max</b> |
|-------------------------------------------|-------------|-----------|------------|------------|
| <i>Panel A: Low-intensity Appliances</i>  |             |           |            |            |
| TV                                        | 108.3       | 111.9     | 1.0        | 1333.3     |
| Fridge                                    | 238.5       | 151.6     | 8.3        | 1416.7     |
| Mobile phone                              | 31.1        | 40.9      | 0.8        | 500.0      |
| Fan                                       | 15.9        | 10.6      | 1.1        | 58.3       |
| <i>Panel B: High-Intensity Appliances</i> |             |           |            |            |
| Air conditioner                           | 248.8       | 120.1     | 26.0       | 616.7      |
| Freezer                                   | 298.1       | 180.2     | 20.0       | 1416.7     |
| Iron                                      | 19.5        | 13.9      | 2.9        | 83.3       |
| Generator                                 | 208.3       | 206.7     | 83.3       | 666.7      |
| Laptop                                    | 232.9       | 158.3     | 16.7       | 1091.7     |
| Microwave                                 | 75.6        | 31.7      | 16.7       | 166.7      |
| Mixer                                     | 38.3        | 19.1      | 13.3       | 83.3       |
| Printer                                   | 103.6       | 80.4      | 41.7       | 283.3      |
| Tablet                                    | 85.8        | 114.5     | 4.2        | 1150.0     |
| Vacuum                                    | 110.4       | 97.0      | 33.3       | 250.0      |
| Washing machine                           | 243.4       | 202.3     | 1.4        | 750.0      |

Notes: This table reports different moments of the price of electrical appliances between 2018 and 2022, as reported by respondents in the EHCVM dataset described in section 3. Only households included in both rounds of the EHCVM surveys (2018/2019 and 2021/2022) are used to compute the statistics (N = 5,864 households).

Table 13: Purchase Price of Non-Electrical Assets Between 2018 and 2022

|                      | Mean   | SD      | Min | Max      |
|----------------------|--------|---------|-----|----------|
| Bed                  | 129.3  | 156.8   | 0.4 | 1666.7   |
| Bicycle              | 285.3  | 346.0   | 5.0 | 2500.0   |
| Car                  | 5809.2 | 4675.2  | 7.3 | 25000.0  |
| Carpet               | 12.6   | 16.6    | 0.2 | 250.0    |
| Closet               | 193.5  | 167.6   | 1.4 | 1083.3   |
| Gaz                  | 13.3   | 7.5     | 0.4 | 33.3     |
| House                | 8224.8 | 10490.3 | 8.3 | 33333.3  |
| Land                 | 2591.4 | 9450.5  | 2.1 | 115000.0 |
| Mattress             | 30.2   | 34.0    | 0.2 | 266.7    |
| Non-electrical iron  | 4.9    | 2.4     | 1.0 | 16.7     |
| Non-electrical mixer | 25.5   | 18.2    | 1.7 | 79.2     |
| Radio                | 12.5   | 15.9    | 0.4 | 250.0    |
| Sofa                 | 390.9  | 297.4   | 8.3 | 2166.7   |
| Stove                | 10.6   | 7.1     | 0.3 | 58.3     |
| Table                | 41.1   | 74.3    | 0.1 | 500.0    |

Notes: This table reports different moments of the price of non-electrical assets between 2018 and 2022, as reported by respondents in the EHCVM dataset described in section 3. Only households included in both rounds of the EHCVM surveys (2018/2019 and 2021/2022) are used to compute the statistics (N = 5,864 households).

Table 14: Correlation between Appliance Ownership

|                 | Connection | Fan  | TV   | Phone | Fridge | Iron | Microwave | Mixer | Freezer | Washing | Vacuum | AC   | Generator | Tablet | Laptop | Printer |
|-----------------|------------|------|------|-------|--------|------|-----------|-------|---------|---------|--------|------|-----------|--------|--------|---------|
| Connection      | 1.00       |      |      |       |        |      |           |       |         |         |        |      |           |        |        |         |
| Fan             | 0.30       | 1.00 |      |       |        |      |           |       |         |         |        |      |           |        |        |         |
| TV              | 0.40       | 0.56 | 1.00 |       |        |      |           |       |         |         |        |      |           |        |        |         |
| Phone           | 0.06       | 0.29 | 0.32 | 1.00  |        |      |           |       |         |         |        |      |           |        |        |         |
| Fridge          | 0.24       | 0.42 | 0.41 | 0.18  | 1.00   |      |           |       |         |         |        |      |           |        |        |         |
| Iron            | 0.07       | 0.16 | 0.14 | 0.06  | 0.19   | 1.00 |           |       |         |         |        |      |           |        |        |         |
| Microwave       | 0.04       | 0.11 | 0.08 | 0.04  | 0.14   | 0.27 | 1.00      |       |         |         |        |      |           |        |        |         |
| Mixer           | 0.04       | 0.10 | 0.08 | 0.04  | 0.14   | 0.18 | 0.31      | 1.00  |         |         |        |      |           |        |        |         |
| Freezer         | 0.14       | 0.26 | 0.24 | 0.11  | 0.12   | 0.17 | 0.14      | 0.09  | 1.00    |         |        |      |           |        |        |         |
| Washing machine | 0.02       | 0.04 | 0.03 | 0.02  | 0.05   | 0.10 | 0.17      | 0.08  | 0.04    | 1.00    |        |      |           |        |        |         |
| Vacuum          | 0.02       | 0.03 | 0.03 | 0.01  | 0.05   | 0.14 | 0.13      | 0.10  | 0.05    | 0.18    | 1.00   |      |           |        |        |         |
| AC              | 0.07       | 0.13 | 0.12 | 0.05  | 0.19   | 0.24 | 0.29      | 0.17  | 0.19    | 0.13    | 0.16   | 1.00 |           |        |        |         |
| Generator       | 0.02       | 0.03 | 0.03 | 0.00  | 0.04   | 0.05 | 0.01      | 0.02  | 0.06    | 0.00    | 0.04   | 0.05 | 1.00      |        |        |         |
| Tablet          | 0.09       | 0.21 | 0.19 | 0.10  | 0.23   | 0.20 | 0.21      | 0.18  | 0.17    | 0.05    | 0.06   | 0.21 | 0.06      | 1.00   |        |         |
| Laptop          | 0.14       | 0.27 | 0.25 | 0.12  | 0.28   | 0.21 | 0.17      | 0.17  | 0.21    | 0.06    | 0.06   | 0.20 | 0.08      | 0.32   | 1.00   |         |
| Printer         | 0.03       | 0.05 | 0.06 | 0.03  | 0.09   | 0.10 | 0.16      | 0.21  | 0.08    | 0.07    | 0.03   | 0.14 | 0.25      | 0.12   | 0.19   | 1.00    |

Note: This table reports the correlation between appliances. Each row/column represents a dummy variable that takes the value of 1 if the respondent owns the appliance stated in that row/column, and 0 otherwise. Data include the sample of respondents in the EHCVM dataset described in section 3, between 2015 and 2023. Only households included in both rounds of the EHCVM surveys (2018/2019 and 2021/2022) are used to compute the statistics (N = 5,864 households).

## 9 Additional Details on the Data

### 9.1 Data Construction

In this subsection, I provide details on the process through which I assemble the different datasets used for the analysis, clean them and match them with one another.

#### Maps of Electricity Network

The source maps are AutoCAD files that I digitize and turn into Comma Separated Value (CSV) files, with one file per year and with one row per distribution station, along with the variables described in the main text of this paper for that year. I also record other characteristics of the stations (e.g. whether they are public, private or mixed), along with the physical infrastructure of the grid around them (e.g. underground lines, aerial lines, electrical devices nearby, etc.).

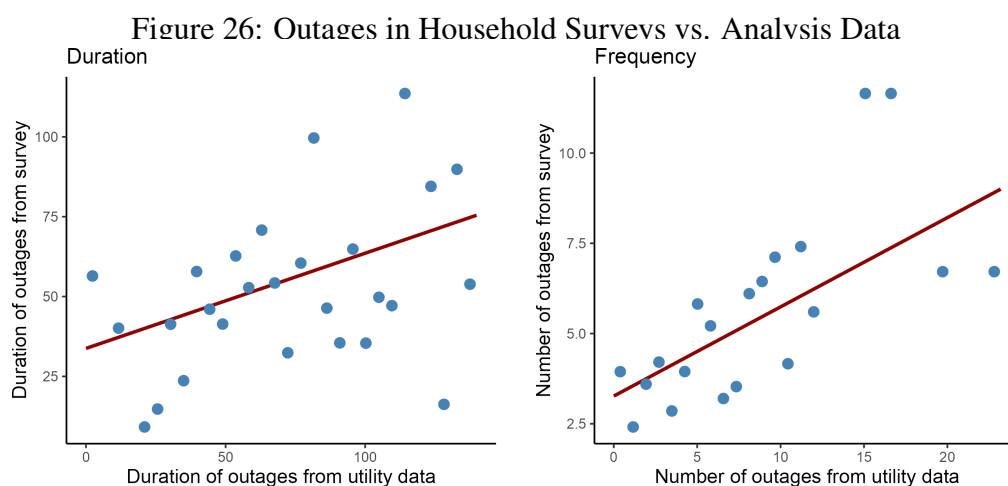
#### Outages

SENELEC uses Supervisory Control and Data Acquisition (SCADA) systems, which are generally used for controlling, monitoring, and analyzing industrial devices and processes. SENELEC uses SCADA systems to get signaled alarms about line interruptions that can be seen by the remote visual operator of the grid network. Whenever there is a signal about a line, the operator puts all relevant information about the interruption in an internal application developed for that matter. The entry records the source station, the feeder line, the distribution stations affected by the interruption, as well as the nature, cause and duration of the interruption. To identify the station affected by the interruptions and those where the default occurs, the operation team follows the internal SCADA signaling systems, and relies on agents sent to the ground to situate the default and determine its extent. There is an operator controlling the system at all hours of the day, even at nights or during holidays to ensure that interruptions during those times are also recorded properly.

There are two issues that could make this collection process incomplete: (i) a break in the information system, and (ii) an incomplete log of outages that occur beyond the level of the sub-stations, at the level of clients (e.g. voltage fluctuations beyond the point of the customer). On point (i), there are sometimes breaks in the information system itself, which causes the alarms related to the interruptions to not be received by the utility. In these cases, the utility company receives the information after the interruptions after the breaks are resolved, which could be hours after the interruptions actually occurred. For the purpose of the analysis, the delay in the logs should not be an issue, so long as the information entered in the delayed logs match the reality on the ground. On point (ii), SENELEC has a 24/7 hotline at the service of customers for them

to report small outages at their level at any time of the day. This service is connected to the outage management team to ensure that they are included in the logs. As with the breaks in the information system, there is often a delay between when these incidents occur and when they are reported. For example, these interruptions that happen post the stations might not be reported on time by customers if they happen while they are asleep, when they have no call network or a bad call network, when they do not know the existence of the hotline to report incidents, or when they simply do not care to report these incidents. In most cases, these interruptions are delayed but end up making it to the log of incidents I use in the analysis. As shown in Table 6, 8.83 percent of incidents relate to these network incidents.

To check whether the data picks up these micro-outages, I compare the recorded outages with outages reported by clients in the EHCVM survey. Figure 26 reports, for the same week, the weekly duration and frequency of outages between the second round of the EHCVM survey conducted in 2021 and the weekly outages from the official data of the utility company. The figure shows a strong positive correlation between the two outages data.



Notes: This map compares the weekly duration and frequency of outages between the surveys (y-axis) and the official data (x-axis) for the year 2021.

I discuss these points to give the reader as much clarity in terms of the outages that are reflected in the data and the potential outages that may be missing from the data. For the analysis, however, the impact of these concerns is relatively minimal. First, I do not directly use outages in any of the regressions but one: the event study figure on the impact of the reliability project on outages. So regardless of the direction and magnitude of the bias that the concerns bring to the outages data, they don't affect any of the tables but that first one. Second, even for the event-study results related to outages, the concerns will only be relevant if somehow the reliability projects increased

the outages missed by the data more for the early-treated stations, relative to the later-treated stations. The conversations with the utility company suggest that, if anything, these small outages would also be reduced by the reliability project, meaning that the estimates would not be lower if these small outages are somehow missing from the data.

## Electricity Usage

The data on the universe of bills has information on the meter ID of all customers during this period, along with their name, their literal address, their phone number, their subscription date, the history on their electricity consumption and purchases, and the size of their account (small vs. medium vs. large).

For customers on postpaid meter, the billing data cover consumption level, bill amount, whether and when they paid those bills, etc. Each row in the billing data shows the meter ID of the customer, their full name, their literal address, their phone number, their subscription date, their electricity consumption for that bill (in kWh), the amount they oI for that bill (in FCFA), the due date of the bill, the size of their unit (small vs. medium vs. large), the type of the customer (residential vs. firm) and the tariff structure they are subject to (block pricing vs. peak load pricing). Each row in the postpaid billing data represents a bill received by a given consumer that is using a postpaid meter between 2012 and 2024. The bill is for consumption over a given period of time.<sup>38</sup> In total, I have approximately 100 million customer-month pairs between 2012 and 2024 across Senegal.

For customers on the prepaid system<sup>39</sup>, the data include their transactions and details pertaining to these transactions (amount, kWh purchased, transaction date, etc.). Each row in the prepaid transaction data represents a transaction (purchase of kWh) made by a given consumer that is using a prepaid meter. Each row shows the meter ID of the customer, their full name, their literal address, their phone number, their subscription date, their electricity (in kWh) corresponding to the amount purchased<sup>40</sup>, the size of their unit (small vs. medium vs. large), the type of the customer (residential vs. firm) and the tariff structure they are subject to (block pricing vs. peak load pricing). The raw prepaid data has 61,205,325 observations. The prepaid transactions are later collapsed at the month level to be comparable to the postpaid billing data. I collapse this dataset so that each row is a consumer-month pair. In total, I have approximately 25 million customer-month pairs between 2014 and 2023.

To construct electricity consumption, I use the amount of kWh reported in the post-

---

<sup>38</sup>Small and medium customers get their bills every 2 months while large clients receive their bills every month.

<sup>39</sup>The prepaid system requires people to pay before they can consume electricity, so the data are ex-ante purchases rather than ex-post bills

<sup>40</sup>Note that the kWh corresponding to a given purchase varies over time because of the block pricing and the reset of the tariff schedule at the beginning of the month

paid billing data and the prepaid transaction data. In the postpaid billing data, I consider the number of days between two bills to get the number of days covered by a bill. I then divide the bill amount by that number of days to get a measure of daily consumption over the time window. In the prepaid transaction data, I consider the number of days between a given transaction and the next transaction to get the number of days covered by the current transaction. I then divide the electricity purchased by the number of days to get a measure of daily consumption over the time window. Implicit in this derivation is the non-accumulation of electricity purchased with prepaid meters. I think this is a reasonable assumption since the block pricing makes it inefficient to buy electricity at higher blocks while having electricity still available.

### **Geocoded Household Surveys**

The households surveys are shared with the centroid of the enumeration areas to which they belongs. In total, the data covers 598 enumeration areas, out of a total of more than 5,827 enumeration areas in Senegal. So enumeration areas cover on average 33.76 squared kilometers. I match households to the stations, by matching the centroid of their enumeration area with the station closest to it.

### **Matching the Datasets**

For my empirical analysis, I need to do three matches: (A) match network maps to outage incidents in order to link reliability projects to outage incidents, (B) match network maps to household surveys in order to link reliability projects to household outcomes, and (C) match electricity bills to network maps in order to link reliability projects to monthly electricity consumption.

For (A), the matching is done through an exact matching of station names in the outage incidents datasets and station names in the network maps. This exact matching is straightforward, after cleaning out some data entry errors on the names of the stations recorded in the maps or in the logs of outage incidents.

For (B), I rely on the geocoded information that I have for the network maps and the household surveys. I match each enumeration area in the household survey to the closest station in the network map. This matching is straightforward since it simply uses the geocoordinates of the stations and the geocoordinates of the enumeration areas.

For (C), I prepare the addresses from the billing data and filter out invalid entries, such as addresses that are purely numeric and those without a region.

The addresses from the billing data are used to match clients to distribution stations using the two procedure: (i) matching literal addresses from the billing data with geocoded distribution stations directly based on the names, and (ii) matching literal addresses from the billing data with geocoded villages and neighborhoods based on the names, and then matching them with geocoded distribution stations.

On the first procedure, I write a script to matches these literal addresses to stations based on: (i) an exact matching of the district they are located, and (ii) a fuzzy matching of the literal addresses and the names of the distribution stations. I use three string distance metrics, the Jaro-Winkler distance, the Levenshtein distance and the Jaccard distance, to calculate the similarity between the literal addresses and station names. The Jaro-Winkler distance measures the similarity between two strings. It considers both the number of matching characters and the transpositions of these characters within the strings. The Jaro-Winkler distance is useful for comparing strings that are similar but may have slight differences, such as typographical errors or slight variations. The Levenshtein distance calculates the minimum number of single-character edits (insertions, deletions, or substitutions) required to change one string into the other. The Levenshtein distance is helpful for identifying strings that are almost identical but may have small differences. The Jaccard distance measures the dissimilarity between two sets. It calculates the Jaccard index, which is the size of the intersection of the sets divided by the size of the union of the sets. The closest matches are identified by finding the minimum distance among these three metrics for each address, and only the matches with a minimum distance of 0.15 are kept. The threshold of 0.15 is relatively low and conservative, but I use it to ensure that all matches that I keep are correct.

On the second procedure, I match literal addresses with geocoded villages and neighborhoods via a script that performs a matching similar to the first matching above: (i) (an exact matching of the department in which the village and the literal address are located, and (ii) a fuzzy matching of the literal addresses and the names of villages and neighborhoods.

## 9.2 Electricity Plans

In this subsection, I provide additional details on the consumer plans available in the electricity market in Senegal.

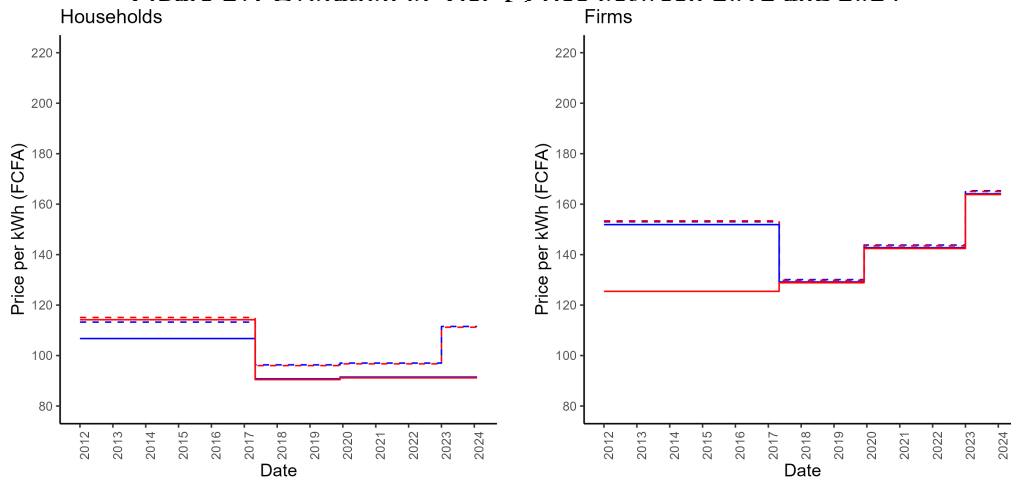
Each electricity plan has its own set of attributes including price and duration. The pricing scheme follows a second-degree price discrimination with increasing blocks of three, so the marginal price increases as I go from the first to the third block. The choice set is restricted to six plans for households and six plans for firms. Clients on time of use pricing, which are subject to peak and off-peak prices at different times of the day, are excluded from the analysis.

I leverage the difference in price between the different electricity plans available to customers for the analysis. The price differences between these electricity plans depend on two factors: customer type and meter type. Customer type depends on the usage type (residential vs. commercial) and the capacity load supplied by the utility (small vs. medium vs. large). The capacity load is categorized as either low (for capacity between 1 kW and 6 kW), medium (between 6 kW and 17 kW) or high (between 17 kW and 34 kW). Importantly, the number of appliances and the type of appliances owned by clients

is a major driver for the plan they choose. Each plan corresponds to a different tariff schedule, with different block cutoffs, and different marginal prices corresponding to these tiers.

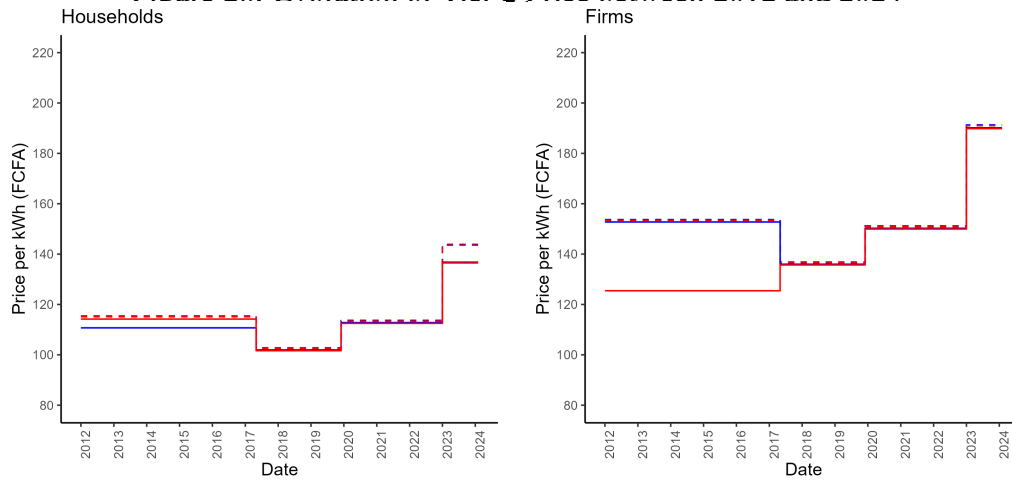
Regarding the meter types, there are two price differences between the two meter types. First, there are level differences in the prices corresponding to the different tiers of the tariff schedule for each meter type. For example, in the period between December for the first and the second blocks of usage, prepaid and postpaid meters have the same price schedule whereas prepaid meter is cheaper for the third block. Second, the duration covered by the quantity to which the tariff schedule is applied differs between the two meter types. For postpaid meters, the tariff schedules cover consumption over two months and the blocks are applied to this two-month consumption. For prepaid meters, the tariff schedules cover consumption over a month only, and the blocks are applied to this one-month consumption. Table 22 in the main text shows the prices corresponding to the different blocks for each meter type and each customer group. Figures 27, 28 and 29 in Appendix E show the evolution of the pooled average price for the different tariff schedules between 2012 and 2023. As highlighted in these figures, there is a sharp discontinuity in the average price of electricity once a new tariff schedule is enacted. Between 2012 and 2023, the average price has hovered between 110 FCFA ( $\sim 20$  cents in US dollars) and 210 FCFA ( $\sim 40$  cents in US dollars).

Figure 27: Evolution of Tier 1 Price between 2012 and 2024



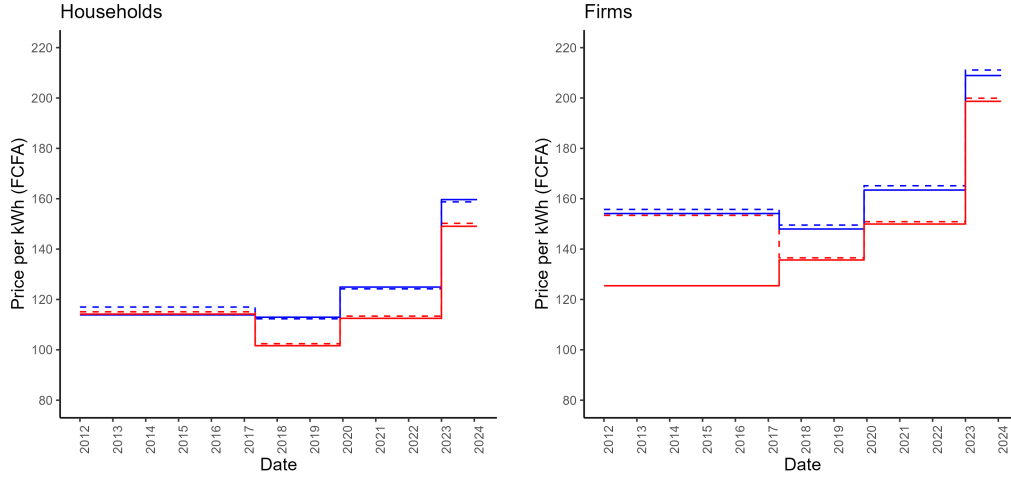
Notes: This figure displays the evolution of the tier 1 price for different groups of clients between 2012 and 2024. The left hand side graph shows the evolution of tier 1 price for residential clients and the right hand side graph shows the evolution of tier 1 price for commercial clients. Line colors are as follows: solid blue for small residential clients using postpaid meters, dashed blue for large residential clients using postpaid meters, solid red for small residential clients using prepaid meters, dashed red for large residential clients using prepaid meters. Small clients are those who receive a voltage level less than 1000 volts (based on their baseline number of appliances). Large clients are those that receive voltage levels between 1000 and 60 000 volts. Residential clients are clients whose main use of electricity is in a residential unit and firms are clients whose main use of electricity is in a commercial establishment. The unit prices are given in the local currency FCFA, which is pegged to the Euro. 1 US dollar is approximately 550 FCFA.

Figure 28: Evolution of Tier 2 Price between 2012 and 2024



Notes: This figure displays the evolution of the tier 2 price for different groups of clients between 2012 and 2024. The left hand side graph shows the evolution of tier 2 price for residential clients and the right hand side graph shows the evolution of tier 2 price for commercial clients. Line colors are as follows: solid blue for small residential clients using postpaid meters, dashed blue for large residential clients using postpaid meters, solid red for small residential clients using prepaid meters, dashed red for large residential clients using prepaid meters. Small clients are those who receive a voltage level less than 1000 volts (based on their baseline number of appliances). Large clients are those that receive voltage levels between 1000 and 60 000 volts. Residential clients are clients whose main use of electricity is in a residential unit and firms are clients whose main use of electricity is in a commercial establishment. The unit prices are given in the local currency FCFA, which is pegged to the Euro. 1 US dollar is approximately 550 FCFA.

Figure 29: Evolution of Tier 3 Price between 2012 and 2024



Notes: This figure displays the evolution of the tier 3 price for different groups of clients between 2012 and 2024. The left hand side graph shows the evolution of tier 3 price for residential clients and the right hand side graph shows the evolution of tier 3 price for commercial clients. Line colors are as follows: solid blue for small residential clients using postpaid meters, dashed blue for large residential clients using postpaid meters, solid red for small residential clients using prepaid meters, dashed red for large residential clients using prepaid meters. Small clients are those who receive a voltage level less than 1000 volts (based on their baseline number of appliances). Large clients are those that receive voltage levels between 1000 and 60 000 volts. Residential clients are clients whose main use of electricity is in a residential unit and firms are clients whose main use of electricity is in a commercial establishment. The unit prices are given in the local currency FCFA, which is pegged to the Euro. 1 US dollar is approximately 550 FCFA.

### 9.3 Stylized Facts

In this subsection, I present additional figures and tables related to a few stylized facts from the data to help motivate the reduced-form and model analysis in the main text.

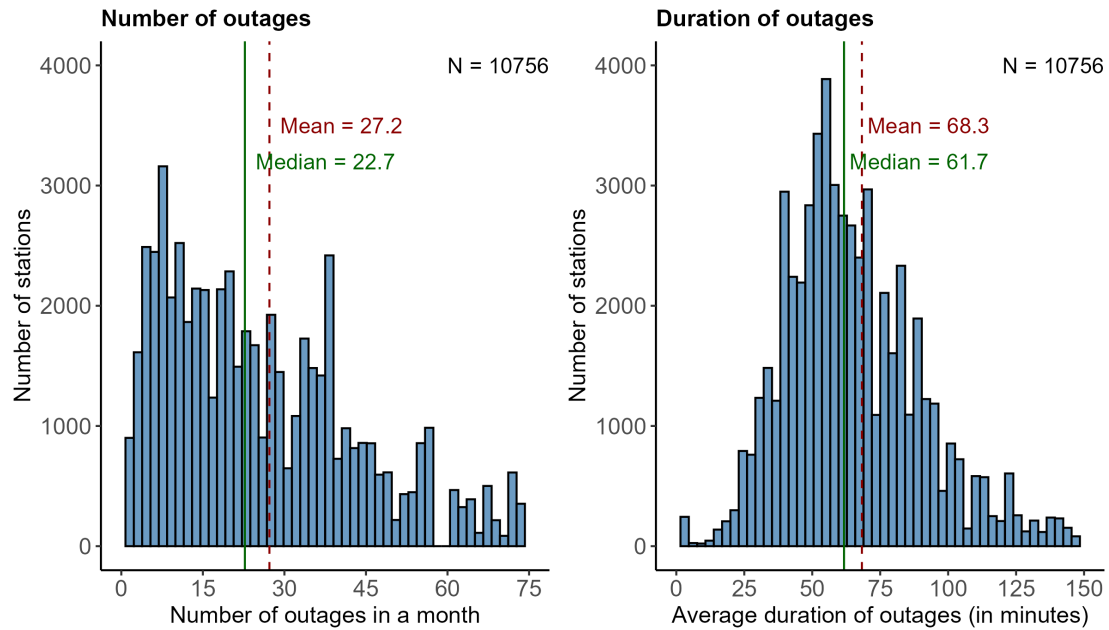
In this section, I present a few stylized facts on outages, appliance usage and electricity consumption from the data used in the analysis, and described in Section 3 of the main text.

#### Outages

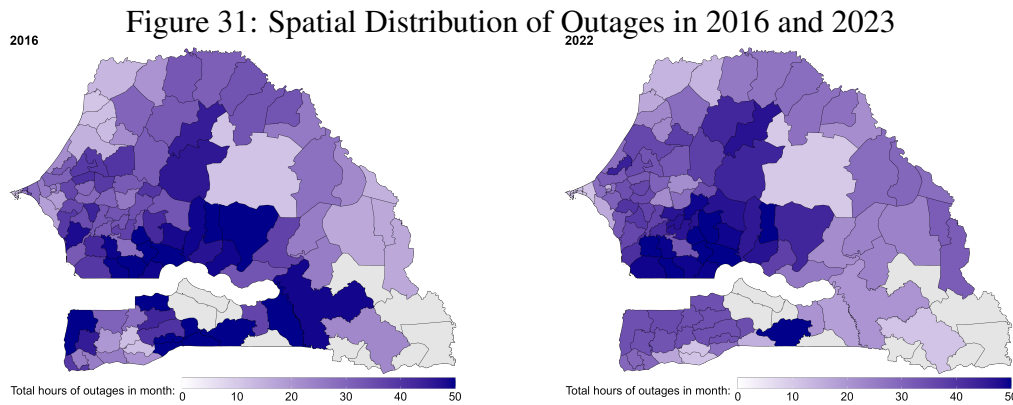
**FACT 1: OUTAGES VARY IN DURATION AND FREQUENCY ACROSS SPACE AND OVER TIME.**

Figure 30 shows the spatial distribution in the frequency and duration of outages across transformer stations in the data. The left hand side of the figure shows significant dispersion in the frequency of outages with some stations have virtually no outages over the course of a year and some stations getting outages on most days of year. The variation in reliability is less stark when I turn to duration of outages. Most outages only last a couple of hours in the settings and rarely do outages last more than 3 hours.

Figure 30: Histogram of Frequency and Duration of Outages across Distribution Stations



Notes: This figure shows histograms of the frequency of outages (as measured by the number of days with outages in a year) and the duration of outages (as measured by the mean duration of outages) across distribution stations. The y-axis represents the count of the number of stations by bin. Data include all outages and stations discussed in Section 3.



Notes: This map displays the spatial distribution of outages across departments, which are the second level of administrative units in Senegal. The left hand side map is for 2016 and the right hand side map is for 2022. The values for each department correspond to the mean total hours of outages in a month across all stations located in the department. Areas in white have low values of the variable and areas in dark blue have a high value of the variable. All distribution stations are included in the calculations. Grey areas are departments with no stations or no data available. Mean values above 50 hours of outages are replaced with 50 hours of outages. Data include all outages and stations discussed in Section 3.

FACT 2: MOST OUTAGES ARE UNPLANNED.

| Table 15: Number of Outages by Incident Type |       |       |       |       |
|----------------------------------------------|-------|-------|-------|-------|
| Type of outage                               | 2018  | 2019  | 2020  | 2021  |
| Line incidents                               | 23400 | 19414 | 19168 | 20989 |
| Supply issues                                | 69    | 286   | 140   | 230   |
| Removal of high-voltage clients              | 112   | 51    | 15    | 31    |
| Maintenance work                             | 6469  | 5655  | 6164  | 6015  |
| High/low voltage                             | 74    | 12    | 33    | 0     |

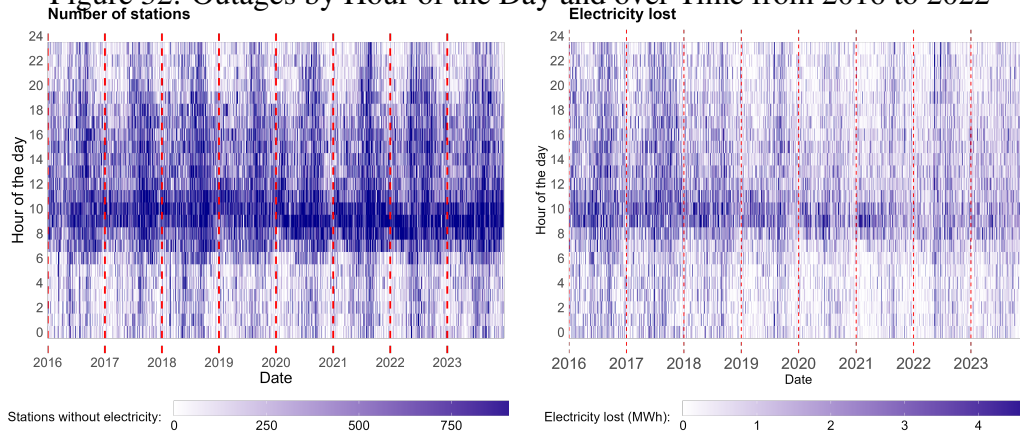
Notes: This table shows the number of outage events by type of incident as officially reported and categorized by the utility company in Senegal, SENELEC, from 2018 to 2021. Line incidents include any incident not recorded as one of the other type of incidents in the table. These line incidents include weather-related incidents, or incidents due to unknown causes. Supply issues refer to shortages in the supply of electricity. Maintenance work includes planned outages for work done on feeders and stations to resolve an incident that has caused an outage in the past or to upgrade lines. High/low voltage includes outages that are due to voltage fluctuations in the distribution lines. Data are sourced from SENELEC's annual reports.

### FACT 3: MOST OUTAGES OCCUR DURING THE DAY.

Another dimension of reliability relates to the timing of the outages throughout day. Figure 32 shows the outages by hour of the day and by day of the study period, both in terms of electricity lost and in terms of number of stations affected. The right hand side panel of the figure shows that most outages happen during the day between 8am and 3pm. This pattern is due to the fact that all maintenance work happen during that time window and most manufacturing firms and heavy users also operate during that time window. Still, the figure also shows that outages do happen at other times of the day and in particular in the evening and at night between 4pm and midnight. The period with the least number of outages is unsurprisingly the period with least demand (i.e., from midnight to 7am).

From Figure 32, I can also note the evolution in outages over time from 2016 to 2022. In particular, while the the heat map becomes slightly whiter over time and the left hand side figure on electricity lost suggests that reliability may have improved, the right hand side figure does not seem to indicate any particular decrease in the number of stations affected by outages. The fact that reliability did not seem to improve significantly speak to persistent reliability challenges faced by the utility, in an environment with growing population and growing electricity demand and muted expansion of supply generation.

**Figure 32: Outages by Hour of the Day and over Time from 2016 to 2022**



Notes: This heat map shows the intensity of outages in terms of number of distribution stations affected (on the left hand side) and in terms of electricity lost (on the right hand side), for each hour of the day and for each day between January 1, 2016 and December 31, 2023. The white parts of the heat map correspond to no outages and no electricity lost, and the dark blue parts of the map correspond to high electricity lost and to a high number of affected stations. Electricity lost during an outage is defined as electricity that would have been served absent the outage based on historical use at the particular times in which the outage occurred. The y-axis corresponds to the hour of the day in 24-hour scale (e.g. 18 corresponds to 6pm). The x-axis shows the day in the study period. Vertical red lines mark the end of the calendar year. Data includes all outages discussed in Section 3. The number of stations affected by outages in an hour is replaced by 5,000 if it is higher than 5,000. The amount of electricity lost in an hour is replaced by 10 MWh if it is higher than 10 MWh.

## Appliances

**FACT 1: CONSUMERS HAVE VERY FEW APPLIANCES IN GENERAL.**

Table 16 below displays the share of households with a given appliance and average number. Most households have at least one member with a phone and a good share of households have a TV and fan. Outside of these common appliance items, the vast majority of households do not have much of the fancier appliances on the table. This simple descriptive is important to set the stage regarding the margins where I would expect reliability to operate through. As the fancier appliances require a lot of capital investments from resource-constrained consumers, even large and significant reliability improvements might not be enough to increase adoption of these fancier appliances in the context.

Table 16: Appliance Roster of Customers in Senegal  
(a) Electrical appliances

| <b>Appliance</b> | <b>Share with<br/>Appliance (%)</b> | <b>Number of<br/>Appliance</b> |
|------------------|-------------------------------------|--------------------------------|
|                  | (1)                                 | (2)                            |
| Phone            | 91.43                               | 3.43                           |
| TV               | 48.23                               | 0.73                           |
| Fan              | 31.25                               | 0.68                           |
| Fridge           | 19.64                               | 0.23                           |
| Laptop           | 7.84                                | 0.11                           |
| Freezer          | 6.85                                | 0.08                           |
| Tablet           | 3.78                                | 0.05                           |
| Iron             | 2.37                                | 0.03                           |
| AC               | 1.45                                | 0.02                           |
| Microwave        | 0.61                                | 0.01                           |
| Wash             | 0.26                                | 0.00                           |
| Generator        | 0.10                                | 0.00                           |
| Vacuum           | 0.09                                | 0.00                           |
| Printer          | 0.44                                | 0.00                           |

(b) Non-electrical appliances

| <b>Appliance</b>  | <b>Share with<br/>Appliance(%)</b> | <b>Number of<br/>Appliance</b> |
|-------------------|------------------------------------|--------------------------------|
|                   | (1)                                | (2)                            |
| Mattress          | 88.02                              | 3.87                           |
| Bed               | 85.39                              | 3.16                           |
| Closet            | 44.95                              | 0.91                           |
| Gaz               | 48.24                              | 0.81                           |
| Carpet            | 31.40                              | 0.74                           |
| Non-Electric Iron | 42.84                              | 0.53                           |
| Radio             | 37.81                              | 0.52                           |
| Bicycle           | 19.71                              | 0.36                           |
| Sofa              | 19.01                              | 0.32                           |
| Table             | 10.36                              | 0.29                           |
| Stove             | 15.44                              | 0.19                           |
| Land              | 8.63                               | 0.13                           |
| Car               | 3.07                               | 0.04                           |
| Mixer             | 0.94                               | 0.01                           |

Notes: For each row corresponding to an appliance type, the tables show the share of households that own such appliance (column (1)) and the average number of this appliance type owned by a household in the sample (column (2)). The data is taken from the first wave of the household surveys. All respondents of the surveys are included in the calculations.

## FACT 2: CONSUMERS DO UPGRADE THEIR APPLIANCES OVER TIME.

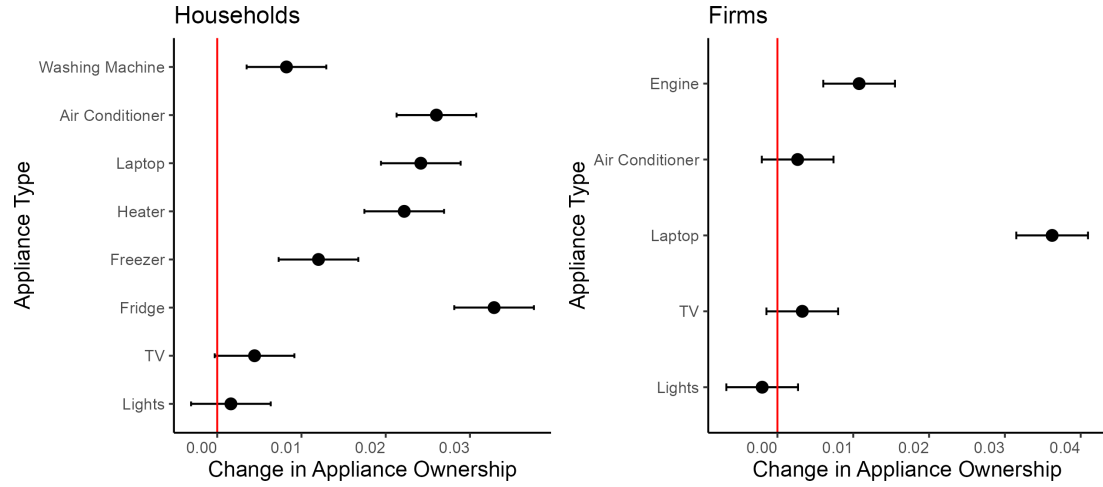
However, despite the low level of appliances in general, households do upgrade their appliances over time. Figure 33 uses the cross-sectional client characteristics data and shows the change in appliance ownership throughout the study period for clients who move between accounts and for which I have appliances at multiple points in time. Each point estimate and the associated confidence intervals is a separate regression of the type:

$$Y_{it,k} = \beta_0 + \beta_1 D_{it,k} + \theta_i + \lambda_t + \epsilon_{it,k} \quad (35)$$

Here,  $Y_{it,k}$  is a dummy variable equal to 1 if customer  $i$  owns appliance  $k$  at time  $t$ .  $D_{it,k}$  is a dummy equals to one if the time period corresponds to the last period for which I have data about the client and is zero otherwise.  $\theta_i$  and  $\lambda_t$  are customer fixed effects and year fixed effects. I run regression (35) separately for each appliance type and separately for households and firms and I record the  $\beta_1$  associated with each regression in Figure 33.

For low-intensity appliances such as light bulbs and TV, there is no significant change in appliance ownership over time: most households and firms already have lights and TV at baseline. However, for other fancier appliances, I see that households (and firms to a lesser extent) increase adoption between the time of subscription and the last period of data. Despite the low change in adoption over time, I show figure 33 in this subsection to show suggestive evidence that there is a possibility that customers in the setting upgrade their appliance roster over time and that they may decide to acquire high-intensity appliances under favorable electricity environments. Even though the fixed-effects account for some of the unobserved determinants of appliance upgrades, Figure 33 is purely correlational since the appliance upgrade may be correlated with the reasons of the move between locations.

Figure 33: Change in Appliance Ownership by Customer Type and by Appliance Type

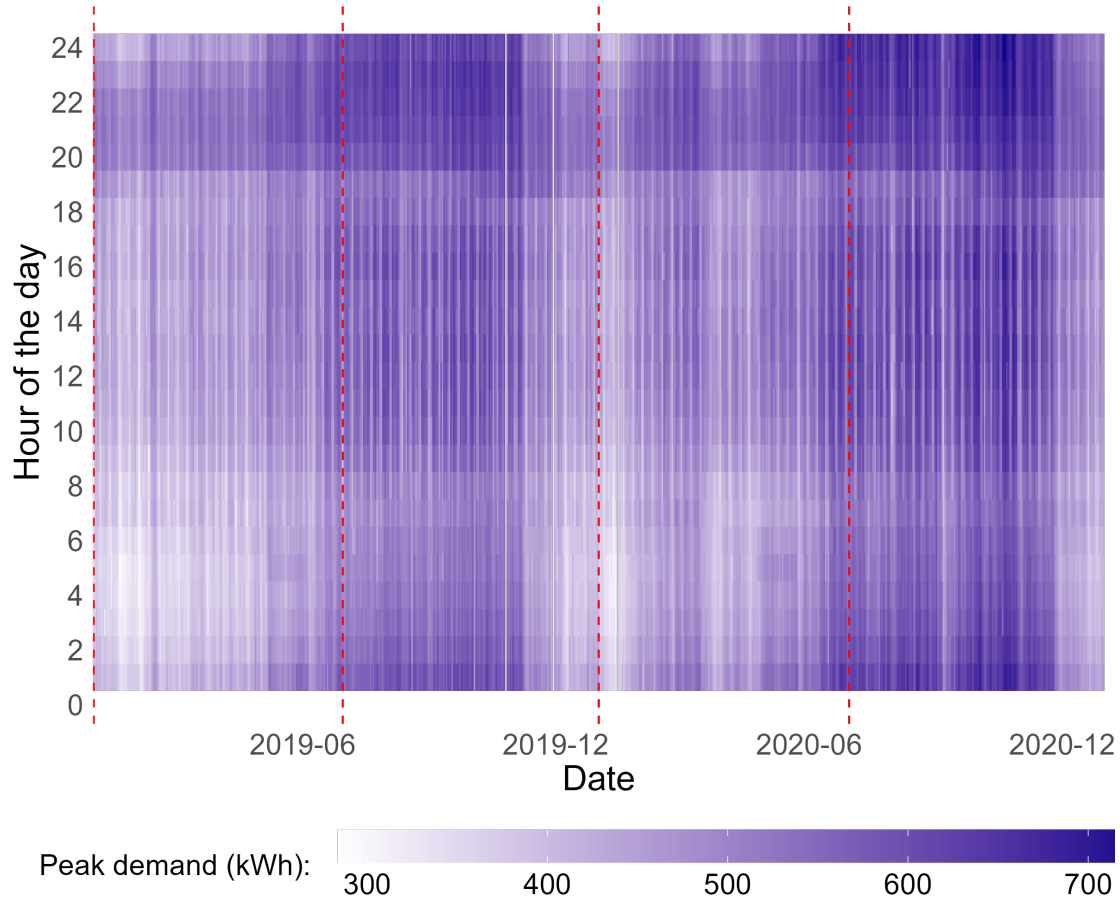


Notes: This figure shows the share of existing customers who adopt new appliances. For each appliance shown on the y-axis, the x-axis shows the average change in appliance ownership from the time the client became a client to the end of the study period. For each appliance type  $k$  on the y-axis, the black dots corresponds to  $\beta_1$  in equation (35) and represent the share of customers who went from having no appliance  $k$  to having at least one of appliance  $k$  between their subscription date and 2022. Horizontal black lines correspond to 95% confidence intervals associated with the estimated change. Data include the clients-characteristics dataset discussed in Section 3. Figure is restricted to clients who move between locations throughout the study period.

## Electricity Consumption

FACT 1: ELECTRICITY CONSUMPTION IS HIGHEST BETWEEN 7PM AND 11PM AND DURING THE SUMMER.

Figure 34: Peak Electricity Demand Throughout the Day and Over Time  
**Electricity demand**



Notes: This heat map shows the peak electricity demand for each hour of each day between January 1, 2019 and December 31, 2020. The white parts of the heat map correspond to lower peak demand, and the parts of the map in dark blue correspond to higher peak demand. The y-axis corresponds to the hour of the day in 24-hour scale (e.g. 18 corresponds to 6pm). The x-axis shows the day in the study period. Vertical red lines mark the end of the calendar year. Data is the aggregate peak demand obtained directly from SENELEC; this data is not included in the analysis and is not discussed in Section 3.

## 10 Robustness Checks

In this section, I conduct a battery of robustness checks to address the main confounders in the empirical analysis: (i) potential violation of SUTVA, (ii) potential violation of parallel trends assumption, (iii) sensitivity of the results to the functional forms of the outcome variables, the particular estimator used, the types of fixed effects included in the regressions, (iv) sensitivity of the results to different control variables that reflect other changes in the electricity grid network over time, and (v) robustness of the results

to a restricted sample of individuals with the same level of exposure to treatment.

## 10.1 SUTVA

the analysis relies on the SUTVA. This assumption might be violated in the context if potential outcomes for untreated stations are affected by the treatment status of stations that are nearby. More specifically, I need to following assumption to hold for the estimates to be unbiased:

$$Y_{it}(D_{it}, D_{-it}) = Y_{it}(D_{it}) \quad (36)$$

Here,  $D_{-it}$  is the treatment status of other distribution stations in the sample. In other words, the potential outcomes  $Y_{it,j}$  of station  $i$  depends only on  $i$ 's own treatment assignment status,  $D_{it}$ , and is not affected by the treatment status of other distribution stations in the sample. Since the outcome variable of interest at the station level is electricity reliability (duration of outages, frequency of outages, etc.), SUTVA here would imply that: (i) whether a distribution station receives the reliability improvement devices does not depends on whether other distribution stations have received them, and (ii) the potential reliability levels of a distribution station is not affected by the treatment status of other distribution stations nearby. SUTVA would thus rule out spillover effects and general equilibrium effects.

Checking the validity of SUTVA is particularly important in my context, since I study distribution stations that are in a network where all objects are directly or indirectly connected with one another. In particular, there are three reasons why local spillover effects might exist. First, there might be a mechanical interconnection between treated and non-treated components of the electricity network, given that all control stations not directly treated will affect and be affected by treated stations, either directly (by being on the same line) or indirectly through the communication of loads, outages and incidents across the network. If this interconnection affects the potential outcomes for treated units, then the estimated effect of reliability projects on outages would be an underestimate, as it would also account for the indirect (arguably positive) effect of the projects on untreated stations. Second, spillover effects might decay with distance, since stations in feeder lines far away will not be as affected as stations in close feeder lines, and stations within the same feeder lines will have a higher level of dependence than stations across feeder lines. To the extent that the placement of the stations is not random, the magnitude of the potential bias might spatially vary in a way that has important economic ramifications. Third, the spillovers are likely to be additive in the context since reliability of a given not-yet-treated station might improve with the number of nearby stations that receive the reliability improvement devices.

To assuage this concern, I note that in my context, the interconnection between transformer stations that are located at different feeder lines is muted by the presence

of circuit breakers at the tails of feeder lines. Though these circuit-breakers are not remotely controlled (unlike the devices I study in this paper), they help reduce the pass-through of outages across feeder lines. The stylized facts that I present in Appendix C3 regarding the heterogeneity in duration and frequency of outages, and the stickiness of this heterogeneity over time speaks to the relatively little communication between stations of different feeder lines. The remaining interconnection that I have to worry about is that between stations located in the same feeder line. Since I cluster the standard errors at the feeder level, I account for the correlation between outcomes of these stations located within the same feeders.

To formally test for spillover effects, we estimate regression 3 discussed in Section 4.1. Table 17 shows the results of our estimation.

| Table 17: Spillover Effects of Reliability Projects on Untreated Stations |                               |                          |
|---------------------------------------------------------------------------|-------------------------------|--------------------------|
|                                                                           | Total hours of outages<br>(1) | Number of outages<br>(2) |
| Early treated x Post                                                      | -2.172**<br>(1.099)           | 2.159<br>(1.522)         |
| Post x Upstream                                                           | 0.2442<br>(1.006)             | 5.848***<br>(1.289)      |
| Early treated x Post x Upstream                                           | -4.146**<br>(1.711)           | -8.480***<br>(1.719)     |
| Pre-Treatment Downstream DV Mean                                          | 16.65                         | 22.97                    |
| Year FEs                                                                  | Yes                           | Yes                      |
| Station FEs                                                               | Yes                           | Yes                      |
| Controls                                                                  | Yes                           | Yes                      |
| Observations                                                              | 160,416                       | 160,416                  |
| R <sup>2</sup>                                                            | 0.39919                       | 0.62997                  |

Notes: Each column shows the results of a triple difference-in-differences regression with the interactions between three dummy variables: (1) a dummy for whether the feeder line received treatment early, (2) a dummy for whether the station is upstream or downstream of the treated station, and (3) a dummy for whether the year is post treatment. Standard errors are clustered at the station level. All regressions include year fixed-effects and feeder fixed-effects. Controls include: number of stations on the same feeder line and number of feeders from the same source station

## 10.2 Sensitivity of Results to Potential Violations of Parallel Trends

There are several valid reasons why I should conduct formal tests of pre-trends despite the main event-study results passing the “eye test” of the parallel trend assumption.

First, despite the fact that I condition the parallel trends checks on observed covariates, I still face the possibility that the selection criteria are correlated with unobserved time-varying factors that affect electricity reliability as well as the firm and household

outcomes that I study. For example, one of the criteria used is the number of clients served by the station and this number could reflect people moving in and out of areas as a result of changing economic conditions in their origin and destination areas. Second, there could be concern about the sensitivity of the parallel trends assumption to the functional form of the outcome variables of interest. The main results show that there are no pre-trends for the outcomes of interest in log forms, but these outcomes are mostly in logs. As documented by [Roth et al. \(2023\)](#), it may not be the case that parallel trends would still hold if I show the same results with the outcome variables presented in levels rather than in logs.

Third, I could have a low power or pre-test bias that would make us statistically reject the existence of pre-trends, even if they truly existed in the data. As discussed by Bilinski and Hatfield (2018) and [Kahn-Lang and Lang \(2020\)](#), there are several issues that could cause this: imprecise pre-treatment coefficients, insufficient power to reject the null, null hypothesis specified in the wrong directions and leading to too many false negatives (if there is not enough data) or too many false positives (if there is a lot of data). In this case, these conditions might mislead us into concluding that there is no violation of the parallel trend assumption. Finally, the analysis sample could be a biased sample of the true data generation process, in a way that parallel trend assumption holds for the sample without holding in the true data.

All of the above concerns call for a more formal test of the pre-trends assumption, or a way to interpret the results if the pre-trends test fail to give us enough assurance about the validity of the parallel trend assumption.

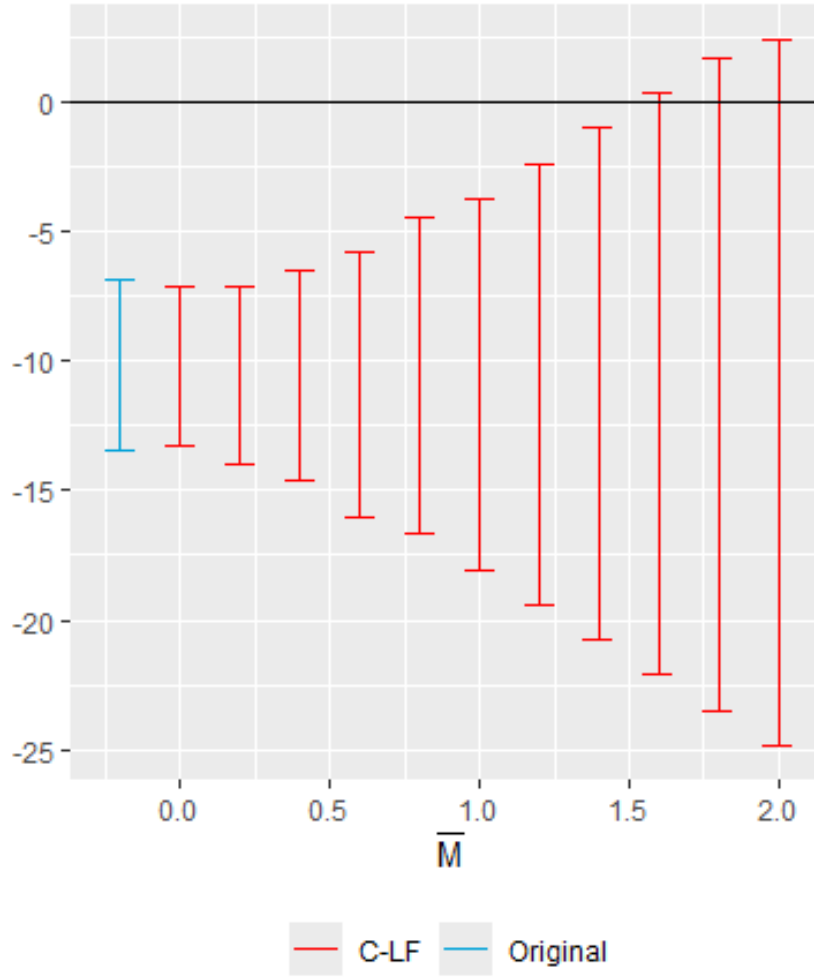
To formally check for the sensitivity of the results to parallel trend violations, I conduct formal tests using [Rambachan and Roth \(2023\)](#)'s proposed techniques to bound potential violations of post-treatment parallel trends, both in terms of their magnitude relative to pre-treatment violation of parallel trend, and in terms of a linear extrapolation of the slope of the pre-trends coefficient. I use the R package `HonestDID` developed by [Rambachan and Roth \(2023\)](#) for this test. This approach thus does not make the assumption that parallel trends holds in the pre-periods and simply check how robust the post-treatment results are conditional on there being a violation of parallel trends in the pre-period.

Figure 35 shows the change in the estimated impact of reliability projects on outages based on different assumptions on the magnitude of the post-treatment violation of parallel trend relative to the worst pre-treatment violation of parallel trends (between consecutive periods). The x-axis shows different upper bounds of this relative magnitude, which I denote as  $\bar{M}$  following the notation in [Rambachan and Roth \(2023\)](#). For example,  $\bar{M} = 1$  means that I impose post-treatment violation of parallel trend to be at most as large as the worst pre-treatment violation of parallel trends between consecutive periods.  $\bar{M} = 0$  corresponds to the default assumption of no post-treatment violation in parallel trends. The y-axis shows the confidence interval of the average treatment effect

of total hours of outages under the assumed  $\bar{M}$ .

Figure 35 shows that the main results on the impact of reliability projects on outages remain significant as long as the post-treatment violation in parallel trends is no more than 1.5 times the worst pre-treatment violation of parallel trends between consecutive periods. I can thus reject a null treatment effect unless I am willing to assume that post-treatment violation of parallel trend are larger than 1.5 times the worse parallel trend violation between consecutive periods in the pre-treatment period.

Figure 35: Sensitivity of Results to Post-Treatment Violation of Parallel Trends (Relative to the Worst Pre-Treatment Violation of Parallel Trends)



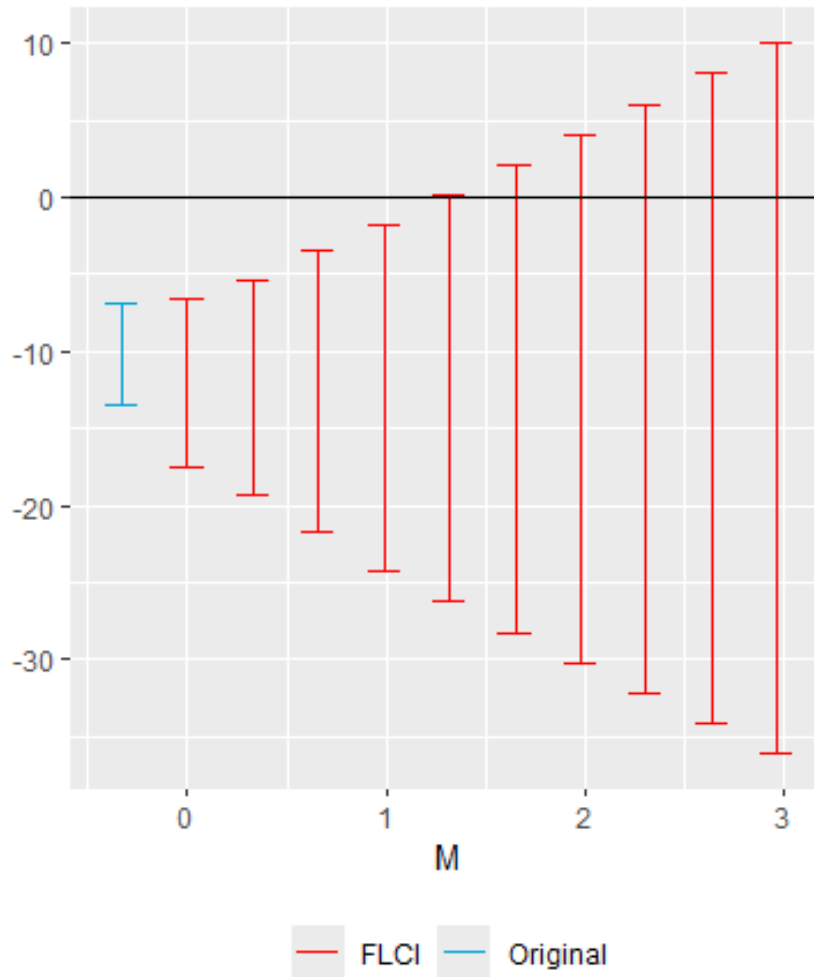
Notes: This figure shows the change in the estimated impact of reliability projects on outages based on different assumptions on the magnitude of the post-treatment violation of parallel trends, relative to the worst pre-treatment violation of parallel trends (between consecutive periods). The x-axis shows different upper bounds of this relative magnitude, which I denote as  $\bar{M}$  following the notation in [Rambachan and Roth \(2023\)](#). For example,  $\bar{M} = 1$  means that I impose post-treatment violation of parallel trend to be at most as large as the worst pre-treatment violation of parallel trends between consecutive periods.  $\bar{M} = 0$  corresponds to the default assumption of no post-treatment violation in parallel trends. For each assumed  $\bar{M}$ , the y-axis shows the confidence intervals of the average  $\lambda_l$  for the post-treatment periods in regression (1) with total hours of outages as the outcome variables and the estimation done using the [Callaway and Sant'Anna, 2021](#)). The blue line corresponds to the baseline results, and the red lines correspond to different confidence intervals for the post-treatment parallel trend violations. The figure is produced using the R package `HonestDID` developed by [Rambachan and Roth \(2023\)](#). “C-LF” stands for “Conditional-least favorable”. See [Rambachan and Roth \(2023\)](#) for more details on the methodology.

Figure 36 shows the change in the estimated impact of reliability projects on outages based on different assumptions on the magnitude of the post-treatment violation of parallel trends relative to the slope of the difference in trends in the pre-treatment periods.

The x-axis shows different upper bounds of this relative magnitude, which I denote as  $M$  following the notation in [Rambachan and Roth \(2023\)](#). For each  $M$ , I impose that the slope of the difference in trends in the post-treatment periods to be no more than  $M$  percentage points that between two consecutive periods in the pre-treatment periods. For example,  $M = 1$  means that I am willing to allow for the linear extrapolation of pre-treatment slope to be off by 1 percentage point across consecutive periods.  $M = 0$  corresponds to the baseline results. The y-axis shows the confidence interval of the average treatment effect of total hours of outages under the assumed  $\bar{M}$ .

Figure 36 shows a breakdown value for a significant effect is  $M = 1$ . This means that the main results on the impact of reliability projects on outages would remain significant even in the presence of parallel trend violations in the pre-treatment periods, as long as I am willing to allow for the linear extrapolation across consecutive periods to be off by less than 1 percentage point.

Figure 36: Sensitivity of Results to Post-Treatment Violation of Parallel Trends (Relative to Slope in Pre-Treatment Periods)



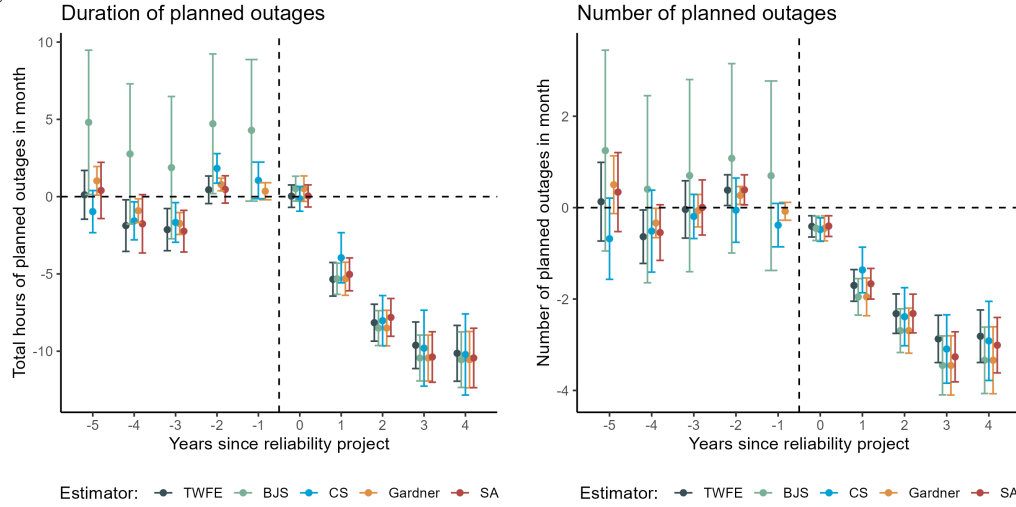
Notes: This figure shows the change in the estimated impact of reliability projects on outages based on different assumptions on the magnitude of the post-treatment violation of parallel trends relative to the slope of the difference in trends in the pre-treatment periods. The x-axis shows different upper bounds of this relative magnitude, which I denote as  $M$  following the notation in [Rambachan and Roth \(2023\)](#). For each  $M$ , I impose that the slope of the difference in trends in the post-treatment periods to be no more than  $M$  percentage points that between two consecutive periods in the pre-treatment periods. For example,  $M = 1$  means that I am willing to allow for the linear extrapolation across consecutive periods to be off by 1 percentage point.  $M = 0$  corresponds to the baseline results. For each assumed  $M$ , the y-axis shows the confidence intervals of the average  $\lambda_l$  for the post-treatment periods in regression (1) with total hours of outages as the outcome variables and the estimation done using the [\(Callaway and Sant'Anna, 2021\)](#). The blue line corresponds to the baseline results, and the red lines correspond to different confidence intervals for different bounds of the post-treatment parallel trend violations. The figure is produced using the R package `HonestDID` developed by [Rambachan and Roth \(2023\)](#). “FLCI” stands for “Fixed-length confidence interval”. See [Rambachan and Roth \(2023\)](#) for more details on the methodology.

### 10.3 Sensitivity of Results to Specific Functional Forms, Estimators, and FEs

In this subsection, I estimate the reduced-form impacts of reliability projects on electricity consumption, on appliance adoption and on economic outcomes for firms and households using alternative packages developed recently. As shown on Table 2 of [Roth et al. \(2023\)](#), there are numerous packages I could use as alternative packages. Since they all rely on different assumptions and specifications, I briefly discuss the relevance and differences between (some of) them to justify the final choice. The main specification uses the package developed by [Callaway and Sant’Anna \(2021\)](#). This is the preferred specification since I want to be as conservative as possible and not take for granted the validity of parallel trends assumptions over longer time horizons. In addition, this paper studies electricity network, which means that I am likely to deal with station-level outcomes that are serially correlated. In contrast, imputation-based estimator such as [Borusyak et al. \(2024\)](#) have the advantage of producing robust estimates. However, I don’t use it in the main specification since it is more suited for contexts where the outcome is not too serially correlated and the parallel trends are not violated across all periods.

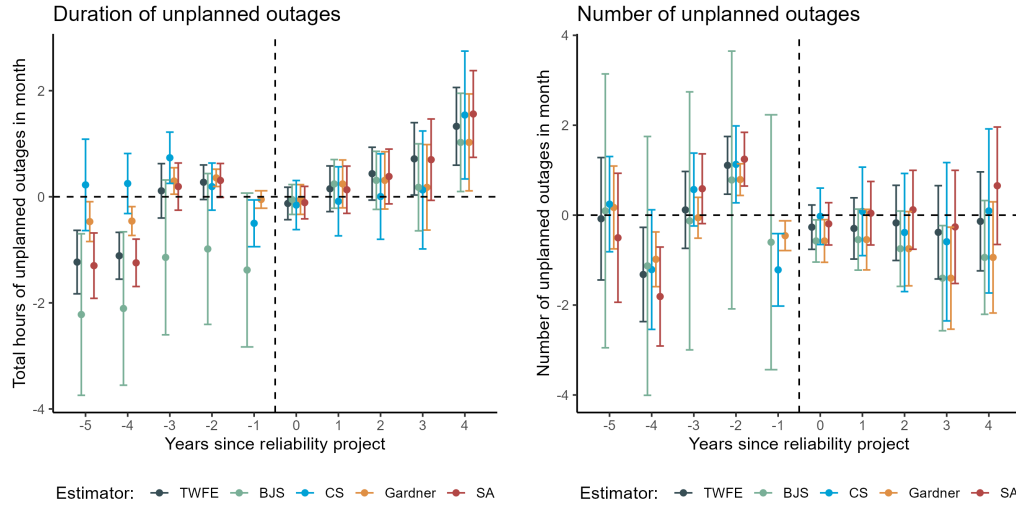
Figure 39 shows the impacts of reliability projects on the duration and frequency of outages using different estimators and show that the results broadly hold for these different estimators. Figures 37 and 38 show similar figures, separately for planned outages and unplanned outages respectively.

Figure 37: Impact of Reliability Projects on Duration and Frequency of Planned Outages



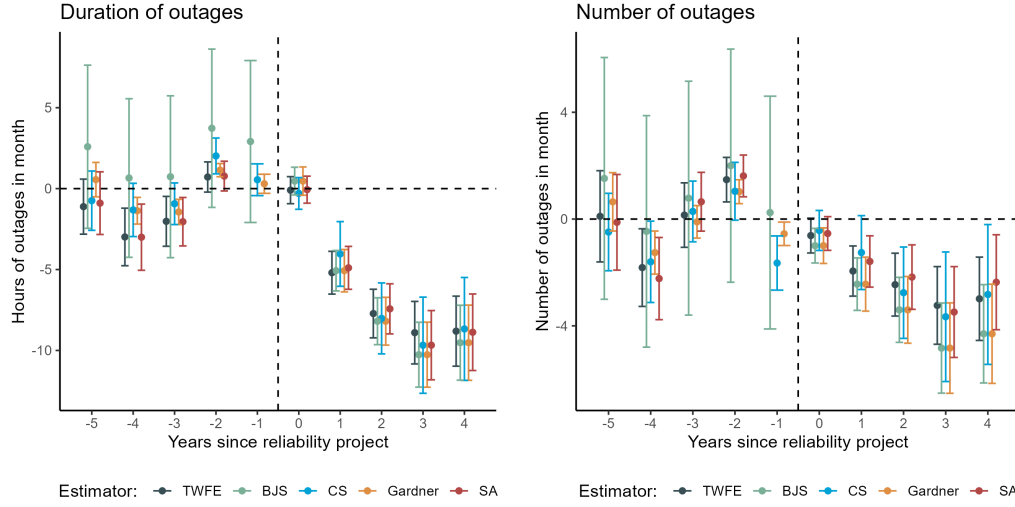
Notes: This figure shows the coefficients  $\lambda_k$  in regression (1) and their corresponding confidence intervals, with the duration and frequency of planned outages as the outcome variables, and the estimation done using the following estimators: Callaway and Sant'Anna, 2021 (CSA); Sun and Abraham, 2021 (SA); Borusyak et al., 2024 (BJS) and Gardner (2021). All coefficients are estimates relative to a baseline value for the year preceding project date. The left-hand side graph shows the duration of unplanned outages in terms of total hours of outages in month. The right-hand side graph shows the frequency of unplanned outages in terms of number of outages in month. For both graphs, the x-axis shows the number of years that has passed since the completion of the reliability project. The vertical dashed grey line corresponds to the year of project date. Data used for estimation is a balanced sample of treated stations with outage observations for all time periods of the analysis, from 5 years before project date to 4 years after project date. The estimation is done at the distribution station level.

Figure 38: Impact of Reliability Projects on Duration and Frequency of Unplanned Outages



Notes: This figure shows the coefficients  $\lambda_k$  in regression (1) and their corresponding confidence intervals, with the duration and frequency of unplanned outages as the outcome variables, and the estimation done using the following estimators: Callaway and Sant'Anna, 2021 (CSA); Sun and Abraham, 2021 (SA); Borusyak et al., 2024 (BJS) and Gardner (2021). All coefficients are estimates relative to a baseline value for the year preceding project date. The left-hand side graph shows the duration of unplanned outages in terms of total hours of outages in month. The right-hand side graph shows the frequency of unplanned outages in terms of number of outages in month. For both graphs, the x-axis shows the number of years that has passed since the completion of the reliability project. The vertical dashed grey line corresponds to the year of project date. Data used for estimation is a balanced sample of treated stations with outage observations for all time periods of the analysis, from 5 years before project date to 4 years after project date. The estimation is done at the distribution station level.

Figure 39: Impact of Reliability Projects on Duration and Frequency of Outages (with Multiple Estimators)

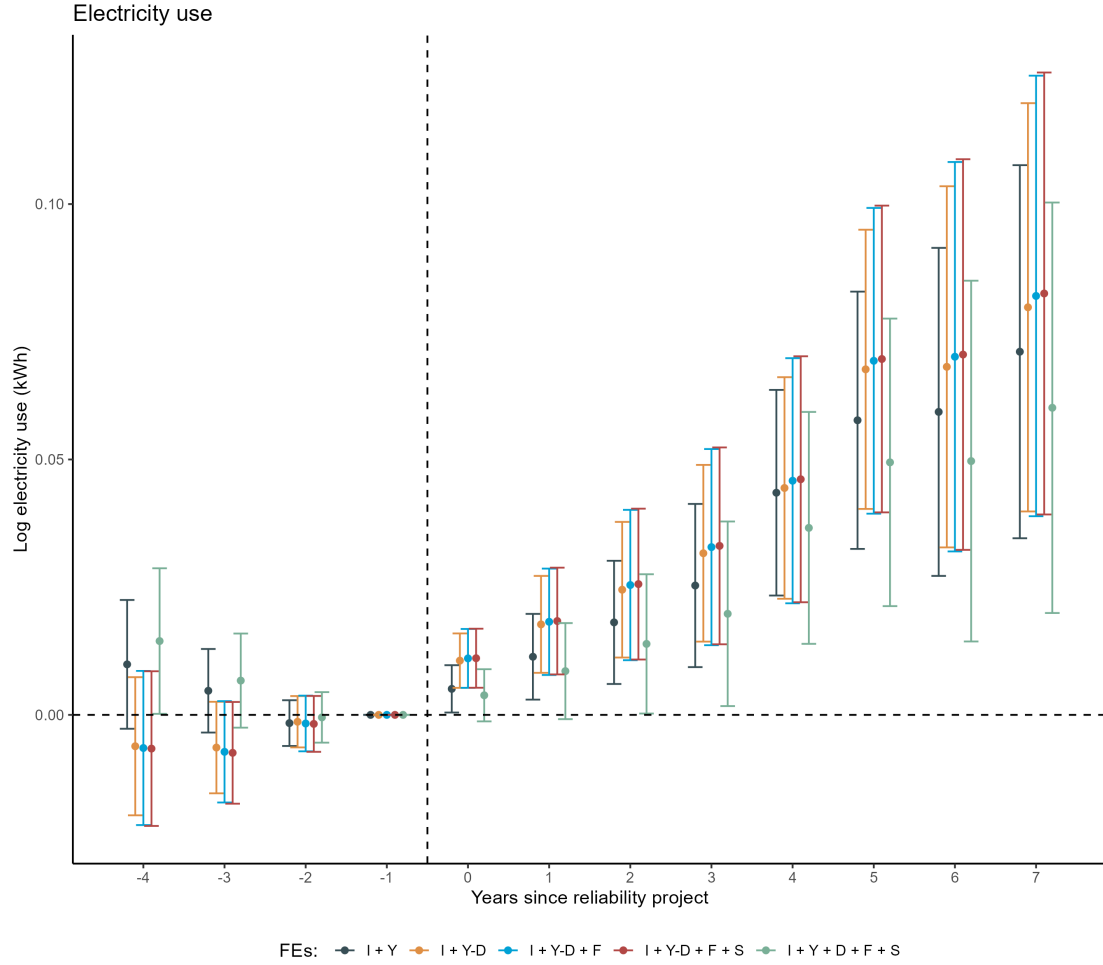


Notes: This figure shows the coefficients  $\lambda_k$  in regression (1) and their corresponding confidence intervals, with the duration and frequency of outages as the outcome variables and the estimation done using the following estimators: Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Borusyak et al., 2024 and Gardner (2021). All coefficients are estimates relative to a baseline value for the year preceding project date. The left-hand side graph shows the duration of outages in terms of total hours of outages in month. The right-hand side graph shows the frequency of outages in terms of number of outages in month. For both graphs, the x-axis shows the number of years that has passed since the completion of the reliability project. The vertical dashed grey line corresponds to the year of project date. Data used for estimation is a balanced sample of treated stations with outage observations for all time periods of the analysis, from 5 years before project date to 4 years after project date. The estimation is done at the distribution station level.

In addition to testing the sensitivity of the results to the use of different estimators, I also run the specifications of the empirical analysis with different types fixed-effects. This is because there may be concerns that stations and individuals treated later may coincide with major events (e.g. COVID-19 period), thus potentially confounding the results. While the year fixed-effects that I include in the regression should help account for any homogeneous effects of nation-wide events, I also check the robustness of the results to different combinations of the following fixed effects: individual fixed effects, year fixed effects, district fixed effects, feeder fixed effects, station fixed effects and district-by-year fixed effects.

Figure 40 shows the effects of the reliability projects on electricity use, with the coefficients  $\lambda_l$  in regression (3) and their corresponding confidence intervals, with the above fixed effects

Figure 40: Impact of Reliability Projects on Electricity Use (with Different Types of Fixed Effects)



Notes: This figure shows the effects of the reliability projects on electricity use, with the coefficients  $\lambda_l$  in regression (3) and their corresponding confidence intervals, estimated under different fixed-effects specifications. All coefficients are estimates relative to a baseline value for the year preceding project date. The y-axis shows electricity use in terms of the mean log electricity use (kWh) consumed in the year. The x-axis shows the number of years since the completion of the reliability project. The vertical dashed grey line corresponds to the year of project date. Data used for estimation is the geocoded billing data, where each observation is a customer of the utility company. Estimation is at the client level. “I+Y” includes individual and year fixed-effects. “I+Y-D” (baseline) adds year-by-district fixed-effects. “I+Y-D+F” adds feeder fixed-effects. “I+Y-D+F+S” adds station fixed-effects. “I+Y+D+F+S” includes individual, year, district, feeder, and station fixed-effects.

## 10.4 Controlling for Other Changes in Electricity Network over Time

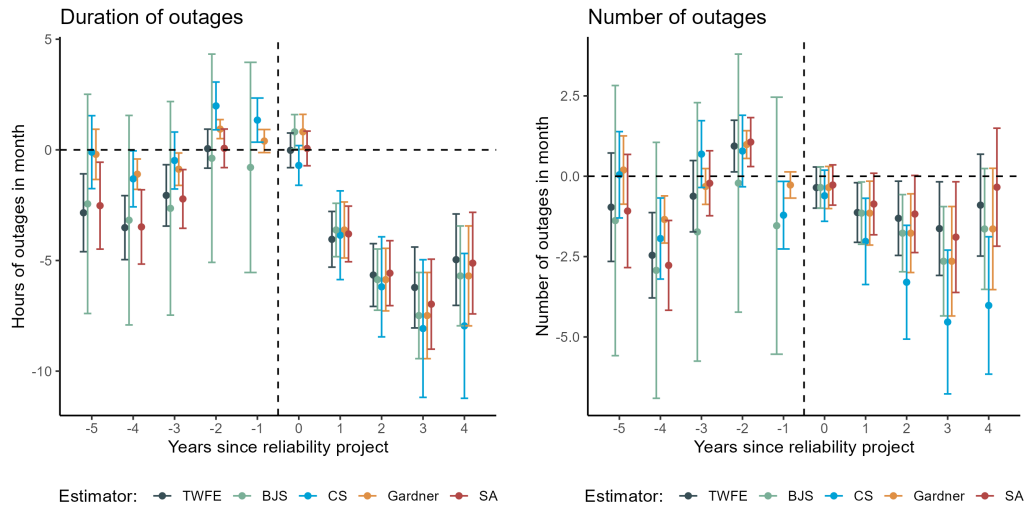
In this sub-section, I show the sensitivity of the results to the inclusion of controls. Given the downward in the estimated coefficients for the impact of reliability projects on outages and the upward trend in the estimated coefficient for the impact of reliabil-

ity projects on electricity use, I may be worried about the additiveness of the treatment (i.e., more stations of the same feeder becoming treated over time, and this resulting in increasing effect over time), the effect of other changes of the electricity grid network (both observed and unobserved to the researcher) on outages and electricity consumption.

In order to address these concerns, I include the following control variables for each distribution station  $l$  in the analysis: the number of treated stations on the feeder line as  $l$ , the number of stations on the same feeder line as  $l$ , the number of feeder lines served by the same sub-station as the sub-station serving  $l$ , the rank of  $l$  among the stations on the same feeder line as  $l$ , the importance of  $l$  in terms of the critical services it serves, the voltage level of  $l$  and the number of sub-lines of  $l$ .<sup>41</sup>

Figure 41 shows the replication of 39 with the inclusion of the above controls. The standard errors are now larger, but the trends and significance of estimates are qualitatively similar.

Figure 41: Impact of Reliability Projects on Duration and Frequency of Outages (with Control Variables)



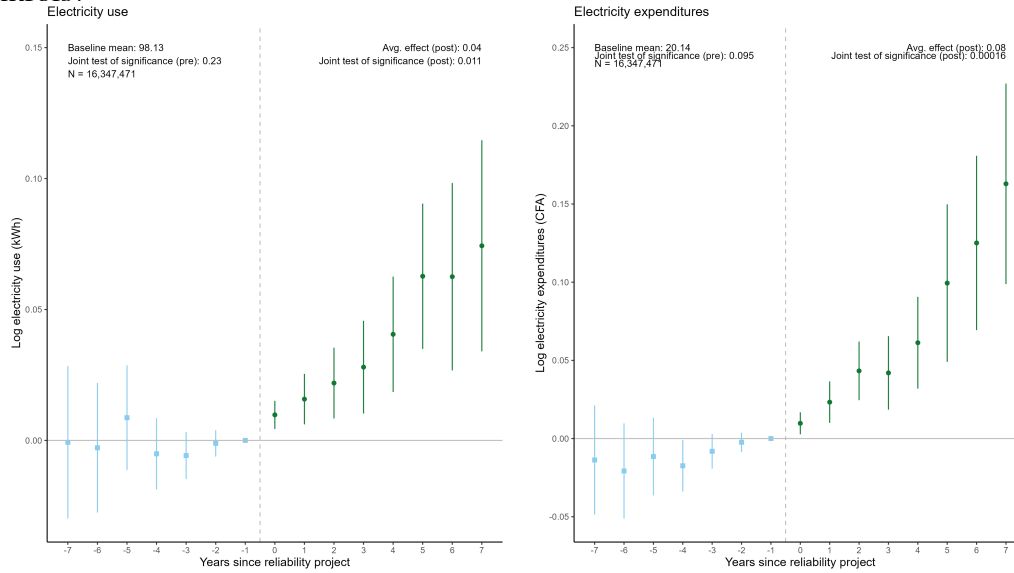
Notes: This figure shows the coefficients  $\lambda_k$  in regression (1) and their confidence intervals, estimated with control variables. Estimators: Callaway and Sant'Anna (CSA); Sun and Abraham (SA); Borusyak et al. (BJS); Gardner (2021). Controls for each station  $l$ : number of treated stations on its feeder; total stations on feeder; number of feeders served by its sub-station; rank of  $l$  on its feeder; critical-service importance; voltage level; number of sub-lines. All coefficients are relative to the year before project date. Left: total hours of outages per month. Right: number of outages per month. X-axis: years since project. Vertical dashed line: project year. Sample: balanced treated-station panel over 5 years pre to 4 years post.

Figure 42 shows the replication of Figure 4 with the inclusion of the above controls. The standard errors are slightly larger, but the trends and significance of estimates are

<sup>41</sup>Please see Section 2 for more details on the technical meaning of these different variables and their relevance for the analysis.

overall similar.

Figure 42: Impact of Reliability Projects on Electricity Use and Expenditures (with Controls)



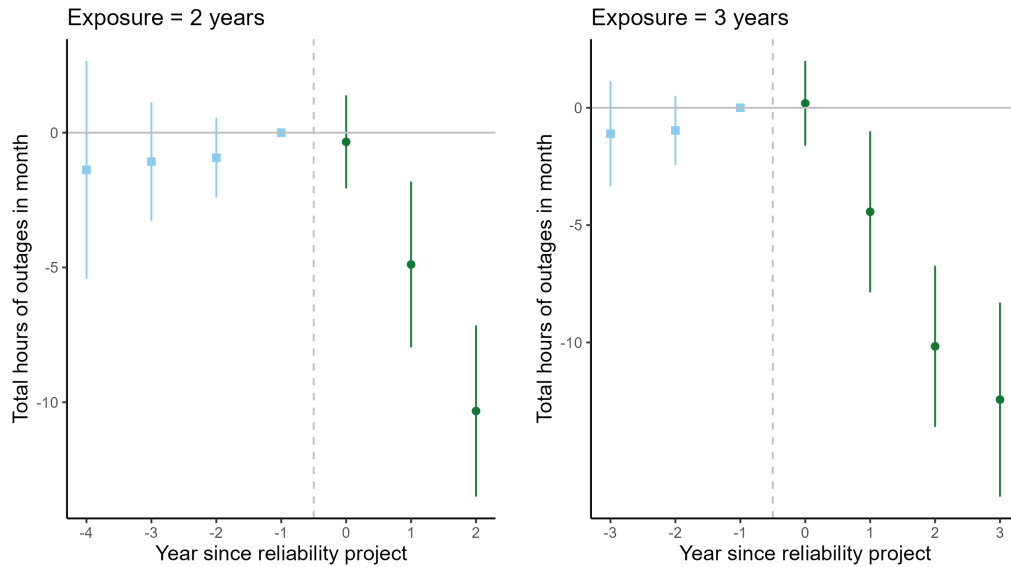
Notes: This figure shows the coefficients  $\lambda_l$  in regression (3) and their confidence intervals, estimated with the same station-level controls listed in Figure 41. Left: mean log electricity use (kWh) per year. Right: mean log electricity expenditures (CFA Franc). All coefficients are relative to the year before project date. X-axis: years since project. Vertical dashed line: project year. Sample: geocoded billing data, client-level panel.

## 10.5 Restricted Sample of Units with Balanced Exposure to Treatment

the results shown in subsection 4.2 rely on a balanced sample of individuals and stations that appear in each of the relative time periods shown on the main figures. However, the exposure to treatment is different across individuals given that I compare early-treated individuals and stations to late-treated individuals and stations. In particular, some individuals are exposed to treatment for much longer, which means that the coefficients for the earliest periods pre-treatment and the latest periods post-treatment only applies to these individuals. This change in the composition of the sample raises challenges in interpreting the different coefficient estimates, and in understanding the trend associated with them.

Figure 43 replicates Figure 16 but includes, in its estimation, only stations with exposure to treatment of two years and stations with exposure to treatment of three years, respectively in its left and right panels. I see the trends and coefficient estimates are broadly similar to those of Figure 16.

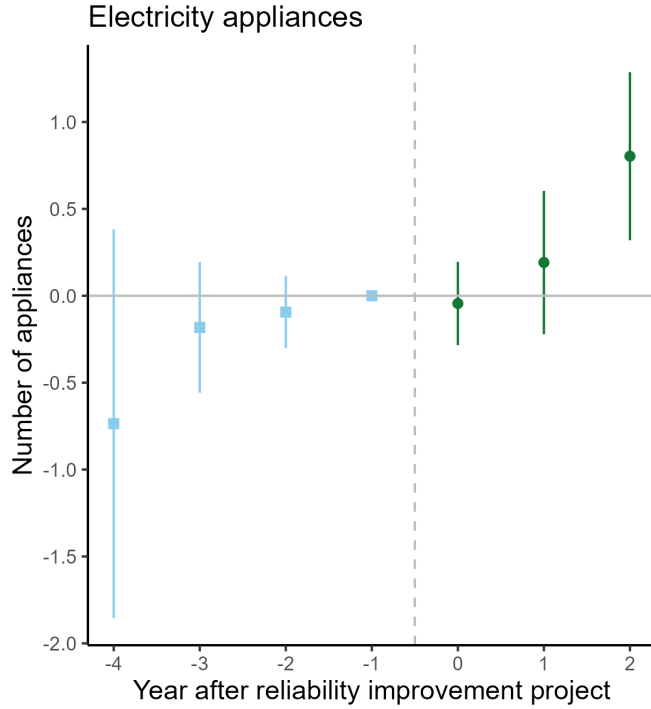
Figure 43: Impact of Reliability Projects on Duration and Frequency of Outages (with Balanced Exposure to Treatment)



Notes: This figure shows the coefficients  $\lambda_k$  in regression (1) and their corresponding confidence intervals, with the duration and frequency of outages as the outcome variables and the estimation done using Callaway and Sant'Anna, (2021). The regressions include the control variables specified in regression (1). All coefficients are estimates relative to a baseline value for the year preceding project date. The left-hand side graph shows the duration of outages in terms of total hours of outages in month. The right-hand side graph shows the frequency of outages in terms of number of outages in month. For both graphs, the x-axis shows the number of years that has passed since the completion of the reliability project. The vertical dashed grey line corresponds to the year of project date. Data used for estimation is a balanced sample of treated stations with outage observations for all time periods of the analysis, from 5 years before project date to 4 years after project date. The estimation is done at the distribution station level. Only stations with exposure to treatment of two years are included in the estimation on the left panel. Only stations with exposure to treatment of three years are included in the estimation on the right panel.

Figure 44 replicates Figure 5 but includes, in its estimation, only households served by stations with exposure to treatment of three years. I also see the trends and coefficient estimates are broadly similar to those of Figure 5.

Figure 44: Impact of Reliability Projects on Electric Appliances (with Balanced Exposure)



Notes: This figure shows the coefficients  $\lambda_t$  in regression (3) and their corresponding confidence intervals. All coefficients are estimates relative to a baseline value for the year preceding project date. The left-hand side graph shows the total number of electric appliances (TV, phone, laptop, fan, etc) while the right hand side show the total number of non-electric appliances (bicycle, car, radio, bed, tables, sofa, etc.). For both graphs, the x-axis shows the number of years that has passed since the reliability project. The vertical dashed grey line corresponds to the year of project completion. Data used for estimation is a balanced sample of treated households with non-missing values for all time periods of the analysis, from 6 years before project date to 4 years after project date. Sample is also balanced on exposure to treatment. The estimation is done at the household level.

## 11 Additional Results

In this section, I present additional results on the impact of improved reliability on the location of consumers and on time use.

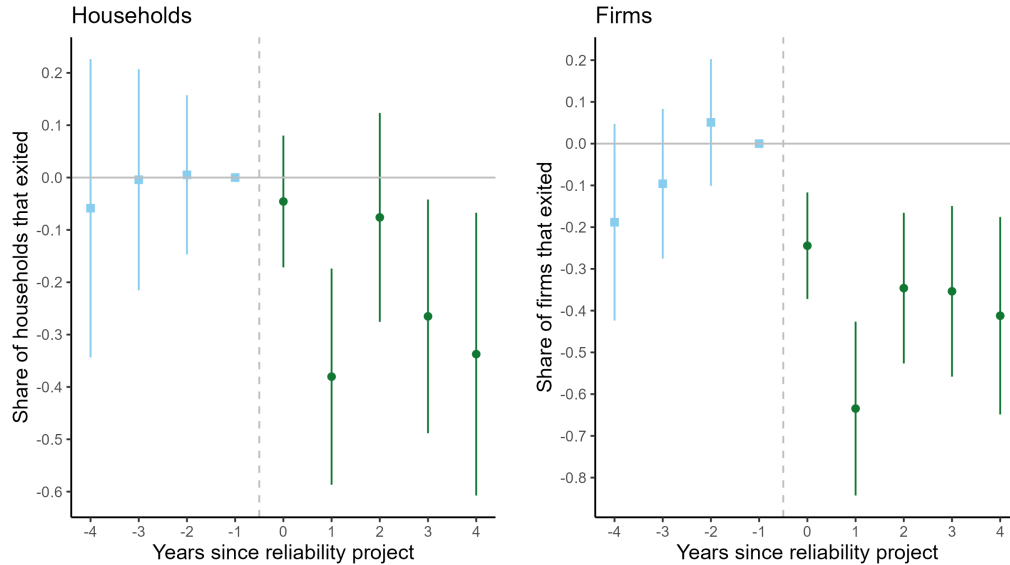
### Location of Consumers

I look at the effect of the reliability projects on the number of clients connected to the grid and on the location of consumers to assess whether reliability changes affect grid connections on the extensive margin.

Figure 45 shows the impact of reliability projects on the likelihood of households and firms leaving their existing place. For any given year  $t$ , I first construct a variable to indicate whether a given customer stays in the same place and is served by the same

distribution station as in previous years. The variable is a dummy variable that takes the value of 1 if a consumer who received bills in year  $t - 1$  did not receive a bill in year  $t$  at the same place and 0 otherwise. I then aggregate these values at the distribution station level to obtain the number of existing clients at  $t - 1$  who exited at  $t$ . The logarithmic transformation of these aggregate values represents the outcome variable in Figure 45. The left-hand side of the figure is for residential clients (or households), while the right-hand side of the figure is for commercial clients (or firms).

Figure 45: Impact of Reliability Projects on the Exit Rate of Households and Firms



Notes: This figure shows the coefficients  $\lambda_l$  in regression (1) and their corresponding confidence intervals. All coefficients are estimates relative to a baseline value for the year preceding the project date. The left-hand side graph shows the likelihood of existing households out-migrating or exiting their market. The right-hand side graph shows the likelihood of existing firms out-migrating or exiting their market. The x-axis shows the number of years that have passed since the completion of the reliability project. The vertical dashed gray line corresponds to the year of the project date. The data used for estimation are geocoded billing data, where each observation is a customer of the utility company. The estimation is performed at the level of the distribution station. Standard errors are clustered at the feeder level.

Figure 45 shows a decrease in the out-migration rate of firms and households served by distribution stations post-reliability project. The decrease is approximately 3% relative to the exit rate in the year preceding project completion for households, and is slightly higher for firms than for households.

Consumers who do not receive bills at the same place may experience one of two things: (i) they could either migrate away to a different place, obtain a new account, and be considered a different/new client in the dataset, or (ii) they could stop their electricity use and exit the market altogether. In this context, I also have data on consumer names and phone numbers. While both of these variables are fuzzy and could change even for the same entity, I use them to show that in this context, firms that do not stay in the same location likely exit the market.

These results suggest that improved reliability encourages firms and households to stay in the same place, instead of migrating away or stopping their use of grid electricity altogether.<sup>42</sup>

### Time Use

Table 18 shows the results of regressions of household economic outcomes on a dummy variable for whether the neighborhood is served by a station that received a reliability project. While I do not find any significant change in the employment status of workers who live in areas that are upgraded, I find that reliability projects led to an increase in hours worked and to a decrease in hours spent on household chores, in hours spent taking care of other household members, and in hours spent searching for wood. These findings are quite meaningful, especially because it has been notoriously difficult to document shifts in time use in low-income countries (Charnes, 2021). Table 18 also confirms the results shown in Figures 3 and 4 on the effect of reliability projects on electricity use and expenditures with a dataset different from the one used for these two figures.

---

<sup>42</sup>This analysis includes only firms and households that were previously clients of the utility company, and it does not capture the potential creation of new businesses in areas with improved reliability.

Table 18: Impact of Reliability Projects on Time Use

| Outcome variable                             | Estimate | S.E.  | P-Value | Obs. | DV Mean |
|----------------------------------------------|----------|-------|---------|------|---------|
| <b>Electricity</b>                           |          |       |         |      |         |
| Monthly hours of outages                     | -10.444  | 2.719 | 0.000   | 7114 | 24.844  |
| Log of electricity expenditures ('000s FCFA) | 1.419    | 0.694 | 0.042   | 5186 | 11.393  |
| <b>Employment</b>                            |          |       |         |      |         |
| Self-employment                              | -0.002   | 0.010 | 0.885   | 7113 | 0.148   |
| Employed by others                           | -0.008   | 0.009 | 0.329   | 7113 | 0.108   |
| <b>Time use</b>                              |          |       |         |      |         |
| Hours worked daily                           | 1.449    | 0.620 | 0.020   | 3104 | 11.950  |
| Daily hours spent on household chores        | -0.746   | 0.294 | 0.012   | 7065 | 5.374   |
| Daily hours spent on child and elderly care  | -0.359   | 0.190 | 0.059   | 7104 | 1.641   |
| Daily hours spent on market transactions     | -1.675   | 1.470 | 0.255   | 7114 | 1.238   |
| Daily hours spent searching for water        | -0.003   | 0.054 | 0.955   | 7114 | 0.657   |
| Daily hours spent searching for wood         | -0.107   | 0.047 | 0.023   | 7114 | 0.422   |

Notes: This table shows the results of regressions of different time use indicators of households on a dummy variable for whether the neighborhood is served by a station that received a reliability project. Each line corresponds to a separate regression where the outcome variable is regressed on a dummy variable for reliability projects and shows the statistics associated with this dummy variable. Analysis is restricted to households connected to the grid and present in all rounds of the household panel data. Household fixed effects and year fixed effects are included in all regressions. Standard errors are clustered at the neighborhood level.

## 12 Model

### 12.1 Model Propositions

In this subsection, I give details on the derivation of proposition 1 presented in the main text, under the assumption that energy efficiency are endogenous (i.e. energy efficiency depends on reliability.)

The goal of this subsection is to derive a closed-form expression for the change in electricity use  $x_{i1lt}$  when electricity reliability  $\rho_{lt}$  changes. I start with the first-order condition (FOC) from maximizing the utility function (4) is:

$$\mu'_i(\tau_{ilt}(\rho_{lt}) \cdot x_{i1lt}) \cdot \tau_{ilt}(\rho_{lt}) = e_t$$

I drop the subscripts and set  $z = \tau(\rho) \cdot x$  for ease of notation. The total differentiation of the FOC with respect to  $x$  and  $\rho$  gives:

$$\begin{aligned}
d[\mu'(z) \cdot \tau(\rho)] &= 0 \\
\mu'(z) \cdot d\tau(\rho) + \tau(\rho) \cdot d\mu'(z) &= 0 \\
\mu'(z) \cdot [\tau'(\rho) \cdot d\rho] + \tau(\rho) \cdot \mu''(z) [\tau(\rho) \cdot dx + x \cdot \tau'(\rho) \cdot d\rho] &= 0 \\
\mu'(z) \tau'(\rho) d\rho + \tau(\rho) \mu''(z) [\tau(\rho) dx + x \tau'(\rho) d\rho] &= 0 \\
\mu'(z) \tau'(\rho) d\rho + \tau(\rho) \mu''(z) \tau(\rho) dx + \tau(\rho) \mu''(z) x \tau'(\rho) d\rho &= 0 \\
[\mu'(z) \tau'(\rho) + \tau(\rho) \mu''(z) x \tau'(\rho)] d\rho + \tau^2(\rho) \mu''(z) dx &= 0 \\
[\mu'(z) + \tau(\rho) \mu''(z) x] \tau'(\rho) d\rho + \tau^2(\rho) \mu''(z) dx &= 0 \\
[\mu'(z) + \mu''(z) z] \tau'(\rho) d\rho + \tau^2(\rho) \mu''(z) dx &= 0
\end{aligned}$$

I then get the following:

$$\frac{\partial x}{\partial \rho} = -\frac{[\mu'(z) + \mu''(z) z] \tau'(\rho)}{\tau^2(\rho) \mu''(z)}$$

I then turn  $\frac{\partial x}{\partial \rho}$  into elasticities as follows:

$$\begin{aligned}
\frac{\partial x}{\partial \rho} &= -\frac{[\mu'(z) + \mu''(z) z] \tau'(\rho)}{\tau^2(\rho) \mu''(z)} \\
\frac{\partial x}{\partial \rho} \cdot \frac{\rho}{x} &= -\frac{[\mu'(z) + \mu''(z) z] \tau'(\rho)}{\tau^2(\rho) \mu''(z)} \cdot \frac{\rho}{x} \\
\varepsilon_{x,\rho} &= -\frac{[\mu'(z) + \mu''(z) z]}{\mu''(z) z} \cdot \frac{\tau'(\rho) \rho}{\tau(\rho)} \\
\varepsilon_{x,\rho} &= -\frac{\mu'(z) + \mu''(z) z}{\mu''(z) z} \cdot \varepsilon_{\tau,\rho} \\
\varepsilon_{x,\rho} &= -\frac{\mu'(z) + \mu''(z) z}{\mu''(z) z} \cdot \varepsilon_{\tau,\rho} \\
\varepsilon_{x,\rho} &= -\left(\frac{\mu'(z)}{\mu''(z) z} + 1\right) \cdot \varepsilon_{\tau,\rho} \\
\varepsilon_{x,\rho} &= -\left(\frac{1}{\varepsilon_{\mu',z}} + 1\right) \cdot \varepsilon_{\tau,\rho}
\end{aligned}$$

Here,  $\varepsilon_{\tau,\rho}$  corresponds to the elasticity of energy efficiency with respect to electricity reliability.  $\varepsilon_{\mu',z} = \mu'' \frac{z}{\mu'}$  measures the responsiveness of the marginal utility  $\mu'(z)$  with respect to changes in  $z$ . It reflects the concavity of the marginal utility function, or how quickly marginal utility decreases with consumption. By construction,  $\varepsilon_{\mu',z}$  is always

non-positive since  $\mu' > 0$  and  $\mu'' < 0$ .  $\varepsilon_{x,\rho}$  measures the elasticity of electricity use with respect to electricity reliability.

I obtain the following proposition:

**Proposition 2** *For any instant  $m$ ,  $\varepsilon_{x,\rho} > 0$  iff  $\varepsilon_{\mu',z} > -1$  and  $\varepsilon_{\tau,\rho} > 0$ , or  $\varepsilon_{\mu',z} < -1$  and  $\varepsilon_{\tau,\rho} < 0$ .*

Intuitively, the behavioral change in electricity use that a consumer experiences as a result of a change in electricity reliability depends on two factors: (i) the change in the efficiency of existing appliances and (ii) the change in the marginal utility of electrical services. An improvement in reliability may lead to two changes: (i) an increase in the efficiency of existing appliances, which causes demand for electricity from existing appliances to decrease, and (ii) a more intense use of appliances since the marginal utility of electrical services is always positive. The increase in the efficiency of existing appliances would come from an increased usability of existing appliances due to reduced power outages and voltage fluctuations. The positive marginal utility would induce consumers to use more electricity through a more intense use of existing appliances or through the purchase of new appliances.

Specifically, under the more plausible assumption that  $\varepsilon_{\tau,\rho} > 0$  (i.e., energy efficiency improves with reliability), this proposition states that the sign of  $\varepsilon_{x,\rho}$  will depend on the magnitude of  $\varepsilon_{\mu',z}$ . In particular, a value of  $\varepsilon_{\mu',z}$  between -1 and 0 indicates that marginal utility decreases with  $z$  but not too rapidly. A marginal utility function that is not too concave means consumers still gain significant additional utility from consuming more electricity, even at higher levels of energy efficiency or electricity use. In this case, marginal utility is not too concave (i.e.,  $-1 < \varepsilon_{\mu',z} < 0$ ), and an increase in  $\rho$  positively affects electricity usage  $x$ . In contrast, if the marginal change in utility decreases, then consumers might prefer to decrease consumption despite efficiency gains from reliability.

## 12.2 Deriving Consumer Surplus

To calculate the consumer surplus (CS) derived from electricity consumption  $x$ , we integrate the inverse demand function and subtract total expenditure. The inverse demand function, derived from the first-order condition, is:

$$e(x) = \alpha\tau^{1-\gamma}x^{-\gamma} \quad (37)$$

where  $e(x)$  is the price of electricity at consumption level  $x$ ,  $\alpha$  is the preference parameter,  $\tau$  is the technical efficiency parameter and  $\gamma$  is the coefficient of relative risk aversion.

The consumer surplus is given by:

$$CS = \int_0^x e(q) dq - e(x) \cdot x \quad (38)$$

First, we compute the area under the demand curve up to quantity  $x$ :

$$\int_0^x e(q) dq = \int_0^x \alpha \tau^{1-\gamma} q^{-\gamma} dq \quad (39)$$

$$= \alpha \tau^{1-\gamma} \int_0^x q^{-\gamma} dq. \quad (40)$$

The integral of  $q^{-\gamma}$  with respect to  $q$  is:

$$\int q^{-\gamma} dq = \frac{q^{1-\gamma}}{1-\gamma} \quad \text{for } \gamma \neq 1 \quad (41)$$

Evaluating the definite integral from  $q = 0$  to  $q = x$ :

$$\int_0^x q^{-\gamma} dq = \left[ \frac{q^{1-\gamma}}{1-\gamma} \right]_{q=0}^{q=x} \quad (42)$$

$$= \frac{x^{1-\gamma}}{1-\gamma} - \lim_{q \rightarrow 0^+} \frac{q^{1-\gamma}}{1-\gamma} \quad (43)$$

When  $\gamma < 1$ , the term  $q^{1-\gamma}$  approaches zero as  $q \rightarrow 0^+$ , so the lower limit evaluates to zero. Thus, the integral simplifies to:

$$\int_0^x q^{-\gamma} dq = \frac{x^{1-\gamma}}{1-\gamma} \quad (44)$$

Substituting back into equation (39):

$$\int_0^x e(q) dq = \alpha \tau^{1-\gamma} \frac{x^{1-\gamma}}{1-\gamma}. \quad (45)$$

The total expenditure on electricity at consumption level  $x$  is:

$$e(x) \cdot x = (\alpha \tau^{1-\gamma} x^{-\gamma}) x = \alpha \tau^{1-\gamma} x^{1-\gamma} \quad (46)$$

Substituting equations (45) and (46) into the consumer surplus expression (38):

$$CS = \alpha \tau^{1-\gamma} \frac{x^{1-\gamma}}{1-\gamma} - \alpha \tau^{1-\gamma} x^{1-\gamma} \quad (47)$$

$$= \alpha \tau^{1-\gamma} x^{1-\gamma} \left( \frac{1}{1-\gamma} - 1 \right) \quad (48)$$

Simplifying the expression inside the parentheses:

$$\frac{1}{1-\gamma} - 1 = \frac{1 - (1-\gamma)}{1-\gamma} = \frac{1-1+\gamma}{1-\gamma} = \frac{\gamma}{1-\gamma} \quad (49)$$

Substituting back into equation (47) yields:

$$CS = \alpha \tau^{1-\gamma} x^{1-\gamma} \cdot \frac{\gamma}{1-\gamma}. \quad (50)$$

### 12.3 From Economic Model to Discrete Choice Model

The goal of this subsection is to start from the economic model of electricity demand and make appropriate assumptions and transformations to derive a discrete choice utility function that captures consumer preferences over different electricity service options. Each component of the final utility function corresponds to variables and parameters in the economic model, allowing for coherent and interpretable estimation of consumer behavior in the context of unreliable electricity supply.

#### Step 1: Optimal Electricity and Service Consumption

I first find the optimal electricity consumption  $x_{i1lt}^*$  that maximizes utility subject to the budget constraint in Equation (4). I set up the Lagrangian:

$$\mathcal{L} = \rho_{lt} \cdot \mu_i(\tau_{ilt}(\rho_{lt}) \cdot x_{i1lt}) + y_{ilt} - \lambda(e_t \cdot \rho_{lt} \cdot x_{i1lt} + y_{ilt} - w_{ilt}). \quad (51)$$

Taking the first-order conditions yields:

$$\frac{\partial \mathcal{L}}{\partial x_{i1lt}} = \rho_{lt} \cdot \mu'_i(\tau_{ilt}(\rho_{lt}) \cdot x_{i1lt}) \cdot \tau_{ilt}(\rho_{lt}) - \lambda \cdot e_t \cdot \rho_{lt} = 0, \quad (52)$$

$$\frac{\partial \mathcal{L}}{\partial y_{ilt}} = 1 - \lambda = 0. \quad (53)$$

From Equation (53), I find  $\lambda = 1$ . Substituting  $\lambda = 1$  into Equation (52) and simplifying, I get:

$$\mu'_i(\tau_{ilt}(\rho_{lt}) \cdot x_{i1lt}) \cdot \tau_{ilt}(\rho_{lt}) = e_t. \quad (54)$$

Solving for  $x_{i1lt}$  yields:

$$\mu'_i(s_i) = \frac{e_t}{\tau_{ilt}(\rho_{lt})}, \quad \text{where } s_i = \tau_{ilt}(\rho_{lt}) \cdot x_{i1lt} \quad (55)$$

As a reminder,  $s_i$  denote the total electrical services consumed. The inverse of the marginal utility function gives:

$$s_i = (\mu'_i)^{-1} \left( \frac{e_t}{\tau_{ilt}(\rho_{lt})} \right) \quad (56)$$

Thus, the optimal electricity consumption is:

$$x_{i1lt}^* = \frac{s_i}{\tau_{ilt}(\rho_{lt})} = \frac{1}{\tau_{ilt}(\rho_{lt})} \cdot (\mu'_i)^{-1} \left( \frac{e_t}{\tau_{ilt}(\rho_{lt})} \right) \quad (57)$$

Substituting the optimal  $x_{i1lt}^*$  into the utility function, I have:

$$u_{ilt} = \rho_{lt} \cdot \mu_i(s_i) + y_{ilt} - e_t \cdot \rho_{lt} \cdot x_{i1lt}^* \quad (58)$$

Here, I used the budget constraint to replace  $y_{ilt}$  with  $y_{ilt} - e_t \cdot \rho_{lt} \cdot x_{i1lt}^*$ .

## Step 2: Specific Utility Function

For analytical tractability, I assume that the utility function  $\mu_i(\cdot)$  is logarithmic:

$$\mu_i(s) = \ln(s) \quad (59)$$

Thus, the marginal utility and its inverse are:

$$\mu'_i(s) = \frac{1}{s}, \quad \text{so} \quad (\mu'_i)^{-1}(z) = \frac{1}{z} \quad (60)$$

Using Equation (56), I find:

$$s_i = (\mu'_i)^{-1} \left( \frac{e_t}{\tau_{ilt}(\rho_{lt})} \right) = \frac{\tau_{ilt}(\rho_{lt})}{e_t} \quad (61)$$

Therefore, from Equation (57), the optimal electricity consumption simplifies to:

$$x_{i1lt}^* = \frac{s_i}{\tau_{ilt}(\rho_{lt})} = \frac{1}{e_t} \quad (62)$$

Substituting  $s_i$  and  $x_{i1lt}^*$  back into the utility function:

$$u_{ilt} = \rho_{lt} \cdot \mu_i(s_i) + y_{ilt} - e_t \cdot \rho_{lt} \cdot x_{i1lt}^* \quad (63)$$

$$= \rho_{lt} \cdot \ln \left( \frac{\tau_{ilt}(\rho_{lt})}{e_t} \right) + y_{ilt} - \rho_{lt} \quad (64)$$

Here, I used the fact that  $e_t \cdot x_{i1lt}^* = 1$ .

### Step 3: From Economic Model to Discrete Choice Model

Consumers choose among different electricity service options indexed by  $j$ . More specifically, consumer  $i$  can choose among the following  $j$  options:

$$j = \begin{cases} 0 & \text{if consumer } i \text{ is not connected to the grid and has no appliances} \\ 1 & \text{if consumer } i \text{ is connected to the grid with only low-intensity appliances} \\ 2 & \text{if consumer } i \text{ is connected to the grid with low-intensity and high-intensity appliances} \\ 3 & \text{if consumer } i \text{ gets electricity from solar panels only} \end{cases}$$

Each option  $j$  is characterized by reliability  $\rho_{ljt}$  (availability of electricity under option  $j$  in location  $l$  at time  $t$ ), electricity price  $e_{jt}$ , appliance efficiency  $\tau_{iljt}$  (efficiency of appliances under option  $j$ ), connection cost  $c_{jt}$ , and appliance cost  $p_{ljt}$ .

First, since  $y_{ilt}$  is constant across choices and does not affect the utility difference when comparing options, I can omit it from the utility function for comparative purposes. Therefore, the utility simplifies to:

$$u_{ilt} = \rho_{lt} \cdot (\ln(\tau_{ilt}(\rho_{lt})) - \ln(e_t) - 1). \quad (65)$$

I then update the utility function to reflect these options:

$$u_{iljt} = \rho_{ljt} \cdot (\ln(\tau_{iljt}) - \ln(e_{jt}) - 1) - c_{jt} - p_{ljt}. \quad (66)$$

Recognizing that reliability and appliance efficiency may vary between peak and off-peak hours, let:

$$\begin{aligned} h_{1,ljt} &= \text{Hours of electricity availability during peak hours,} \\ h_{2,ljt} &= \text{Hours of electricity availability during off-peak hours.} \end{aligned}$$

Total reliability is  $\rho_{ljt} = h_{1,ljt} + h_{2,ljt}$ . Similarly, let:

$$\begin{aligned} \tau_{iljt}^{\text{peak}} &= \text{Appliance efficiency during peak hours,} \\ \tau_{iljt}^{\text{off}} &= \text{Appliance efficiency during off-peak hours.} \end{aligned}$$

I rewrite the utility function to separate peak and off-peak components:

$$\begin{aligned} u_{iljt} &= h_{1,ljt} \cdot \left( \ln(\tau_{iljt}^{\text{peak}}) - \ln(e_{jt}) - 1 \right) \\ &\quad + h_{2,ljt} \cdot \left( \ln(\tau_{iljt}^{\text{off}}) - \ln(e_{jt}) - 1 \right) - c_{jt} - p_{ljt} \end{aligned} \quad (67)$$

I define the coefficients for the discrete choice model:

$$\beta_{1,iljt} = \ln(\tau_{iljt}^{\text{peak}}) - \ln(e_{jt}) - 1 \quad (68)$$

$$\beta_{2,iljt} = \ln(\tau_{iljt}^{\text{off}}) - \ln(e_{jt}) - 1 \quad (69)$$

The utility function becomes:

$$u_{iljt} = \beta_{1,iljt} \cdot h_{1,ljt} + \beta_{2,iljt} \cdot h_{2,ljt} - c_{jt} - p_{ljt}. \quad (70)$$

To include the electricity price  $e_{jt}$  explicitly, I linearize the logarithmic term using a first-order Taylor approximation around a reference price  $\bar{e}$ :

$$\ln(e_{jt}) \approx \ln(\bar{e}) + \frac{1}{\bar{e}}(e_{jt} - \bar{e}) \quad (71)$$

Substituting back, I update the coefficients:

$$\beta_{1,iljt} \approx \ln(\tau_{iljt}^{\text{peak}}) - \ln(\bar{e}) - \frac{e_{jt} - \bar{e}}{\bar{e}} - 1, \quad (72)$$

$$\beta_{2,iljt} \approx \ln(\tau_{iljt}^{\text{off}}) - \ln(\bar{e}) - \frac{e_{jt} - \bar{e}}{\bar{e}} - 1 \quad (73)$$

Let  $\gamma = \frac{1}{\bar{e}}$ , so the price effect is linear in  $e_{jt}$ .

The utility function now becomes:

$$\begin{aligned} u_{iljt} = & \left( \ln(\tau_{iljt}^{\text{peak}}) - \ln(\bar{e}) - 1 - \gamma(e_{jt} - \bar{e}) \right) h_{1,ljt} \\ & + \left( \ln(\tau_{iljt}^{\text{off}}) - \ln(\bar{e}) - 1 - \gamma(e_{jt} - \bar{e}) \right) h_{2,ljt} - c_{jt} - p_{ljt}. \end{aligned} \quad (74)$$

Combining terms and recognizing that constants can be absorbed into fixed effects, I define:

$$\beta_{3,ilt} = \ln(\tau_{iljt}^{\text{peak}}) - \ln(\bar{e}) - 1 + \gamma\bar{e} \quad (75)$$

$$\beta_{4,ilt} = \ln(\tau_{iljt}^{\text{off}}) - \ln(\bar{e}) - 1 + \gamma\bar{e} \quad (76)$$

$$\beta_5 = -\gamma \quad (77)$$

Thus, the utility function simplifies to:

$$u_{iljt} = \beta_{3,ilt} h_{1,ljt} + \beta_{4,ilt} h_{2,ljt} + \beta_5 e_{jt} (h_{1,ljt} + h_{2,ljt}) - c_{jt} - p_{ljt} \quad (78)$$

#### Step 4: Final Econometric Adjustments

To account for unobserved factors, I include fixed effects and an error term:

$$u_{iljt} = \beta_{3,ilt}h_{1,ljt} + \beta_{4,ilt}h_{2,ljt} + \beta_5e_{jt} + \beta_6c_{jt} + \beta_7p_{ljt} + \xi_{jt} + \xi_l + \xi_{ij} + \epsilon_{iljt} \quad (79)$$

I also make two additional adjustments. First, I unrestrict  $\beta_6$  and  $\beta_7$ , the coefficients on connection cost and appliance cost, respectively, to allow them to be estimated from the data. In the above derivation, coefficients on connection cost and appliance cost were set to  $-1$  for simplicity. However, there is no theoretical necessity to fix these coefficients to  $-1$ . Instead, I allow  $\beta_6$  and  $\beta_7$  to be parameters to be estimated. This flexibility acknowledges that consumers may value connection costs and appliance costs differently, and their sensitivities can be empirically determined. Second, I drop the interaction term on electricity costs to make the coefficient easier to interpret.

Our final utility function, suitable for estimation in a discrete choice framework, is:

$$u_{iljt} = \beta_{3,ilt}h_{1,ljt} + \beta_{4,ilt}h_{2,ljt} + \beta_5e_{jt} + \beta_6c_{jt} + \beta_7p_{ljt} + \xi_{jt} + \xi_l + \xi_{ij} + \epsilon_{iljt} \quad (80)$$

Each coefficient corresponds to a specific component from the economic model.  $\beta_{3,ilt}$  is the marginal utility of electricity availability during peak hours, incorporating appliance efficiency and base electricity price.  $\beta_{4,ilt}$  is the marginal utility of electricity availability during off-peak hours.  $\beta_5$  is the marginal disutility of electricity price  $e_{jt}$  per hour of availability.  $\beta_6$  is the coefficient on connection costs  $c_{jt}$ ,  $\beta_7$  is the coefficient on appliance costs  $p_{ljt}$ .

The fixed effects  $\xi_{jt}$ ,  $\xi_l$ , and  $\xi_{ij}$  capture unobserved attributes specific to products, locations, and individual preferences, respectively. The error term  $\epsilon_{iljt}$  accounts for random shocks and unobserved factors.

## 12.4 Calibration of Energy Efficiency Factor

This subsection shows an example of the energy efficiency factor is calculated using the billing data and engineering estimates from external reports.

I begin by determining the efficiency improvements (EI) for each appliance type. For lighting, the transition from incandescent bulbs to compact fluorescent lamps (CFLs) and light-emitting diodes (LEDs) significantly enhances efficiency. Incandescent bulbs have an efficiency of approximately 15 lumens per watt (lm/W), whereas CFLs and LEDs average around 75 lm/W. The efficiency improvement for lighting is thus calculated as:

$$EI_{\text{Light}} = \frac{75 \text{ lm/W}}{15 \text{ lm/W}} = 5.$$

For refrigerators, modern energy-efficient models consume about 300 kWh per year compared to older models consuming around 500 kWh per year. The efficiency improvement for refrigerators is calculated as:

$$EI_{\text{Refrigerator}} = \frac{500 \text{ kWh/year}}{300 \text{ kWh/year}} \approx 1.67.$$

In the case of televisions, the shift from cathode ray tube (CRT) models to LED televisions leads to energy savings. Assuming average power consumption of 100 W for CRT TVs and 60 W for LED TVs, the efficiency improvement is:

$$EI_{\text{TV}} = \frac{100 \text{ W}}{60 \text{ W}} \approx 1.67.$$

Next, I obtain usage weights (UW) to each appliance based on their proportion of total household electricity consumption using the baseline appliance composition data obtained from the utility company.<sup>43</sup> This baseline appliance dataset allows me to get a sense of usage profile at the consumer level. Below, I show usage weights for a consumer for whom lighting accounts for approximately 50% of electricity usage (due, for example, to extended use during evening hours and its ubiquity in households), Refrigerators for 30% of consumption (as they operate continuously in households that own them) and televisions representing around 20% of consumption.

With the efficiency improvements and usage weights established, I calculate the contribution (C) of each appliance to the technical efficiency parameter  $\tau$  of this consumer. For lighting, the contribution is:

$$C_{\text{Light}} = EI_{\text{Light}} \times UW_{\text{Light}} = 5 \times 0.50 = 2.50$$

For refrigerators:

$$C_{\text{Refrigerator}} = EI_{\text{Refrigerator}} \times UW_{\text{Refrigerator}} = 1.67 \times 0.30 = 0.50$$

For televisions:

$$C_{\text{TV}} = EI_{\text{TV}} \times UW_{\text{TV}} = 1.67 \times 0.20 \approx 0.33$$

Summing these contributions provides an initial estimate of the technical efficiency parameter:

$$\tau = 2.50 + 0.50 + 0.33 = 3.33$$

However, this consumer may not have adopted energy-efficient appliances. To adjust for market penetration rates, I consider the adoption rates of efficient technologies in Senegal. In the calibration, I use a uniform distribution to model adoption rates given

<sup>43</sup>See Section 3 for more details on this dataset.

the lack of data on this. Below, I show calculations assuming that this consumer has 60% of their lighting appliances as CFLs or LEDs, 40% of energy-efficient refrigerators, and 50% of LED televisions. The adjusted efficiency (AE) improvement for lighting for this consumer becomes is calculated as:

$$AE_{\text{Light}} = 1 + 0.60 \times (EI_{\text{Light}} - 1) = 1 + 0.60 \times (5 - 1) = 3.4$$

The adjusted contribution (AC) to  $\tau$  for lighting becomes:

$$AC_{\text{Light}} = AE_{\text{Light}} \times UW_{\text{Light}} = 3.4 \times 0.50 = 1.70$$

Similarly, for refrigerators:

$$AE_{\text{Refrigerator}} = 1 + 0.40 \times (EI_{\text{Refrigerator}} - 1) = 1 + 0.40 \times (1.67 - 1) \approx 1.27$$

$$AC_{\text{Refrigerator}} = AE_{\text{Refrigerator}} \times UW_{\text{Refrigerator}} = 1.27 \times 0.30 \approx 0.38$$

For televisions:

$$AE_{\text{TV}} = 1 + 0.50 \times (EI_{\text{TV}} - 1) = 1 + 0.50 \times (1.67 - 1) \approx 1.335$$

$$AC_{\text{TV}} = AE_{\text{TV}} \times UW_{\text{TV}} = 1.335 \times 0.20 \approx 0.27.$$

By summing the adjusted contributions, I obtain the calibrated value of the technical efficiency parameter for this consumer:

$$\tau = 1.70 + 0.38 + 0.27 = 2.35$$

## 12.5 Elasticities of Electricity Demand

In this section, I present the estimation of the short-run price elasticities of electricity demand.

In this sub-section, I estimate the price elasticity of demand using specific contextual details of the electricity market that I study. I start with the following regression to estimate price elasticity of electricity demand:

$$\ln x_{ilt} = \alpha_i + \mu_{lt} + \lambda_{lg} \ln P_{ilt} \cdot \mathbb{1}\{G_{il} = g\} + \Gamma X_{ilt} + \epsilon_{ilt} \quad (81)$$

In equation (81),  $\ln x_{ilt}$  is the log of electricity use (in kWh) for customer  $i$  served by distribution station  $l$  in year  $t$ .  $\alpha_i$  and  $\mu_{lt}$  are customer fixed effects and district-year fixed effects, respectively. The dummy  $\mathbb{1}\{G_{il} = g\}$  takes the value of 1 if the client is classified in group  $g$ . There are 6 groups of customers based on electricity use and usage

type: small residential, medium residential, large residential, small firms, medium firms and large firms.  $X_{ilt}$  represent a set of observable demand shifters: number of rooms in establishments, number of appliances, characteristics of establishment (whether the establishment is an apartment, house, store, etc), characteristics of the client (age, gender, tenure as a client, etc.). Standard errors are clustered at the client level.

$\lambda_{lg}$  expresses the price elasticity of demand, but it is not identified. More formally, if I assume that demand and supply are characterized by the following pair of simultaneous equations:

$$Q = D(X, P, \epsilon) \quad (82)$$

$$P = C(W, Q, V) \quad (83)$$

Here,  $X$  and  $W$  are respectively observable demand and supply shifters,  $P$  and  $Q$  represent price and quantity, and  $U$  and  $V$  are unobserved shocks to demand and to marginal cost, respectively. Here, equation (82) is simply a generalization of equation (81). Even under the assumption that only prices and quantities are endogenous (i.e.,  $(X, W) \perp (U, V)$ ), the presence of  $\epsilon$  and  $V$  causes the typical endogeneity problem in the estimation of demand because of the problem of simultaneity (i.e., prices are correlated with unobserved demand shocks) and mean reversion.

Identification of demand can be obtained through elements of the set of cost shifters  $W$  different from demand shifters  $X$  and excluded from the demand function in equation (82). These cost shifters need to satisfy an appropriate independence condition with respect to  $\epsilon$ .

I need to find valid instruments that exogeneously shift prices and supply schedules in order to identify the demand schedule. In the context, these shifts are especially hard to find because there is one utility company and prices are set nationally, so there is generally no sticky spatial variation in prices that can be exploited. In addition, standard time fixed-effects will soak the temporal variation in prices that I am left with.

I follow the literature on demand estimation in markets where prices are endogeneously determined by consumption levels. In particular, I rely on a simulated instrument using consumption in a period midway between  $t$  and  $t - k$  (Saez et al., 2012; Ito, 2014; Sears, 2022). I consider the following differenced model to estimate the price elasticity of consumption using quasi-exogenous variation in price schedules over time.

$$\Delta \ln(x_{ilt}) = \mu_{lt} + \lambda \Delta \ln(P_{ilt}) + h(Q_{il,t-6}) + \Gamma X_{ilt} + \epsilon_{ilt} \quad (84)$$

$\ln(x_{ilt})$  is the log of electricity consumption for customer  $i$  in location  $l$  at month  $t$  and  $\Delta \ln(x_{ilt}) = \ln(x_{ilt}) - \ln(Q_{il,t-k})$  where  $k = 12$  in the main specification to get a year-long difference. Similarly,  $\ln(P_{ilt})$  is the log price for customer  $i$  in location

$l$  at month  $t$  and  $\Delta \ln(P_{ilt}) = \ln(P_{ilt}) - \ln(P_{il,t-12})$ . The year-long difference eliminates customer-by-month fixed effects.  $\ln(P_{i,t-1})$  is the log of price paid in the previous month. The existing literature on energy consumption documents that households primarily respond to average price, as opposed to marginal price (Ito, 2014; Wichman, 2014). Accordingly, I use average price in the main specification but also test for responses to marginal prices. Because of the year-long difference,  $\epsilon_{ilt}$  is the difference in electricity demand shocks between the current month and the same month in the previous year.  $\mu_{lt}$  is location by month-year fixed effects, which would account for time-varying district-level factors (such as weather) that affect electricity demand, vary by month, and are common to customers within the same locations.  $X_{ilt}$  is the same set of observable demand shifters defined above. The differencing model mechanically removes the time-invariant customer characteristics.

To address issues of price endogeneity and mean reversion, I construct the simulated price instrument as follows:

$$\widehat{\Delta \ln(P_{il,t})} = \ln(P_t(Q_{il,t-6})) - \ln(P_{t-12}(Q_{il,t-6})) \quad (85)$$

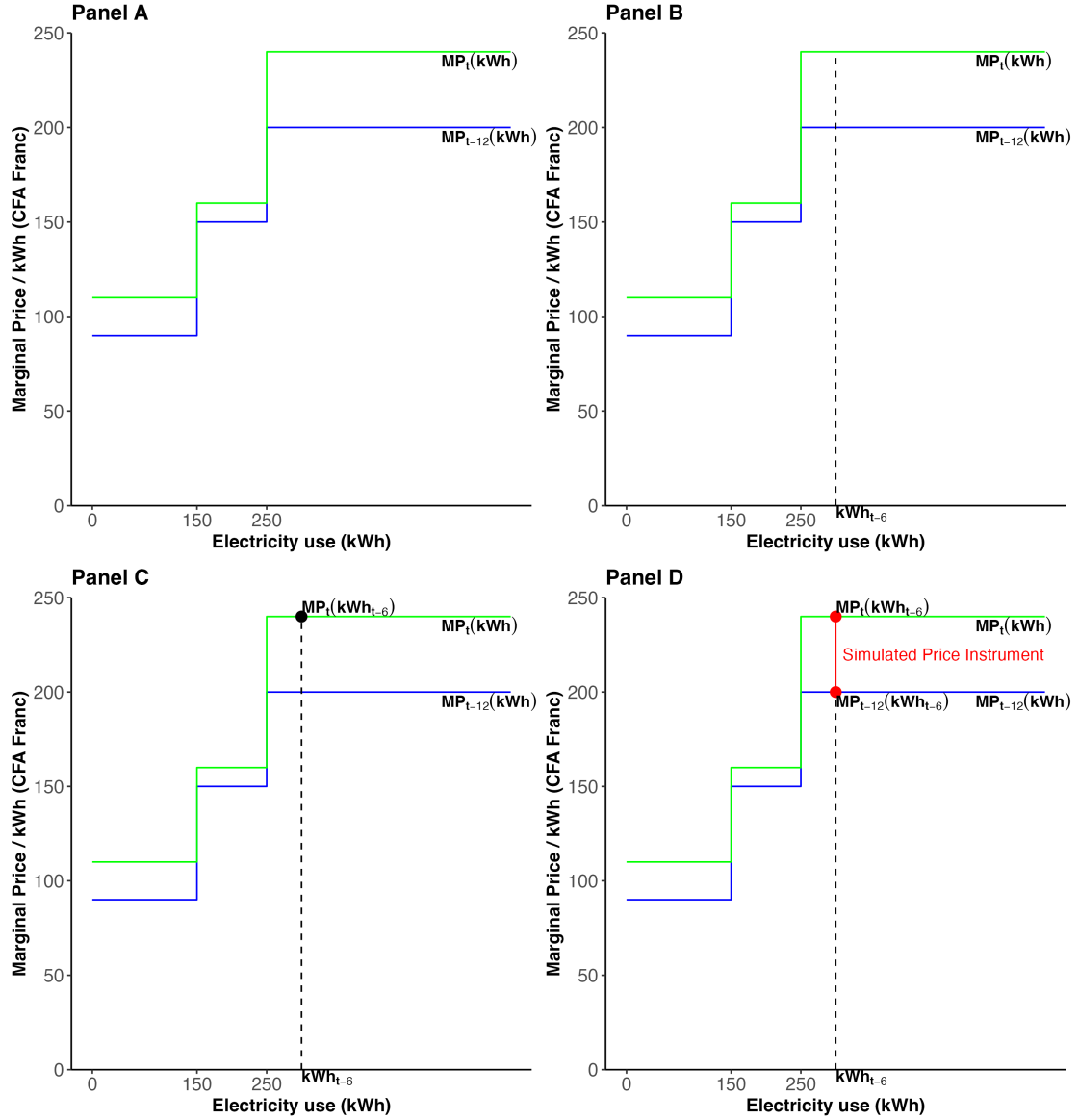
The simulated change in price  $\widehat{\Delta \ln(P_{i,t})}$  is calculated as the difference between: (i) the log of current price for a consumption level equal to that of six months ago, and (ii) the log of price twelve months ago for a consumption level equal to that of six months ago.

Figure 22 show the tariff schedule for the different groups of customers in 2023. In the study period from 2012 to 2024, there have been three price changes: in April 2017, December 2019, and January 2023.<sup>44</sup> Figure 46 shows a visual example of how the simulated price instrument is constructed based on this tariff schedule.

---

<sup>44</sup>Figures 27, 28 and 29 in Appendix C2 show the evolution of the marginal price for the three tiers of the different tariff schedules between 2012 and 2023.

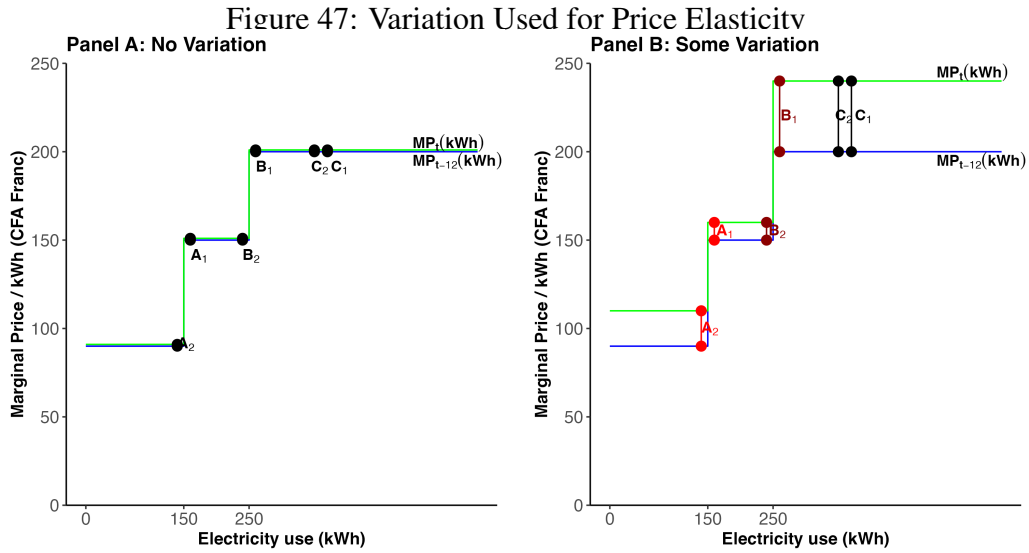
Figure 46: Construction of Simulated Price Instrument



Notes: The four panels in this figure display the steps taken to construct the simulated price instrument. The y-axis for all four panels show the marginal price of 1 kWh, and are given in the local currency of CFA Franc, which is pegged to the Euro (1 US dollar is approximately 550 CFA Franc). The x-axis for all four panels show the electricity use in kWh. Panel A shows the marginal price schedule in the current period  $t$ ,  $MP_t$ , and the marginal price schedule at  $t - 12$  (12 months prior),  $MP_{t-12}$ , that a client faces. Panel B reproduces the figure in Panel A, and adds a vertical line corresponding to the electricity use at time  $t - 6$  (six months prior),  $kWh_{t-6}$ . Panel C reproduces Panel A and B, and adds the marginal price corresponding to  $kWh_{t-6}$  on the marginal price schedule in the current period  $t$ ,  $MP_t(kWh_{t-6})$ . Panel D reproduces Panel A, B, and C and adds the marginal price corresponding to  $kWh_{t-6}$  on the marginal price schedule at  $t - 12$ ,  $MP_{t-12}(kWh_{t-6})$ . The value of the simulated price instrument for this consumer is shown in red and corresponds to:  $MP_t(kWh_{t-6}) - MP_{t-12}(kWh_{t-6})$ .

Figure 47 shows a few examples of the variation exploited by the simulated price instrument. Panel A of figure 47 shows the simulated price instrument for a scenario

in which the marginal price schedule does not change between  $t - 12$  and  $t$ . For this scenario, all simulated price instrument are 0, by construction and there is no variation in price left for the estimation. Panel B of figure 47 shows the simulated price instrument in a scenario in which the marginal price schedule does change between  $t - 12$  and  $t$ . In this case, all simulated price instruments are positive, though only the simulated price instruments for the pairs  $\{A_1, A_2\}$  and  $\{B_1, B_2\}$  provide us with price variation that can be used for the estimation. The estimation of the regression parameters thus relies on the within-location difference in price between consumers in the same consumption percentiles using the same meter type (prepaid vs. postpaid) and subject to the same usage type (residential vs. commercial). This variation thus picks up a difference in price faced by consumers living in the same area, with similar consumption levels at  $t - 6$  but who have different values for the simulated price instrument.



Notes: The two panels in this figure display the simulated price instruments for three pairs of consumers with similar consumption levels at  $t - 6$ , and who could fall under the same decile of consumption in the regression (87).  $A_1$  and  $A_2$  are clients electricity use of 140 kWh and 160 kWh at  $t - 6$  respectively, and are thus comparable.  $B_1$  and  $B_2$  are clients with electricity use of 240 kWh and 260 kWh at  $t - 6$  respectively, and are thus comparable.  $C_1$  and  $C_2$  are clients with electricity use of 350 kWh and 370 kWh at  $t - 6$  respectively, and are thus comparable. The staircase function in blue corresponds to the marginal price schedule in the current period  $t$ ,  $MP_t$ , and staircase function in green the marginal price schedule at  $t - 12$  (12 months prior),  $MP_{t-12}$ . Each colored vertical line corresponds to the value of a simulated price instrument for a given consumer, in line with the steps shown in Figure 46. Panel A shows the simulated price instrument for a scenario in which the marginal price schedule does not change between  $t - 12$  and  $t$ . For this scenario, all simulated price instrument are 0, by construction and there is no variation in price left to estimate regression (87). Panel B shows the simulated price instrument for a scenario in which the marginal price schedule does change between  $t - 12$  and  $t$ . In this case, all simulated price instruments are positive, though only the simulated price instruments for the pairs  $\{A_1, A_2\}$  and  $\{B_1, B_2\}$  provide us with price variation that can be used to estimate regression (87).

This construction of the simulated instrument allows me to deal with the correlation between past and present price levels and unobserved demand shocks  $\epsilon_{ilt}$  and with mean

reversion.<sup>45</sup>

I run the following two-stage least square regressions:

$$\text{First stage: } \Delta \ln(P_{ilt}) = \theta_{lt} + \gamma \Delta \ln(\widehat{P_{il,t}}) + h(Q_{il,t-6}) + \delta X_{ilt} + \nu_{ilt} \quad (86)$$

$$\text{Second stage: } \Delta \ln(x_{ilt}) = \mu_{lt} + \lambda \Delta \ln(P_{ilt}) + h(Q_{il,t-6}) + \Gamma X_{ilt} + \epsilon_{ilt} \quad (87)$$

The underlying assumption needed for  $\lambda$  to give us unbiased estimates of the price elasticity of electricity demand are:  $Cov(\Delta \ln(P_{ilt}), \Delta \ln(\widehat{P_{il,t}})) \neq 0$  and  $Cov(\Delta \ln(\widehat{P_{il,t}}), \epsilon_{ilt}) = 0$ . The exclusion restriction requires that the unobserved idiosyncratic demand shocks at time  $t - 6$  be uncorrelated with the unobserved idiosyncratic demand shocks  $\epsilon_{ilt}$ . In other words, I need the simulated price instrument  $\Delta \ln(\widehat{P_{il,t}})$  to be uncorrelated with unobservable demand shocks at time  $t$  and unobservable demand shocks at time  $t - 12$ :  $Cov(Q_{il,t-6}, \epsilon_{il,t-12}) = 0$ , and (ii)  $Cov(Q_{il,t}, \epsilon_{il,t-12}) = 0$ .

The use of the simulated instrument does not guarantee the absence of a correlation between the residual  $\epsilon_{ilt}$  and past consumption and between residual  $\epsilon_{ilt}$  and past prices. For example, the changes in consumption distribution over time could lead to different trends for customers at different levels of electricity use (e.g. high-usage customers experiencing different demand shocks or responding differently to demand shocks compared to low-usage customers). To address this issue, I flexibly introduce lagged consumption variables  $h(Q_{il,t-6})$  as controls for customers' relative position in the electricity use distribution and for the different structural shifts in electricity demand based on the position of the customer in the distribution. Following Ito (2014), I non-parametrically estimate this function with usage percentile fixed effects: for each month, I calculate the usage percentile of each customer's consumption six month prior and interact corresponding percentile dummies with month-by-year fixed effects. These month-by-year fixed effects and the percentiles dummies for past consumption control for most variation in price in the setting.

In addition to controlling for the position of clients in the consumption distribution, I also check for the correlation between residualized electricity use at time  $t - 6$  and residualized electricity use at time  $t$ . Table 19 shows this correlation.<sup>46</sup> The very low correlation between the two residualized values provide suggestive evidence that the assumption could be reasonable in the setting. Overall, since the variation that I rely on compare consumers with similar levels of electricity use at time  $t - 6$ , I rule out

<sup>45</sup>If a customer experiences a very large shock at time  $t - 12$ , the shock could shift the customer into a higher tier corresponding to a higher marginal price and a higher average price. Once the shock is relaxed and their electricity use returns to a normal level in subsequent periods, this reduction in consumption to mean consumption level could incorrectly be attributed to the increase in price at  $t - 12$ .

<sup>46</sup>Here, the residualized values are calculated as the difference between the actual and predicted electricity use, where the predicted electricity use is based on a regression of electricity use on district-year-month fixed effects interacted with consumption deciles fixed effects plus consumer fixed effects.

consumers that experience large shocks that create a large gap between their electricity use and that of consumers with similar past consumption behavior. In addition, all unobserved price shocks and all unobserved aggregate demand shocks are absorbed by the fixed effects in the regression.

I run the regression specified in equation (87) separately for the following 6 groups of customers based on electricity use and usage type (small residential, large residential, small firms, and large firms), both before and after the reliability projects.

Table 19 reports the results of regressions of residualized electricity at time  $t$  and simulated price instrument.

Table 19: Correlation between Residualized Electricity Use at time  $t$  and Simulated Price Instrument

|                                     | Residualized consumption at $t$ |                    |                    |
|-------------------------------------|---------------------------------|--------------------|--------------------|
|                                     | (1)                             | (2)                | (3)                |
| Residualized consumption at $t - 6$ | 0.0014***<br>(0.0005)           |                    |                    |
| $\Delta \log(\widehat{MP})$         |                                 | 0.0029<br>(0.0039) |                    |
| $\Delta \log(\widehat{AP})$         |                                 |                    | 0.0009<br>(0.0049) |
| District-Year-Month FEs             | Yes                             | Yes                | Yes                |
| Meter Type FEs                      | Yes                             | Yes                | Yes                |
| Past Cons. Controls                 | Yes                             | Yes                | Yes                |
| Consumer FEs                        | Yes                             | Yes                | Yes                |
| Observations                        | 40,161,594                      | 34,291,908         | 34,291,908         |
| R <sup>2</sup>                      | 0.04302                         | 0.04999            | 0.04999            |

Notes: This table reports the results of regressions of past electricity use on current electricity use. Column (1) regresses electricity use at time  $t$  on residualized electricity use at time  $t - 6$ . Columns (2) regresses residualized electricity reuse at time  $t$  on  $\Delta \log(\widehat{MP})$ , the simulated marginal price instrument constructed using consumption levels from six months earlier. Columns (3) regresses residualized electricity reuse at time  $t$  on  $\Delta \log(\widehat{AP})$ , the simulated average price instrument constructed using consumption levels from six months earlier. These price instruments are described in equation (86) and visualized in Figures 46 and 47. All regressions include district by month by year fixed effects interacted with dummies for consumption deciles at  $t - 6$ , along with consumer fixed effects. The regressions are run using 94,475,034 bills between 2012 and 2024. Standard errors are clustered at the consumer level.

\* :  $p < 0.1$ , \*\* :  $p < 0.05$  and \*\*\* :  $p < 0.01$ .

Table 20 reports the first-stage results price elasticity of electricity use for residential and commercial clients between 2012 and 2023.

Table 20: First-Stage Results from the Regression of Price on Simulated Price Instrument

|                            | $\Delta \log$ marginal electricity price / kWh |                       |
|----------------------------|------------------------------------------------|-----------------------|
|                            | Before<br>(1)                                  | After<br>(2)          |
| <i>Panel A: Households</i> |                                                |                       |
| All                        | 0.7480***<br>(0.0013)                          | 0.5322***<br>(0.0017) |
| Low Consumption            | 0.8411***<br>(0.0013)                          | 0.6187***<br>(0.0019) |
| High Consumption           | 0.5372***<br>(0.0032)                          | 0.3674***<br>(0.0031) |
| <i>Panel B: Firms</i>      |                                                |                       |
| All                        | 0.9339***<br>(0.0016)                          | 0.9266***<br>(0.0017) |
| Low Consumption            | 0.9467***<br>(0.0022)                          | 0.9326***<br>(0.0018) |
| High Consumption           | 0.9260***<br>(0.0023)                          | 0.9163***<br>(0.0031) |
| District-Year-Month FEs    | Yes                                            | Yes                   |
| Meter Type FEs             | Yes                                            | Yes                   |
| Past Cons. Controls        | Yes                                            | Yes                   |
| Consumer FEs               | Yes                                            | Yes                   |

Notes: This table reports the first-stage results price elasticity of electricity use for residential and commercial clients between 2012 and 2023. Each coefficient and its associated standard error are for the regression of  $\Delta \log(MP)$  on  $\Delta \log(\hat{MP})$  on the sample described in the first column of the table. In each regression, the dependent variable is  $\Delta \log(MP)$  and the independent variable is  $\Delta \log(\hat{MP})$ , the simulated marginal price instrument constructed using consumption levels from six months earlier, as described in equation (86) and visualized in Figures 46 and 47. Low consumption households are households with mean monthly consumption below 170 kWh and high consumption households are households with mean monthly consumption above 170 kWh. Low consumption firms are households with mean monthly consumption below 60 kWh and high consumption firms are households with mean monthly consumption above 60 kWh. Columns (1) reports the first-stage coefficient before the reliability project, while column (2) report the first-stage coefficient after the reliability project. All regressions include district by month by year fixed effects interacted with dummies for consumption deciles at  $t - 6$ , along with consumer fixed effects. The regressions are run using 94,475,034 bills between 2012 and 2024. Standard errors are clustered at the consumer level. \* :  $p < 0.1$ , \*\* :  $p < 0.05$  and \*\*\* :  $p < 0.01$ .

Table 21 reports the short-run price elasticity of electricity demand for residential clients served by SENELEC between 2012 and 2023.

| Table 21: Short-Run Price Elasticity of Demand (Households) |                                             |                        |                        |                      |                        |                        |                        |
|-------------------------------------------------------------|---------------------------------------------|------------------------|------------------------|----------------------|------------------------|------------------------|------------------------|
|                                                             | $\Delta \log$ monthly electricity use (kWh) |                        |                        |                      |                        |                        |                        |
|                                                             | (1)                                         | (2)                    | (3)                    | (4)                  | (5)                    | (6)                    | (7)                    |
| $\Delta \log(AP)$                                           | -0.3514***<br>(0.0010)                      |                        |                        |                      |                        |                        |                        |
| $\Delta \log(\widehat{AP})$                                 |                                             | -0.3196***<br>(0.0059) | -0.5195***<br>(0.0121) |                      |                        |                        | -0.5140***<br>(0.0160) |
| $\Delta \log(MP)$                                           |                                             |                        |                        | 4.787***<br>(0.0021) |                        |                        |                        |
| $\Delta \log(\widehat{MP})$                                 |                                             |                        |                        |                      | -0.1751***<br>(0.0034) | -0.3327***<br>(0.0080) | -0.0058<br>(0.0099)    |
| District-Year-Month                                         | Yes                                         | Yes                    | Yes                    | Yes                  | Yes                    | Yes                    | Yes                    |
| Meter type FEs                                              | Yes                                         | Yes                    | Yes                    | Yes                  | Yes                    | Yes                    | Yes                    |
| Past Cons. Controls                                         | No                                          | No                     | Yes                    | No                   | No                     | Yes                    | Yes                    |
| Observations                                                | 54,178,589                                  | 53,576,564             | 53,576,564             | 54,210,222           | 53,576,564             | 53,576,564             | 53,576,564             |
| R <sup>2</sup>                                              | 0.05428                                     | 0.04467                | 0.06015                | 0.24536              | 0.04465                | 0.06013                | 0.06015                |
| # Fixed Effects                                             | 2,699                                       | 2,652                  | 24,890                 | 2,700                | 2,652                  | 24,890                 | 24,890                 |

Notes: This table reports the short-run price elasticity of electricity demand for all residential clients served by SENELEC between 2012 and 2023. Columns (1)-(3) report the average price elasticity while columns (4)-(6) report the marginal price elasticity. The dependent variable is the change in the log of monthly electricity use (in kWh) by a customer between the current billing cycle and the same period in the previous year.  $\Delta \log(AP)$  is the change in the log average price per kWh between the current billing cycle and the same period in the previous year.  $\Delta \log(\widehat{AP})$  is the simulated price instrument constructed using consumption levels from six months earlier. Panel B reports estimates for identical specifications that use the corresponding marginal price measures. Past Consumption controls specify the percentiles consumption six months prior interacted with the month-year fixed effects. Standard errors are clustered at the consumer level. \* :  $p < 0.1$ , \*\* :  $p < 0.05$  and \*\*\* :  $p < 0.01$ .

Table 22 reports the short-run price elasticity of electricity demand for commercial clients served by SENELEC between 2012 and 2023.

Table 22: Price Elasticity of Demand (Firms)

|                             | $\Delta \log$ monthly electricity use (kWh) |                        |                        |                      |                        |                        |                        |
|-----------------------------|---------------------------------------------|------------------------|------------------------|----------------------|------------------------|------------------------|------------------------|
|                             | (1)                                         | (2)                    | (3)                    | (4)                  | (5)                    | (6)                    | (7)                    |
| $\Delta \log(AP)$           | -1.329***<br>(0.0034)                       |                        |                        |                      |                        |                        |                        |
| $\Delta \log(\widehat{AP})$ |                                             | -0.1731***<br>(0.0242) | -0.5436***<br>(0.0351) |                      |                        |                        | -0.4849***<br>(0.0726) |
| $\Delta \log(MP)$           |                                             |                        |                        | 15.73***<br>(0.0577) |                        |                        |                        |
| $\Delta \log(\widehat{MP})$ |                                             |                        |                        |                      | -0.1415***<br>(0.0193) | -0.4974***<br>(0.0312) | -0.0594<br>(0.0611)    |
| District-Year-Month FEs     | Yes                                         | Yes                    | Yes                    | Yes                  | Yes                    | Yes                    | Yes                    |
| Meter type FEs              | Yes                                         | Yes                    | Yes                    | Yes                  | Yes                    | Yes                    | Yes                    |
| Past Cons. Controls         | No                                          | No                     | Yes                    | No                   | No                     | Yes                    | Yes                    |
| Observations                | 12,230,424                                  | 12,028,399             | 12,028,399             | 12,241,492           | 12,028,399             | 12,028,399             | 12,028,399             |
| R <sup>2</sup>              | 0.13224                                     | 0.07297                | 0.09317                | 0.28166              | 0.07297                | 0.09317                | 0.09317                |
| # Fixed Effects             | 2,699                                       | 2,652                  | 24,890                 | 2,700                | 2,652                  | 24,890                 | 24,890                 |

Notes: This table reports the short-run price elasticity of electricity demand for all commercial clients served by SENELEC between 2012 and 2023. Columns (1)-(3) report the average price elasticity while columns (4)-(6) report the marginal price elasticity. The dependent variable is the change in the log of monthly electricity use (in kWh) by a customer between the current billing cycle and the same period in the previous year.  $\Delta \log(AP)$  is the change in the log average price per kWh between the current billing cycle and the same period in the previous year.  $\Delta \log(\widehat{AP})$  is the simulated price instrument constructed using consumption levels from six months earlier. Panel B reports estimates for identical specifications that use the corresponding marginal price measures. Past Consumption controls specify the percentiles consumption six months prior interacted with the month-year fixed effects. Standard errors are clustered at the consumer level. \* :  $p < 0.1$ , \*\* :  $p < 0.05$  and \*\*\* :  $p < 0.01$ .

Table 23 report the price elasticity of electricity demand for all clients between 2012 and 2023, with either the average price and the marginal price, and with different fixed effects. In the preferred and more robust specifications in columns (3) and (6) of these tables, I estimate a price average price elasticity and marginal price elasticity of -0.52 and -0.33 for households, and a price average price elasticity and marginal price elasticity of -0.54 and -0.50 for firms. In line with the latest estimates in the literature, the encompassing regressions in column (7) of both tables show that consumers are more sensitive to average price than to marginal price.

Table 21 and 22 respectively report the price elasticity of electricity demand for all residential and commercial clients between 2012 and 2023, with either the average price and the marginal price, and with different fixed effects.

Table 23: Price Elasticity of Demand (All)

|                             | $\Delta \log$ monthly electricity use (kWh) |                        |                        |                      |                        |                        |                        |
|-----------------------------|---------------------------------------------|------------------------|------------------------|----------------------|------------------------|------------------------|------------------------|
|                             | (1)                                         | (2)                    | (3)                    | (4)                  | (5)                    | (6)                    | (7)                    |
| $\Delta \log(AP)$           | -0.3930***<br>(0.0008)                      |                        |                        |                      |                        |                        |                        |
| $\Delta \log(\widehat{AP})$ |                                             | -0.3298***<br>(0.0064) | -0.5120***<br>(0.0093) |                      |                        |                        | -0.2632***<br>(0.0142) |
| $\Delta \log(MP)$           |                                             |                        |                        | 5.002***<br>(0.0031) |                        |                        |                        |
| $\Delta \log(\widehat{MP})$ |                                             |                        |                        |                      | -0.2335***<br>(0.0038) | -0.4640***<br>(0.0071) | -0.2552***<br>(0.0094) |
| District-Year-Month         | Yes                                         | Yes                    | Yes                    | Yes                  | Yes                    | Yes                    | Yes                    |
| Meter type FEs              | Yes                                         | Yes                    | Yes                    | Yes                  | Yes                    | Yes                    | Yes                    |
| Past Cons. Controls         | No                                          | No                     | Yes                    | No                   | No                     | Yes                    | Yes                    |
| Observations                | 66,409,013                                  | 65,604,963             | 65,604,963             | 66,451,714           | 65,604,963             | 65,604,963             | 65,604,963             |
| R <sup>2</sup>              | 0.05854                                     | 0.04475                | 0.06093                | 0.22145              | 0.04473                | 0.06094                | 0.06094                |

Note: This table reports the price elasticity of electricity demand for all clients served by SENELEC between 2012 and 2023. Columns (1)-(3) report the average price elasticity while columns (4)-(6) report the marginal price elasticity. The dependent variable is the change in the log of monthly electricity use (in kWh) by a customer between the current billing cycle and the same period in the previous year.  $\Delta \log(AP)$  is the change in the log average price per kWh between the current billing cycle and the same period in the previous year.  $\Delta \log(\widehat{AP})$  is the simulated price instrument constructed using consumption levels from six months earlier. Panel B reports estimates for identical specifications that use the corresponding marginal price measures. Past Consumption controls specify the percentiles consumption six months prior interacted with the month-year fixed effects. Standard errors are clustered at the consumer level. \* :  $p < 0.1$ , \*\* :  $p < 0.05$  and \*\*\* :  $p < 0.01$ .

## 13 Standard Logit Model

In this section, I show the results of a standard logit by allowing the coefficients in equation (23) to be common across all consumers:

$$U_{iljt} = \beta_1 h_{1,ljt} + \beta_2 h_{2,ljt} + \beta_3 e_{jt} + \beta_4 p_{iljt} + \xi_j + \xi_{lt} + \epsilon_{iljt} \quad (88)$$

The standard logit model allows me to estimate the MWTP directly using the variation induced by reliability projects.

### 13.1 Descriptive Statistics

Table 24 reports descriptive statistics on the share of households using different energy sources, along with different compositions of appliances over time. Between 2018 and 2022, the share of households with no connection decreased by 8.4 percentage points, the share of households with grid connection increased by 7.5 percentage points, and

the share of households with solar panels only increased by a 0.9 percentage points.<sup>47</sup> The share of households with no electric appliances is lower than the share of households unconnected to the grid because many households have mobile phones but are unconnected to the grid.

**Table 24: Energy Source and Appliance Composition over Time**

|                                           | Year |      |      |      |      |      |      |      |
|-------------------------------------------|------|------|------|------|------|------|------|------|
|                                           | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 |
| <i>Panel A: Energy Source</i>             |      |      |      |      |      |      |      |      |
| No connection                             | –    | –    | –    | 29.9 | –    | –    | –    | 21.5 |
| Grid connection                           | –    | –    | –    | 59.0 | –    | –    | –    | 66.5 |
| Solar panels                              | –    | –    | –    | 11.1 | –    | –    | –    | 12.0 |
| <i>Panel B: Appliance Composition</i>     |      |      |      |      |      |      |      |      |
| No electrical appliances                  | 22.3 | 9.0  | 6.2  | 4.8  | 3.0  | 2.5  | 1.8  | 1.3  |
| Mobile phones only                        | 44.5 | 48.8 | 44.4 | 40.5 | 38.0 | 35.2 | 33.8 | 33.6 |
| Low-intensity appliances only             | 8.3  | 4.3  | 3.7  | 3.1  | 2.0  | 1.8  | 1.5  | 0.9  |
| Low-intensity appliances + mobile phones  | 20.3 | 31.5 | 37.7 | 42.0 | 45.4 | 46.6 | 48.0 | 48.7 |
| Low-intensity + high-intensity appliances | 4.6  | 6.4  | 8.1  | 9.6  | 11.6 | 13.8 | 14.9 | 15.4 |

Notes: This table reports the energy source and the appliance compositions for the sample of respondents in the EHCVM dataset described in Section 3 between 2015 and 2023. Each value reports the share of households that uses the energy source stated in Panel A or owns the type of appliances listed in Panel B. Low-intensity appliances include TVs, fans, refrigerators, laptops, and tablets. High-intensity appliances includes air conditioners, freezers, irons, generators, microwaves, mixers, printers, vacuums, and washing machines. The time series of ownership of a given appliance X is constructed based on two survey questions in the EHCVM: (1) do you own appliance X? If so, since when have you owned appliance X? For appliances with more than one unit of ownership, only the last year of purchase is reported. Thus, the values reported in the table represent the share of households that owns a given appliance with no replacement or upgrade of that asset for the stated year. Only households included in both rounds of the EHCVM surveys (2018/2019 and 2021/2022) are used to compute the statistics (N = 5,864 households).

Table 25 reports the mean values of variables related to household characteristics, electricity use and expenditures, income, consumption expenditures, and employment status separately for different groups of clients. As expected, unconnected households are overall more likely to have a low income and to live in rural areas, while connected households with sophisticated appliances are more likely to be more urban, to live in larger establishments, and to have a high income. The last column of the table shows

<sup>47</sup>The share of households with grid connection includes households connected to the grid only as well as those with additional defensive investments such as generators. Generator ownership among households is relatively low, as only 3–4% of all households have one. Only very a small share of households uses both grid electricity and solar panels.

households using solar minigrids. These households tend to be rural households that own their establishments and are better off than the typical rural household.

**Table 25: Characteristics of Households by Connection and Appliance Type (in 2018)**

| Connection type:                                | None   | Grid       | Grid             | Solar            |
|-------------------------------------------------|--------|------------|------------------|------------------|
| Appliance type:                                 | None   | Basic Only | Basic + High-End | Basic + High-End |
|                                                 | (1)    | (2)        | (3)              | (4)              |
| <b>Household Characteristics</b>                |        |            |                  |                  |
| Household size                                  | 9.56   | 9.52       | 9.74             | 11.38            |
| Owns the house                                  | 0.88   | 0.73       | 0.75             | 0.91             |
| Number of rooms in HH                           | 4.11   | 4.57       | 5.54             | 5.22             |
| Lives in urban area                             | 0.22   | 0.73       | 0.83             | 0.16             |
| <b>Electricity</b>                              |        |            |                  |                  |
| Electricity expenditures (000s FCFA)            | –      | 8.77       | 17.14            | –                |
| Number of phones                                | 2.71   | 3.54       | 3.99             | 3.06             |
| Number of basic appliances                      | –      | 6.87       | 9.40             | 5.61             |
| Number of high-end appliances                   | –      | 0.00       | 1.99             | 0.08             |
| Number of outages in a week                     | –      | 5.24       | 5.54             | –                |
| Duration of outages in a week (in minutes)      | –      | 44.94      | 41.12            | –                |
| Subscription costs for the grid (000s FCFA)     | –      | 43.69      | 43.34            | –                |
| Connection costs for the grid (000s FCFA)       | –      | 22.18      | 22.55            | –                |
| <b>Expenditures</b>                             |        |            |                  |                  |
| Household income (000s FCFA)                    | 248.50 | 851.29     | 2217.30          | 392.90           |
| Household revenue (000s FCFA)                   | 20.65  | 110.97     | 301.52           | 66.83            |
| Expenditures on perishable goods (000s FCFA)    | 358.21 | 623.54     | 1070.30          | 462.66           |
| Expenditures on nonperishable goods (000s FCFA) | 924.26 | 1099.30    | 1385.90          | 1163.20          |
| <b>Employment</b>                               |        |            |                  |                  |
| Is self-employed                                | 0.11   | 0.18       | 0.15             | 0.13             |
| Is employed by others                           | 0.06   | 0.13       | 0.18             | 0.06             |
| Number of observations                          | 1753   | 2596       | 862              | 653              |

Notes: This table reports the mean values of variables related to household characteristics, electricity use and expenditures, income, consumption expenditures, and employment status. Column (1) is for households not connected to the grid. Column (2) is for households connected to the grid and using only low-intensity appliances. Column (3) is for households connected to the grid and using both low-intensity and high-intensity appliances. Column (4) is for households not connected to the grid and owning solar panels. Low-intensity appliances include TVs, fans, refrigerators, laptops, and tablets. High-intensity appliances include air conditioners, freezers, irons, generators, microwaves, mixers, printers, vacuums, and washing machines. All amounts are reported in thousand CFA franc. 1 US dollar is approximately 550 CFA franc.

## 13.2 Estimation

The main challenge in my estimation of the discrete choice model is that two variables in the demand estimation are likely to be endogenous in the EHCVM. First, electricity reliability depends on observed and unobserved economic conditions and can there-

fore be correlated with omitted variables. To address this problem, we leverage the reliability improvement projects discussed in Section 2, which give us plausibly exogenous variation in location served by different feeder lines with different exposures to treatment. Second, the price of electricity and that of appliances are also potentially endogenous. Since we use data from many markets in Senegal, our product and market fixed effects absorb unobserved demand factors that are product specific or market specific. Thus, the endogeneity that remains is that driven by market/product-specific omitted factors that correlate with market/product-specific price variation. Importantly, the main product-specific driver of price that is often omitted in these models is product quality, which I observe in my setting and model for. My goal in this subsection is to obtain a revealed-preference estimate of the WTP for reliable electricity while accounting for these endogeneity issues.

For electricity availability, I instrument electricity reliability with reliability improvement projects. In particular,  $T_l$  is a variable that takes the value of 1 if market  $l$  received reliability projects before 2018 and 0 otherwise, and  $P_{lt}$  is a dummy variable that takes the value of 1 if the year is post-treatment and 0 otherwise.  $h_{lt}$  is the total number of hours for which grid electricity is available during time  $t$ .<sup>48</sup> I estimate the first-stage effect of reliability projects on electricity availability using the following equation:

$$h_{lt} = \phi_1 T_l + \phi_2 P_{lt} + \phi_3 T_l P_{lt} + \nu_l + \nu_t + \eta_{lt} \quad (89)$$

For the appliance price, I also consider a distance-based instrument for the weighted-average appliance price described above. Specifically, I consider  $D_{lt}$  the distance between the port of entry and the market, with the implicit assumption that all appliances are imported.  $D_{lt}$  should thus reflect the transportation costs between the point of entry and end consumers.

I then construct the market shares for option  $j$  directly from the EHCVM data since these data are a nationally representative sample of households. I define the market share for the outside option (i.e., no connection) as  $s_{0lt} = 1 - \sum_{k=1}^J s_{klt}$ . Considering the difference between  $\ln s_{jlt}$ , the log market share for option  $j$ , and  $\ln s_{0lt}$ , the log market share for the outside option, and the fact that  $\ln s_{0lt}$  is absorbed by location-time fixed effects gives the following:

$$\ln s_{jlt} = \beta_1 h_{1,jt} + \beta_2 h_{2,jt} + \beta_3 e_{jt} + \beta_4 p_{jlt} + \xi_j + \xi_{lt} + \epsilon_{iljt} \quad (90)$$

Following Ito and Zhang (2020), I consider less stringent fixed effects and use the shares from equation (90) in a two-stage least squares estimation as follows:

---

<sup>48</sup> $h_{lt}$  could be for either the window of peak hours or the window of off-peak hours; alternatively, it could include all hours.

$$\text{First-stage: } h_{lt} = \phi_1 T_l + \phi_2 P_{lt} + \phi_3 T_l P_{lt} + \nu_l + \nu_t + \eta_{lt} \quad (91)$$

$$\text{Second-stage: } \ln s_{jlt} = \lambda_1 h_{lt} E_j + \lambda_2 e_{jt} + \lambda_3 p_{jlt} + \xi_l + \xi_j + \xi_t + \epsilon_{jlt} \quad (92)$$

$$\text{Reduced Form: } \ln s_{jlt} = \kappa_1 T_l E_j + \kappa_2 P_{lt} T_l E_j + \kappa_3 e_{jt} + \kappa_4 p_{jlt} + \xi_l + \xi_j + \xi_t + \epsilon_{jlt} \quad (93)$$

where  $E_j$  is a dummy variable that takes the value of 1 if option  $j$  relies on grid electricity and 0 otherwise. Therefore,  $E_j = 1$  for  $j = 1$  and  $j = 2$  and 0 for other values of  $j$ .  $\xi_l$ ,  $\xi_j$ , and  $\xi_t$  are location, product, and time fixed effects respectively.

### 13.3 Results

Table 26 presents the first-stage estimation results for electricity reliability and appliance prices. Panel A examines the impact of the treatment on electricity availability, distinguishing between peak, off-peak, and all hours. The treatment variable “Treated x Post” represents markets that received the reliability intervention in years following the treatment. The coefficient on “Treated x Post” is positive and statistically significant across all three columns, indicating that treated markets experienced an increase in electricity availability. Specifically, electricity availability increased by 0.72 hours during peak hours, 15.47 hours during off-peak hours, and 16.19 hours overall in treated markets relative to untreated ones, with significance levels of 1% for peak hours and 5% for off-peak and total hours.

Panel B of Table 26 reports the first-stage results for appliance prices. The key explanatory variable here is the distance from the port of entry to each market, which serves as an instrument under the assumption that appliance prices reflect import and transport costs. While the coefficients for the distance variable are negative across all specifications, the coefficient is generally insignificant, suggesting that distance does not play a key role in determining the appliance price. In addition, the coefficients associated with most socioeconomic controls are insignificant, suggesting that within markets, these socioeconomic characteristics do not affect the appliance price.

**Table 26: First-Stage Estimation for Electricity Reliability and Appliance Price**

|                                          | (1)                                          | (2)                | (3)                 |
|------------------------------------------|----------------------------------------------|--------------------|---------------------|
|                                          | Dependent variable: electricity availability |                    |                     |
|                                          | Peak                                         | Off-peak           | All                 |
| <b>Panel A. Electricity availability</b> |                                              |                    |                     |
| Treated x Post                           | 0.7224***<br>(0.1434)                        | 15.47**<br>(4.644) | 16.19**<br>(4.697)  |
| Observations                             | 40,608                                       | 40,608             | 40,608              |
| R <sup>2</sup>                           | 0.12                                         | 0.37               | 0.38                |
| Year FEs                                 | Yes                                          | Yes                | Yes                 |
| Month FEs                                | Yes                                          | Yes                | Yes                 |
| Market FEs                               | Yes                                          | Yes                | Yes                 |
|                                          | Dependent variable: appliance price          |                    |                     |
|                                          | Price                                        | Price              | Price               |
| <b>Panel B. Appliance price</b>          |                                              |                    |                     |
| Distance between port and market         | -4.950<br>(2.007)                            | -2.073<br>(0.3630) | -1.915*<br>(0.2446) |
| Observations                             | 11,720                                       | 11,720             | 11,307              |
| R <sup>2</sup>                           | 0.12                                         | 0.25               | 0.26                |
| Year FEs                                 | Yes                                          | Yes                | Yes                 |
| Product FEs                              |                                              | Yes                | Yes                 |
| Socioeconomic controls                   |                                              |                    | Yes                 |

Notes: This table presents the results of equation (91). The dependent variables in Panel A are the total number of hours for which electricity is available during peak hours, off-peak hours, and across all hours. “Treated” is a dummy variable that takes the value of 1 if the feeder was treated before or during 2018 and 0 otherwise. “Post” is a dummy variable that takes the value of 1 for years post-treatment. Each observation in Panel A is a market-month pair. The dependent variable in Panel B is the weighted-average appliance price. The distance variable reflects the distance between the point of entry of the appliance and the market (with the assumption that all appliances are imported from abroad). Socioeconomic controls include consumption expenditure, household size, the number of low-intensity appliances, and the number of high-intensity appliances. Each observation in Panel B is a client in the EHCVM dataset. Both panels use the EHCVM and outage data described in Section 3. \*: significance at the 10% level, \*\*: significance at the 5% level, and \*\*\*: significance at the 1% level.

Table 27 reports the reduced-form and second-stage estimation results for the standard logit model, which analyzes how electricity availability and appliance prices impact market shares. In Panel A, the reduced-form results indicate that grid connection has a significant impact on market shares, with the interaction term “Treated x Grid connection” showing positive and significant coefficients in both columns. Furthermore, “Treated x Post x Grid connection” positively affects market shares, suggesting that post-treatment periods in treated markets see higher rates of grid electricity adop-

tion. The coefficients for “Weighted-average electricity price” and “Weighted-average appliance price” are negative and statistically significant, indicating that higher prices reduce market share.

Table 27: Standard Logit: Reduced-Form Results

|                                                   | Dependent variable: $\ln(\text{Share})$ |                                          |
|---------------------------------------------------|-----------------------------------------|------------------------------------------|
|                                                   | (1)                                     | (2)                                      |
| Treated x Grid connection ( $\kappa_1$ )          | 0.0505*<br>(0.0067)                     | 0.1072**<br>(0.0082)                     |
| Treated x Post x Grid connection ( $\kappa_2$ )   | 0.0063***<br>( $8.34 \times 10^{-6}$ )  | 0.0064***<br>( $1.02 \times 10^{-14}$ )  |
| Weighted-average electricity price ( $\kappa_3$ ) | -0.0479**<br>(0.0026)                   | -0.0191***<br>( $3.3 \times 10^{-15}$ )  |
| Weighted-average appliance price ( $\kappa_4$ )   | -0.0029***<br>( $4.94 \times 10^{-8}$ ) | -0.0029***<br>( $1.37 \times 10^{-16}$ ) |
| Observations                                      | 4,768                                   | 4,768                                    |
| R <sup>2</sup>                                    | 0.29                                    | 0.29                                     |
| Year FEs                                          | Yes                                     | Yes                                      |
| Product FEs                                       | Yes                                     | Yes                                      |
| Market FEs                                        |                                         | Yes                                      |

Notes: This table presents the results of the reduced-form and the second-stage results shown in equations (92) and (93). The dependent variables in all regressions is the log of market shares, as defined in equations (90). “Treated” is a dummy variable that takes the value of 1 if the feeder line serving the market was treated before or during 2018 and 0 otherwise. “Post” is a dummy variable that takes the value of 1 for years post-treatment. “Grid electricity” is a dummy variable that takes the value of 1 if the option involves connecting to the grid. The regressions use the EHCVM and outage data described in Section 3. \*: significance at the 10% level, \*\*: significance at the 5% level, and \*\*\*: significance at the 1% level.

Panel A of Table 28 presents the second-stage results, estimating the effect of total electricity availability and appliance prices on market shares after accounting for potential endogeneity via instrumental variables. The coefficient for “Total electricity available x Grid connection” is positive and statistically significant, implying that increased electricity availability enhances the grid connection market share. Both “Weighted-average electricity price” and “Weighted-average appliance price” have negative and statistically significant coefficients, suggesting that higher prices reduce market shares. The F statistics from the first-stage regressions are robustly high, demonstrating the strength of the instruments, while the Wu–Hausman and Sargan tests indicate that the instruments are valid and that the endogenous regressor is correctly identified.

Panel B of Table 28 calculates the MWTP for grid reliability. The baseline MWTP is calculated as the ratio of the coefficient on “Treated x Grid connection” to the negative sum of the coefficients on electricity and appliance prices, yielding values of \$4.53 and \$22.27 for the two specifications, respectively; the change in MWTP resulting from reliability improvements is likewise calculated and presented, with values of \$0.57 and \$1.34, respectively. These changes in MWTP values confirm our reduced-form results

on the positive change in the economic value that consumers place on reliability following improvements in grid reliability.

Table 28: Standard Logit: Second-Stage Estimation Results

|                                                               | Dependent variable: $\ln(\text{Share})$ |                                       |
|---------------------------------------------------------------|-----------------------------------------|---------------------------------------|
|                                                               | (1)                                     | (2)                                   |
| <b>Panel A. Second-stage results:</b>                         |                                         |                                       |
| Total electricity available x Grid connection ( $\lambda_1$ ) | 0.0004**<br>( $2.33 \times 10^{-5}$ )   | 0.0004**<br>( $2.19 \times 10^{-5}$ ) |
| Weighted-average electricity price ( $\lambda_2$ )            | -0.6881*<br>(0.0806)                    | -0.1342**<br>(0.0031)                 |
| Weighted-average appliance price ( $\lambda_3$ )              | -0.0103*<br>(0.0013)                    | -0.0103*<br>(0.0014)                  |
| Observations                                                  | 4,072                                   | 4,072                                 |
| $R^2$                                                         | 0.29                                    | 0.38                                  |
| Year FEs                                                      | Yes                                     | Yes                                   |
| Market FEs                                                    |                                         | Yes                                   |
| First-stage F statistic                                       | 515,972.3                               | 642,870.9                             |
| Wu–Hausman statistic                                          | 54.8                                    | 70.4                                  |
| Sargan statistic (p-value)                                    | 0.044                                   | 0.40                                  |
| <b>Panel B. MWTP:</b>                                         |                                         |                                       |
| Baseline MWTP                                                 | 4.53065<br>(0.3719)                     | 22.26947<br>(1.6941)                  |
| Change in MWTP from reliability projects                      | 0.5701<br>(0.0300)                      | 1.3373<br>( $2.34 \times 10^{-12}$ )  |

Notes: This table presents the results of the reduced-form and the second-stage results shown in equations (92) and (93). The dependent variables in all regressions is the log of market shares, as defined in equations (90). “Treated” is a dummy variable that takes the value of 1 if the feeder line serving the market was treated before or during 2018 and 0 otherwise. “Post” is a dummy variable that takes the value of 1 for years post-treatment. “Grid electricity” is a dummy variable that takes the value of 1 if the option involves connecting to the grid. In Panel A, I report the first-stage F test results for the endogenous variable (electricity availability) to check whether my instrument (reliability project) is strong for predicting electricity reliability in the first stage. I also report the Wu–Hausman test results to assess the endogeneity of my endogenous variable (electricity availability). I also report the Sargan test results to check for the correlation between instrument validity and the error term. In Panel B, baseline MWTP is calculated as the ratio of  $\kappa_1$  and  $-(\kappa_3 + \kappa_4)$ . Change in MWTP from reliability projects is calculated as the ratio of  $\kappa_2$  and  $-(\kappa_3 + \kappa_4)$ . The standard errors for the MWTP estimates are calculated based on the delta method. The regressions use the EHCVM and outage data described in Section 3. \*: significance at the 10% level, \*\*: significance at the 5% level, and \*\*\*: significance at the 1% level.