Universidad de Navarra

Facultad de Ciencias Económicas y Empresariales

Working Paper nº 01/20

The Effect of Blackouts on Households' Electrification Status: evidence from Kenya

Raul Bajo-Buenestado
University of Navarra

The Effect of Blackouts on Households' Electrification Status: evidence from Kenya

Raul Bajo-Buenestado
April 2020

ABSTRACT

A number of countries in Sub-Saharan Africa have recently deployed billions of dollars to improve their electricity infrastructure. However, aggregate data shows that the relative number of households with an electricity connection at home has barely increased. In this paper we study the role of blackouts to partially explain why there have been relatively few additional households with electricity access despite the increase in electrification expenditure. Using geo-localized survey data from Kenya, we find that households that live in neighborhoods in which power outages are relatively more frequent are (at least) about 6%-9% less likely to have electricity at home. We also find that households that have electricity access but which experience frequent power outages are also less likely to purchase electrical appliances.

Raul Bajo-Buenestado
University of Navarra

The Effect of Blackouts on Households' Electrification Status: evidence from Kenya*

Raúl Bajo-Buenestado[†1, 2, 3]

[1]Department of Economics, University of Navarra

[2]Baker Institute for Public Policy, Rice University

[3]Navarra Center for International Development (NCID), University of Navarra

April 22, 2020

Abstract

A number of countries in Sub-Saharan Africa have recently deployed billions of dollars to improve their electricity infrastructure. However, aggregate data shows that the relative number of households with an electricity connection at home has barely increased. In this paper we study the role of blackouts to partially explain why there have been relatively few additional households with electricity access despite the increase in electrification expenditure. Using geo-localized survey data from Kenya, we find that households that live in neighborhoods in which power outages are relatively more frequent are (at least) about 6%-9% less likely to have electricity at home. We also find that households that have electricity access but which experience frequent power outages are also less likely to purchase electrical appliances.

JEL Classification Numbers: L94, O13, Q41, Q48.

Keywords: Energy poverty, Electricity access, Electrification rates, Sub-Saharan Africa.

---

*This paper was partially written while I was visiting the Strathmore University Business School in Nairobi (Kenya). I would like to thank them, and especially Robert Mudida, for their hospitality. I particularly benefited from conversations with Alex Armand, Miguel Angel Borrella-Mas, Luke Chicoine, Lucia Corno, Carlos Reyes, Marcello Sartarelli, Zaki Wahhaj, and with representatives from the Kenya Electricity Generating Company (KenGen), the Kenya Power and Lighting Company (KPLC), and the Energy Regulatory Commission of Kenya (ERC). Thanks also to participants at seminars at the Complutense University of Madrid and at an internal seminar at the University of Navarra. Special thanks to Wendy Schreiber Ayres, who helped me to access some information. I am also grateful to the Ramón Areces Foundation which has provided support for this research (grant number CISP16A4779). All views and any errors in this article are my own.

†Department of Economics, Edificio Amigos, University of Navarra 31009 Pamplona, Spain, Email: rbajo@unav.es

1 Introduction

Back in 2000, around 1.7 billion people lived without electricity; about one third of them were located in Sub-Saharan Africa (SSA) —IEA (2015) and IEA (2017). Spurred by the Millennium Development Goals, and more recently by the Sustainable Development Goals (SDG),[1] governments —typically in cooperation with third parties— deployed billions of dollars to improve infrastructure and expand grid coverage —Enerdata (2017). As a result, electricity access —defined as having an electric grid within reach— has risen (and is expected to increase) throughout the developing world. For the particular case of SSA, the electricity access rate increased from 51% to 65% between 2005 and 2015 in a group of 18 selected countries tracked by the Afrobarometer —Oyuke et al. (2016). In some of these, such as Nigeria and Kenya, the overall access rate by 2015 was above 80%, and close to universal access was achieved for households in urban settings.

These promising figures on electricity access, though, stand in stark contrast to the on-the-ground reality, which suggests that all these "grid expansion" efforts are not always fully translated into actual household benefits: as of 2018, around 602 million people in SSA (i.e., about 48% of the population) still lack an electricity connection in their homes, and more than 700 million people in SSA are projected to require it by 2040 —World Bank (2017) and IEA (2018).[2] In other words, even though a large proportion of African households are now "under grid" —i.e., close enough to connect to a low-voltage line at a relatively low cost (Lee et al., 2016)—, particularly in (and around) urban areas, many of them remain "unelectrified". Exploring why this happens is, therefore, crucial in order to identify cost-effective investment projects, and to alleviate the documented (costly) mismatch between the supply and demand of electricity infrastructure.

In the present paper we focus on the role of the lack of supply reliability as a factor that explains why many households remain unelectrified in developing countries.[3] As is widely documented, the current electricity infrastructure in SSA (as in many other developing regions) is highly unreliable. As a result, blackouts and brownouts are the norm across SSA —Jacome et al. (2019). If this is

[1] In particular, the SDG #7 aims at ensuring universal "access to affordable, reliable, [...] energy" by 2030.

[2] Similarly, Palit and Bandyopadhyay (2016) document that while India and Bangladesh have 97.4% and 62% of the villages covered through grid supply, rural household connection levels are around 74% and 48% respectively.

[3] Lee et al. (2019) study in their experiment whether outages affected the demand for grid connections. However, they recognize that "there is a need for research in several areas, including on the impacts of [...] reliability". Some other papers have focused on other causes that affect the demand for on-grid connections in developing countries —Blimpo and Cosgrove-Davies (2019). For instance, Blimpo et al. (2018) study the role of connection charges.

the case, we argue that households will perceive as lower the benefits of electrification, resulting in lower demand for electricity connections.

To address this policy-relevant empirical question, we use data from Kenya. This country is of particular interest as it presents two key ingredients that are often observed across most of the developing countries, namely, *(i)* a marked gap between the electricity access rate and the percentage of connected households, and *(ii)* a relatively low-reliable service.[4] We attempt to establish the causal link between the reliability of the supply and the demand for electricity connections using cross-sectional geo-localized data obtained from a survey of over 14,000-households conducted in Kenya in 2012 and 2013. Based on the average blackouts reported by households in a typical week, we build a variable that captures the frequency of outages (reliability) at the neighborhood level. This household survey data is combined with geographical data on electricity transmission lines, and on the location of power generators.

Using a linear probability model (LPM), we find a negative impact of the frequency of outages at the neighborhood-level on the probability with which households have an electricity connection, after controlling for a battery of household-level and neighborhood-level characteristics. In our preferred model specification, we find that moving from a neighborhood that experiences no outages to another one that experiences outages on a weekly basis decreases the probability of having a connection by 6%-9%. These estimates are robust to different specifications and alternative samples. Moreover, a falsification test suggests that they are unlikely to be driven by unobservable variables that may be correlated with the frequency of outages.

These results, however, should be interpreted as a lower bound of the true impact of the frequency of outages on households' electrification decisions. This is because local amenities and the quality of the infrastructure (such as the reliability of electricity) are relevant variables that affect the neighborhood choices of a household. Therefore, households that are less likely to have electricity may have a higher propensity to live in neighborhoods that frequently experience outages —creating, thus, a potential *selection bias* effect. Consistently, we document that the estimated effect is larger (more negative) when using propensity-score matching, which corrects the afore-

[4] As mentioned above, while the majority of Kenyans have an electric grid within reach, Oyuke et al. (2016) estimate that, in 2014-15, about 57% of households did not have a connection at home. Moreover, according to them, about 22% of the households that had electricity reported that their connections work either never, occasionally, or about half of the time. This figure suggests that, as in many other developing countries, the service is quite unreliable (additional figures are provided in Section 2).

mentioned problem.

Although the impact of the lack of a reliable supply on firms' electrification decisions —Steinbuks and Foster (2010) and Oseni and Pollitt (2015)—, and even on firms' productivity, revenue, and costs is well-documented —De Nooij et al. (2007), Fisher-Vanden et al. (2015), Allcott et al. (2016), and Cole et al. (2018)—, surprisingly few studies have focused on the effect of power outages on households' electrification decisions.[5] To the best of our knowledge, the only exception is by Millien (2017), who finds a negative impact of outages on the probability of connection for households in Kenya —controlling for households' poverty levels— using a probit model with cross-sectional data aggregated at the district level. However, we follow the arguments by Allcott and Taubinsky (2015) and Angrist and Pischke (2008) and estimate instead an LPM that includes a rich battery of household-level and neighborhood-level control variables that may explain household decisions concerning electrification, and as a consequence, could bias the estimates if not accounted for. Moreover, our matching estimation also accounts for the possibility of a household-location selection, an effect that Millien (2017) ignores.

It is worthwhile recalling that an electricity connection (*per se*) does not have an intrinsic value attached to it, but it is rather a necessary condition to provide a number of benefits to households (e.g. appliances access, health benefits). However, it may not be a sufficient one. In fact, recent reports by the World Bank —Bhatia and Angelou (2015)— and by the United Nations (2019) have criticized the frequently-used "binary metric" of whether people have/do not have a connection, as it can be misleading: the supply of electricity may not be reliable enough to power commonly-used electricity-utilizing assets. Bearing this criticism in mind, in this paper we go beyond the question of the impact of supply reliability on *"pure" electrification* decisions by exploring also whether, among those households that have a connection, frequent outages do also have an impact on those subsequent *post-electrification* decisions, such as on the purchase of domestic appliances.

In particular, we focus on the ownership of refrigerators and televisions, which —as documented by Lee et al. (2016)— are some of the most desired appliances among Kenyans. Moreover, these appliances are typically associated with a substantial increase in welfare in developing countries[6], as

[5] Some studies have explored the impact of the lack of supply reliability on households' willingness to pay for electricity. For instance, Dzansi et al. (2018) use data from Ghana to provide evidence of a negative impact of outages on subsequent electricity bill payments. Using survey data from 714 villages in rural India, Kennedy et al. (2019) document a positive relationship between quality of service and willingness to pay for an electricity connection.

[6] Using data from Honduras, the Inter-American Development Bank (2019) estimates a consumer surplus associated

they are directly linked to greater food safety and lower incidence of stomach diseases (for the case of refrigerators) —Heller et al. (2005)—, and to lower acceptability of domestic violence and greater women's autonomy (for the case of television) —Jensen and Oster (2009). In line with the previous results, we find evidence that frequent outages discourage the ownership of these appliances among households with a connection. The effect remains negative and significant when using instrumental variables, which correct for a potential *reverse causality* problem —as overload-related outages are more likely in neighborhoods where a substantial number of households own these appliances.[7]

Our empirical results deliver (at least) two key insights that inform energy-policy decisions. First, they suggest that the ratio of "additional population with connections" to "additional population with access" of grid expansion projects may be substantially below one if the reliability of the service is poor. This is important, as policymakers that ignore this gap may (systematically) overestimate the benefits of these projects. Second, they suggest that the gains from improving the quality of the existing infrastructure could be substantial if we consider that greater reliability increases households' likelihood of connection and appliance ownership. Therefore, these results refer directly to the well-acknowledged tradeoff between extending basic access to more people and enhancing the access of those already served —Inter-American Development Bank (2019).

The rest of the paper proceeds as follows. Section 2 provides some background on the persistence of outages in SSA. In Section 3 we study the impact of outage frequency on households' demand for electricity connections. In Section 4 we study the impact of outage frequency on households' demand for appliances. Finally, Section 5 discusses some policy implications and concludes.

2 Background: pervasive outages in SSA and consequences

Although power markets all over the world are exposed to unplanned outages, in developed countries these outages are, in general, *rara avis*, and are typically caused by extreme weather events —see, for instance, Table B.1 in EIA (2019). By contrast, blackouts and brownouts are the rule rather than the exception in developing countries. For the particular case of SSA, the International Energy Agency (IEA) estimates that power was unavailable about 540 hours per year in 2014 —IEA (2014). In the same vein, according to the most recent data provided by the World Bank's

with television ownership of around $17.1 per household per year, and of $23.3 associated with refrigerator ownership.

[7] Some additional robustness checks are also discussed in Section 4.

Enterprise Surveys, firms experience on average 9 outages per month. The case of Kenya is not an exception: according to Farquharson et al. (2018), power is unavailable in this country about 420 hours per year, and there are about 6.3 electrical outages in a typical month.

As explained by an independent evaluation group of the World Bank —IEG (2015)—, the main reason behind the persistence of outages in SSA is the lack of commercial viability of the electricity business. According to Huenteler et al. (2017), electric utilities in developing countries are usually well-connected to government authorities, who see electricity users not as customers but as voters. Hence, underpricing of electricity is popular for these authorities and, if implemented, is difficult (politically-costly) to remove. In fact, it is common to observe electricity tariffs set well below full recovery cost across SSA.[8] This under-recovery of costs is further exacerbated by certain other supply and demand-related issues such as droughts,[9] political and social conflict, lack of private (pro-business) investors, and billing and collection inefficiencies —Eberhard et al. (2008).

All these revenue-constraint issues experienced by electric utilities result in lack of investment in operation and maintenance (O&M) of the existing infrastructure (and also in additional facilities), leading thus to excessive transmission losses and recurrent power outages —IEG (2015). In line with this idea, Taneja (2017) documents that more than half of the outages experienced in the city of Nairobi (Kenya) in a typical year are caused by lack of O&M-related issues, such as equipment failures, loss of supply, problems in transformers, cables, poles, and the contact of objects.[10]

This lack of reliability of the electricity supply has notorious negative implications for firms in terms of profits and productivity in developing countries —Allcott et al. (2016) and Cole et al. (2018). Consequently, as Steinbuks and Foster (2010) explain, firms perceive the benefits of electrification as being lower. In this paper, we argue that an analogous argument applies to households: if power outages are frequent, the well-documented benefits of electrification —Bernard (2010) and Chakravorty et al. (2014)— will be perceived as being lower. This, in turn, results in *(i)* lower household electrification rates,[11] and *(ii)* lower demand for electric-powered appliances among the

[8] According to Eberhard et al. (2008) (at the time when they were writing) "nowhere in Sub-Saharan Africa do residential or commercial and industrial customers pay full cost-recovery prices". Similar issues are experienced in India —Burgess et al. (2020).

[9] Many countries in SSA are highly dependent on hydropower —WEC (2013).

[10] In addition, this author documents that about one third of the outages are due to local overloads on transformers. This information will be important in Section 4 —when we study the effect of outages on domestic appliances ownership— as local overloads are more likely to occur in areas with an abundant presence of high-power domestic appliances (such as refrigerators), leading thus to an aforementioned potential *reverse causality* problem.

[11] Some preliminary evidence is also provided in the appendix, where we plot an index of perceived quality of supply

households that have a connection. Throughout the rest of the paper we develop a thorough empirical strategy to analyze these two hypotheses using the aforementioned household-level survey data from Kenya.

3 Effect of Outages on Electricity Connections

3.1 Empirical framework

The first regression model that we estimate examines the effect of the frequency of outages in a neighborhood on the probability that households located in such a neighborhood are connected to electricity.[12] In this regression model, the unit of observation is a household i situated in neighborhood j —which is technically called Enumeration Area (EA)[13]—, in city c, and province p.[14] The effect of the frequency of outages in EA j on the probability that household i in EA j has an electricity connection at home is measured in the following cross-sectional regression:

$$y_{i,j,c,p} = \alpha_0 + \alpha_1 \theta_{j,c,p} + \alpha_2 \boldsymbol{X}_{i,j,c,p} + \alpha_3 \boldsymbol{Z}_{j,c,p} + \lambda_p + \lambda_c + \varepsilon_{i,j,c,p} \tag{1}$$

where $y_{i,j,c,p}$ is a dummy variable that is equal to 1 if household i is connected to electricity and 0 otherwise; $\theta_{j,c,p}$ captures the frequency of outages in EA j; $\boldsymbol{X}_{i,j,c,p}$ is a vector of household characteristics; $\boldsymbol{Z}_{j,c,p}$ is a vector of EA characteristics; λ_p and λ_c denote province and city fixed effects respectively; and $\varepsilon_{i,j,c,p}$ is the usual error term.

We estimate equation 1 by using ordinary least squares (OLS), with robust standard errors clustered at the EA level —Moulton (1990). Considering that the dependent variable is a dummy one, the usage of OLS means that the estimated technique takes the form of a linear probability model (LPM). Although the LPM might present some problems in comparison with the usual probit (or logit) regression models —Horrace and Oaxaca (2006)—, Greene (2002) and Bellemare et al.

against the average number of households with electricity for each country and year. In line with our argument, the raw aggregate data and the fitted values suggest a positive relationship between these two variables.

[12] We focus on households that have/do not have an electricity connection to a grid (either to the national grid or to a micro-grid). As we explain later, we rule out from our analysis those few households that have alternative, on-side generation devices, such as solar panels or mini-generators.

[13] An EA is the smallest administrative level included in the sample (similar to census blocks in the United States).

[14] Kenya was divided into eight provinces until 2013, when the provinces were replaced by a system of 47 counties. Since the sample was based on 2009 census data, we keep the provinces in our analysis. We do not include county dummies because all the cities in the sample but one are located in different counties.

(2015) argue that the LPM is a better suited technique in the presence of fixed effects. Moreover, as explained by Allcott and Taubinsky (2015), in typical cases where the true probability model is not known (as is our case), Angrist and Pischke (2008) advocate for using the LPM instead of an arbitrary non-linear model (i.e., probit or logit).[15]

3.2 Potential threats to identification

In equation 1, our coefficient of interest is α_1, which measures the causal impact of the reliability of the electricity supply at the EA level on a household's probability of having a connection. The identification of the causal effect hinges on the assumption that, after controlling for a set of relevant households and EA characteristics that potentially influence households' electrification decisions, more frequent outages reduce their probability of having a connection. Nonetheless, this empirical strategy highlights at least three major threats to identification of which we should be aware.[16]

The first potential threat to identification arises from the concern that the "allocation" of households across EAs is not random. Therefore, one may be concerned that those households that are more willing to purchase an electricity connection (for instance, because they are richer, better educated, or live closer to the grid) are more likely to live in EAs that are less exposed to outages. If this is the case, there may be, thus, a *selection effect* problem, resulting in biased estimates.

To address this potential concern, we use propensity score-matching[17] —Rosenbaum and Rubin (1983). We identify covariates that are likely to affect households' EA choices (income, education, etc.) and estimate, for each household, the *propensity score* of dwelling in an EA that is highly exposed to outages. We end up with sub-samples of households with similar characteristics both in EAs highly affected by outages (*treated*) and in EAs that are not affected by outages (*control*), reducing thus the potential bias created by the aforementioned selection effect.

An additional potential threat to identification relates to the concern that, in some particular neighborhoods, a *marginal household* connected to electricity is likely to cause additional outages by overloading the local system.[18] In particular, this is the case in neighborhoods where informal

[15] We have also checked that our results are not substantially different if we employ instead a probit model —as done by Millien (2017). The set of results using a probit regression model are included in the appendix.

[16] Additional and alternative concerns regarding sample selection are discussed in Section 3.4.

[17] Dehejia and Wahba (2002) explain that propensity score-matching is an appropriate technique to reduce the sample selection bias in non-experimental environments with a rich set of covariates (as it is our case).

[18] As discussed above, a number of outages in Kenya are reported as being caused by local overloads on transformers.

connections are relatively common —as informal connections are typically associated with frequent outages; see Jamil (2013) and Lewis (2015)—, and in neighborhoods where households are likely to use electricity to plug-in high-power appliances (such as refrigerators). These neighborhoods are, thus, a potential source of *endogeneity* ("reverse causality") that bias our OLS coefficients. We attempt to mitigate this concern by re-estimating equation 1 after excluding from the sample EAs in which there is at least one household aware of the existence of informal connections, and in which there is at least one household that owns a refrigerator.[19]

A final potential threat to identification arises due to the fact that there might be some unobserved characteristics that are potentially correlated with the frequency of outages at the EA level ($\theta_{j,c,p}$). If this were the case, one could argue that our main results are not driven by outages frequency, but by such ($\theta_{j,c,p}$-correlated) unobserved characteristics (*hidden bias*).

To address this potential concern, we perform a falsification test using a *fake outcome* that is of "similar nature" as our actual outcome variable $y_{i,j,c,p}$ (have/do not have an electricity connection) but that is known to be unaffected by the main explanatory variable of interest (the frequency of outages) —Rosenbaum (2002). In particular, we use a dummy that captures whether a household has a piped water connection as a fake outcome. By estimating equation 1 on this alternative (fake) outcome, the result of the "treatment" —i.e., exposure to frequent outages— would shed some insight on whether our main result is driven by a hidden bias: if this fake outcome fails to replicate the observed performance for the actual treatment, then we succeed in falsifying the claim that the estimated coefficient of interest (α_1) is driven by the unobservable characteristics.

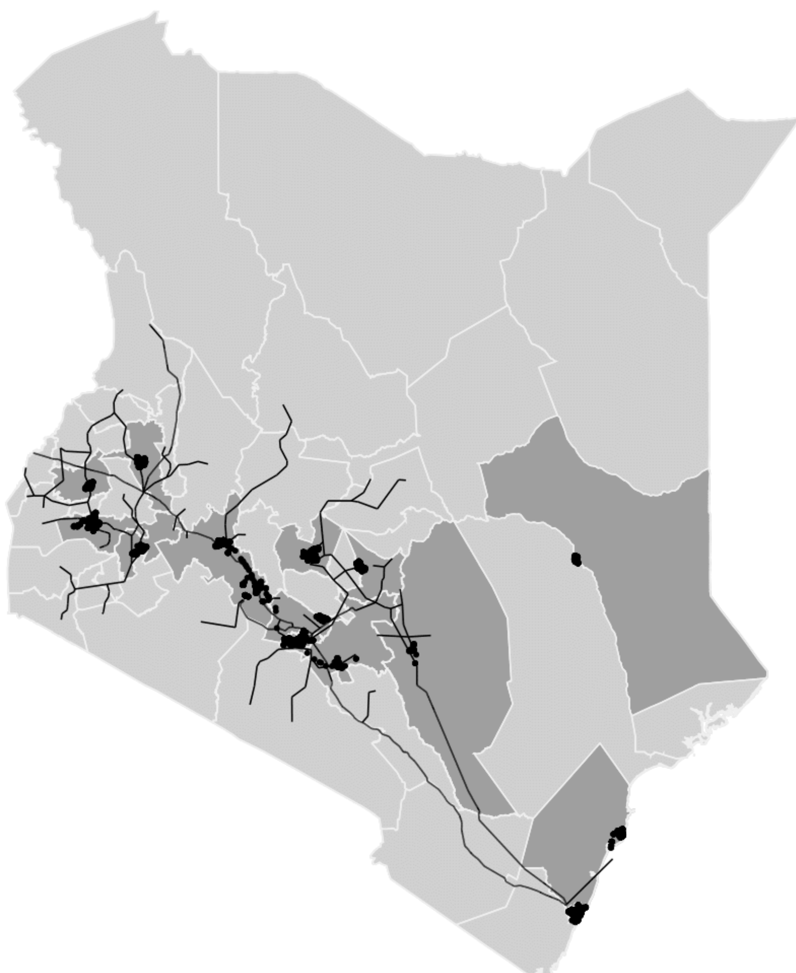
3.3 Data sources

To estimate the coefficients of the previous regression model, we use data from different sources. Our main source of data is the "Kenya - Cities Baseline Survey 2012-2013", collected by the University of Chicago's National Opinion Research Center (NORC) under a contract with the World Bank —Gulyani et al. (2012). The survey contains detailed demographic, infrastructure access, and socio-economic information, as well as the geo-localization of a randomly selected representative

[19] Ideally, as we do in Section 4, one should *instrument* the frequency of outages ($\theta_{j,c,p}$) using a variable that is highly correlated with $\theta_{j,c,p}$ but that does not affect households' electrification decisions, as a potential solution for this concern that does not imply losing some data points. Unfortunately, the instrument that we use in Section 4 —the average number of informal connections in each EA— is unlikely to satisfy the latter condition, as the presence of informal connections potentially increases households' willingness to have a (cheap and illegal) connection at home.

sample of over 14,000 households in and around 15 major urban centers across the eight (former) provinces of Kenya, both in slum and non-slum areas, and in rural and urban areas. The sample of households was selected at the EA level[20] —the smallest geographic census unit in Kenya— in (and around) each metropolitan area, selected from the 2009 census by the Kenya National Bureau of Statistics (KNBS).[21] Figure 1 contains a map with the households included in this study (black dots), and the counties for which data is available (indicated in dark gray).

Figure 1: Map of Kenya with households used in the empirical analysis



Note: The figure shows the map of the households included in the survey. Kenyan counties for which there are households included in the survey are drawn in dark grey (counties for which there are no households included in the survey are drawn in light grey). Each dot on the map represents a household. The black lines represent high voltage or medium voltage transmission lines in Kenya, obtained from the World Bank's Africa Infrastructure Country Diagnostics (AICD) team —AICD (2009).

[20] According to the 2009 Kenya census, there are on average 92.6 households in each of the EAs included in the survey, and 29.2 surveyed households on average per EA. We drop from our sample 495 households for which the identifier of the EA is not available.

[21] Gulyani et al. (2014) and Salon and Gulyani (2019) provide details on the stratification and sampling procedure.

The dataset contains the answers of all the surveyed households to the following question:[22] "is your dwelling unit connected to electricity?" Using the answers to this question, we create our dummy variable of main interest, which is equal to 1 if the answer of household i is "Yes", and is equal to zero if the answer is "No".

Next, we build a variable that measures the frequency of outages for each EA. We use the information on the average weekly outages reported by households that use electricity as their primary source of lighting.[23] In particular, we create a dummy variable that equals 1 if a household that uses electricity as the main source of lighting experiences outages at least once a week, and that equals 0 if a household that uses electricity as the main source of lighting rarely experiences outages. Then, our outages-frequency variable, $\theta_{j,c,p}$, is simply calculated as the average of this dummy at the EA level. That is, for each household i that lives in EA j, city c and province p, we calculate

$$\theta_{j,c,p} = \frac{\sum_{i=1}^{n_{j,c,p}} \mathbb{1}\{\text{Outages at least once a week}\}_{i,j,c,p}}{n_{j,c,p}} \tag{2}$$

where $n_{j,c,p}$ is the total number of households that use electricity as the main source of lighting in each EA.

Figure 2 contains two heatmaps of the outages-frequency variable ($\theta_{j,c,p}$) for all the EAs considered in this study in the Nairobi county —Subfigure 2a— and in the Mombasa county —Subfigure 2b. Areas shaded in dark colors indicate that $\theta_{j,c,p}$ is closer to 1 (greater outages frequency), while areas shaded in light colors indicate that $\theta_{j,c,p}$ is closer to 0 (lower outages frequency). For the particular case of the Nairobi county, darker colors are mostly observed in poorer areas and slums (such as Kibera and Mathare), while lighter colors are observed in wealthier and commercial areas, such as Gigiri and the Central Business District (CBD).

Considering that the outages-frequency variable at the EA level is based on information provided by the households that use electricity as their primary source of lighting, we drop from our sample the households that live in EAs where no household uses electricity for lighting purposes.[24] Still, one might be concerned that in some EAs there are too few households that use electricity to

[22] We drop two households for which there is no answer to this question.

[23] We focus on households that use electricity as the primary source of lighting to create our outages-frequency variable because these households are expected to be more aware of the frequency with which outages occur. As shown in the appendix, our results prove robust to alternative definitions of the outages-frequency variable.

[24] There are 868 households (less than 6%) that live in EAs where no one uses electricity for lighting purposes.

Figure 2: Outages-frequency heatmap by EA in selected counties

(a) Nairobi county

(b) Mombasa county

Note: The figure shows a heatmap of the frequency of outages for each of the EAs that we use in the study both in the Nairobi county (a) and in the Mombasa county (b). The heatmap was built using the average frequency of outages per EA ($\theta_{j,c,p}$), where the geo-localization of each EA was based on the centroid of the surveyed households that live in each EA. Areas shaded in dark colors indicate greater frequency of outages ($\theta_{j,c,p}$ close to 1), while areas shaded in light colors indicate lower frequency of outages ($\theta_{j,c,p}$ close to 0).

light the house (maybe one or two households). If that is the case, one could argue that the main explanatory variable of interest ($\theta_{j,c,p}$) might not be "informative" enough. Therefore, as a robustness check, we also estimate our model after removing those EAs in which less than 10% of the surveyed households use electricity as the main source of lighting.[25]

The survey contains additional household and EA characteristics that are also likely to impact households' electrification decisions. In particular, we include in our estimation the following control variables: household size (number of members); education of the head of the household (a dummy indicating completion of primary education); expenditure in a typical month, in Kenyan Shillings (KSh),[26] two dummies indicating if there are informal connections and street lights in the street where the household lives, a dummy indicating if the EA is in a rural or urban area, a dummy indicating if it is in a slum, number of months that the household has been occupying the house, a dummy indicating if the walls of the house are permanent/robust,[27] and city and province dummies. At the time when the survey was conducted, connection charges were uniform (KSh 35,000) throughout the whole country —Ministry of Energy, Government of Kenya (2012) and Lee et al. (2016).

Finally, following the arguments provided by Van de Walle et al. (2013), Squires (2015) and Blimpo et al. (2018), one might also argue that the proximity of households to electricity infrastructure may also have an impact on whether a household is connected to electricity. Thus, we combine our household-level survey data with geographical information both on the location of high voltage and medium voltage transmission lines, and on the location of power-generating facilities in Kenya. The geo-localization of the transmission lines —which are also included in the map provided in Figure 2— was obtained from the World Bank's Africa Infrastructure Country Diagnostics (AICD) team —AICD (2009).[28] The data on power-generating facilities was obtained from different sources, namely, the World Bank (2014) and the World Resources Institute (2018).[29] Additional details on

[25]Results prove robust to alternative percentages.

[26]Monthly expenditure is computed as the sum of the expenditure in the last month on the following items: food, fuel, clothing and footwear, household supplies, domestic services, transportation, recreation, tobacco, alcohol, insurance, education, taxes, home furnishing and maintenance, and vehicle repair.

[27]Permanent/robust walls are walls built using brick, block or stones (as opposed to non-permanent/non-robust walls, which are those built using mud, wood, tin, or corrugated iron sheets).

[28]This dataset was released in 2009, and (according to the website) the last updated was in 2017. However, the transmission lines included in this dataset exactly coincide with those that were installed at the time when the survey was conducted (2012-2013). For a visual comparison, see Figure 2 and Figure 1 in the report by the Ministry of Energy and Petroleum, Government of Kenya (2016) (which includes a map with the transmission lines back in 2013).

[29]We excluded those power plants that were commissioned after 2013 (when the survey was conducted).

the power plants considered (and a map of them) are included in the appendix.

Using the aforementioned geo-localized data, we compute the distance (as the crow flies) from each household's dwelling unit both to the closest transmission line and to the closest power plant. These two variables ("distance to the closest grid line" and "distance to the closest power plant"), which are also included as controls in our regression analysis, were computed for all the households in the dataset except for those that live in the Garissa county. The reason being that, as shown in the map included in Figure 2, as of 2013 consumers in Garissa county (the Eastern county drawn in dark grey in the map) were not served by the national grid.[30]

3.4 Sample and summary statistics

The "Cities Baseline Survey 2012-2013" contains information on 14,581 households in Kenya. However, besides excluding from our sample 868 households that live in EAs where no one uses electricity as the main source of lighting, we also exclude 90 additional households with off-grid generators at home, and/or with portable batteries that can be charged elsewhere, and then used at home. The reason is that these households are less likely to have on-grid connections, regardless of the frequency of outages.[31] Moreover, we also drop from our sample two households that are on a waiting list to get one, as these households are (presumably) willing to get a connection but, due to bureaucratic and technical barriers, do not have it.

According to the information provided by the survey, approximately half of the households are reported as "not living on a permanent basis in their dwelling units" at the time when the survey was conducted. This issue might be problematic since, for the obvious reasons, one could assume that non-permanent households potentially have less incentives to invest in an electricity connection, regardless of the reliability of the network. Therefore, considering, on the one hand, that these households may potentially underestimate the coefficient of interest, but bearing in mind, on the other hand, that this subset of households is quite large (and potentially informative as well), we present our estimates using both the full set of households and the subset of "permanent" ones.

[30] As of 2016, construction was being undertaken to connect Garissa to the grid (the 132kV transmission line Kindaruma-Mwingi-Garissa line) —Ministry of Energy and Petroleum, Government of Kenya (2016). Meanwhile, most of the households in Garissa get their electricity from local mini-grids (mostly solar).

[31] We acknowledge that this "substitutability" between on-grids connections and off-grid generators/batteries is not unambiguous. In fact, Lee et al. (2016) argue that home solar systems are not a "substitute" for grid power, but rather a "complement" (i.e., as a "back-up" device when on-grid connections fail). Thus, as a robustness check, we include in the appendix our results including also those households that own batteries, solar panels, and/or mini-generators.

Next, we keep in our sample the subset of households that rent a house and whose electricity bill is included in the rent paid to the landlord. We assume that, if a landlord includes the electricity bill in the rent, it is because she expects the electricity supply to be reliable. However, one might also argue that landlords' electrification decisions could have been made irrespective of the frequency of outages —for instance it could have been made in order to charge a higher rent (Choi and Kim, 2012). Therefore, as a robustness check, we also estimate our regression model after removing those households whose rents include the monthly electricity bill (see the appendix). Finally, we also keep in our sample 184 households that do not pay directly to the utility company either, but that rather use prepaid cards. Our results also remain robust after dropping these households.

Table 1: Summary statistics (by sub-sample of households)

	Panel A1: Full sample				
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Hh has electricity	0.682	0.466	0	1	13,102
Outages frequency	0.372	0.302	0	1	13,102
Slum	0.292	0.455	0	1	13,102
Urban	0.851	0.356	0	1	13,102
(log) Month expenditure	8.535	0.996	0	13.874	13,102
Robust/permanent wall	0.42	0.494	0	1	13,102
Distance grid (no Garissa)	6.135	8.521	0.011	35.024	10,337
(log) Months in dwelling	3.459	1.369	0	6.951	13,078
Hh size	2.967	1.824	1	17	13,040
Informal connections	0.08	0.272	0	1	12,894
Head of hh primary educ	0.884	0.32	0	1	13,102
Street lights	0.241	0.428	0	1	13,102
Distance plant (no Garissa)	35.272	29.482	0.236	104.432	10,337

	Panel A2: Permanent households				
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Hh has electricity	0.859	0.348	0	1	4,177
Outages frequency	0.327	0.272	0	1	4,177
Slum	0.258	0.438	0	1	4,177
Urban	0.956	0.205	0	1	4,177
(log) Month expenditure	8.746	1.068	0	13.874	4,177
Robust/permanent wall	0.649	0.477	0	1	4,177
Distance grid (no Garissa)	6.589	6.787	0.011	21.724	2,650
(log) Months in dwelling	3.387	1.302	0	6.733	4,169
Hh size	3.02	1.861	1	17	4,161
Informal connections	0.062	0.241	0	1	4,127
Head of hh primary educ	0.919	0.273	0	1	4,177
Street lights	0.324	0.468	0	1	4,177
Distance plant (no Garissa)	21.117	22.211	0.236	65.504	2,650

Note: The table shows summary statistics for all the households included in our different (sub-)samples. Across panels, the table shows summary statistics for the full sample of households used in the empirical analysis in Panel A1; and summary statistics for all the households included in Panel A1 but excluding non-permanent households in Panel A2. Across columns within each panel, the first column shows means of the variables for all households, the second column includes the standard deviations, the third column includes the minima, the fourth column shows the maxima, and the last column includes the number of observations for each variable. Sections 3.3 and 3.4 offer additional information on the institutional context and the data.

After dropping and keeping all the aforementioned households, our final sample contains 13,102 households, for which we have information on the electrification status, and the city and province in which they live. Summary statistics for all the aforementioned variables for the final sample that we use in our empirical study are included in Table 1, Panel A1. In Panel A2 we include summary statistics after removing the subset of "non-permanent" households.

3.5 Empirical results

Our main empirical results are included in Table 2, which shows the estimated coefficients of the outages-frequency variable ($\theta_{j,c,p}$) —and of all the other controls, when included— on the dummy indicating if a household has an electricity connection ($y_{i,j,c,p}$) —equation 1. We include the (OLS) estimated coefficients with standard errors clustered at the EA level in parenthesis.

Columns (1)–(3) include the estimation results using all the households in our final sample. First, we present the results if no control variables are included —Column (1). In this case, the estimated coefficient is negative and significant at the 5% level. Our estimation suggests that households that live in EAs frequently exposed to power outages on a weekly basis are about 8% less likely to have a connection relative to households that live in EAs that experience no outages.

Next, column (2) includes the set of households' characteristics as controls. In particular, we control for household monthly expenditure (income), education of the head of the household, quality of the dwelling unit (proxied by the quality of the walls), household size, and the number of months that the household has been living in the dwelling unit, together with city and province dummies. Then, in Column (3) we include the full set of household-level and EA-level controls discussed in Section 3.3. In these two cases, our estimates suggest that households living in EAs where power outages are more frequent are about 6% less likely to have an electricity connection.

Finally, in columns (4)-(6) we estimate equation 1 after removing from our sample households that are not reported as permanent dwellers. In this case, our estimates suggest that for households living in EAs where outages are more frequent the probability of having an electricity connection decreases by 9%. This coefficient is significant at the 5% level if just the household characteristics and no other control variables are included; and significant at the 10% level if the full set of controls is included. Consistent with the above discussion, we see that removing those households that do not live on a permanent basis in their dwelling units increases (makes more negative) the estimated

Table 2: Impact of Outages on Electrification Decisions

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
Outages frequency	-0.0789**	-0.0583**	-0.0571*	-0.0921**	-0.0879**	-0.0902*
	(0.0350)	(0.0294)	(0.0328)	(0.0447)	(0.0431)	(0.0509)
(log) Month expenditure		0.129***	0.125***		0.0587***	0.0538***
		(0.00736)	(0.00828)		(0.00836)	(0.0103)
Head of hh primary educ		0.207***	0.208***		0.189***	0.206***
		(0.0147)	(0.0159)		(0.0299)	(0.0379)
Robust/permanent wall		0.228***	0.236***		0.0248	0.0580***
		(0.0150)	(0.0168)		(0.0192)	(0.0215)
Hh size		-0.0240***	-0.0188***		-0.0146***	-0.00409
		(0.00278)	(0.00301)		(0.00480)	(0.00517)
(log) Months in dwelling		-0.0152***	-0.0128***		-0.00640	-0.00937
		(0.00337)	(0.00379)		(0.00537)	(0.00706)
Distance grid (no Garissa)			0.00603*			0.00879*
			(0.00356)			(0.00471)
Informal connections			0.0647***			0.0281
			(0.0210)			(0.0260)
Street lights			0.0699***			0.0717***
			(0.0188)			(0.0189)
Distance plant (no Garissa)			-0.00130			0.00102
			(0.00179)			(0.00323)
Slum			-0.0879***			-0.0889***
			(0.0228)			(0.0239)
Urban			0.112***			0.167***
			(0.0241)			(0.0557)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	13,102	13,016	10,121	4,177	4,153	2,603
R^2	0.048	0.209	0.228	0.032	0.088	0.124

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

effect of outages on households' electrification decisions in all the model specifications and restricted sub-samples relative to the estimates found in columns (1)-(3).

3.6 Robustness checks

3.6.1 Endogeneity and propensity score-matching

As explained above, for the regression model specified in equation 1, there may be some concern with regard to a potential "household-EA" selection problem. In particular, one might argue that households that are more likely to have a connection —for instance, those with higher income or better education— are also more likely to live in EAs where outages are less frequent.[32] Thus, one could argue that the previous coefficients are biased, as they are not capturing the true impact of outages on having electricity at home, but rather the "systematic" differences that exist between households that live in EAs highly exposed to outages and those that live in EAs less exposed to outages.

To address this potential concern, we use propensity score-matching —Rosenbaum and Rubin (1983) and Rosenbaum and Rubin (1984). As explained by Dehejia and Wahba (2002), this technique is an appropriate one to mitigate the selection effect bias if there is availability of a rich set of covariates that explains the selection of individuals into "treatment" (in our case, the selection of households into EAs highly affected by outages). Thus, if relevant differences in households' EA choices are captured by these covariates, denoted X_i (income, education, etc.), matching yields an unbiased estimator. The key assumption for identification is that the information included in X_i is sufficient to make the choice of having an electricity connection independent of the EA choice.

We, thus, divide households into two different groups, depending on whether they live in an EA highly exposed to outages (the *treated group*, denoted $T_i = 1$), or in an EA that never experiences outages (the *control group*, denoted $T_i = 0$). Then, we estimate the probability of living in an EA frequently exposed to outages ($T_i = 1$) conditional on covariates X_i (using a probit regression), and we calculate each household's predicted *propensity score*, denoted $\widehat{p}_i$. Rosenbaum and Rubin (1983) show that the conditional independence result extends to the use of propensity scores. That is, given an outcome of interest y_i (in our case, having an electricity connection),[33] if $y_i \perp T_i | X_i$

[32]This idea is consistent with the findings by Aidoo and Briggs (2019).

[33]For the sake of simplicity and expositional clarity, we omit the additional sub-indices (j, c, and p).

then $y_i \perp T_i | \widehat{p}_i(X_i)$. Therefore, after conditioning the choice of EA on the propensity scores, the estimation of the effect of living in an EA frequently exposed to outages on the probability of having a connection does not capture this potential selection effect bias anymore.

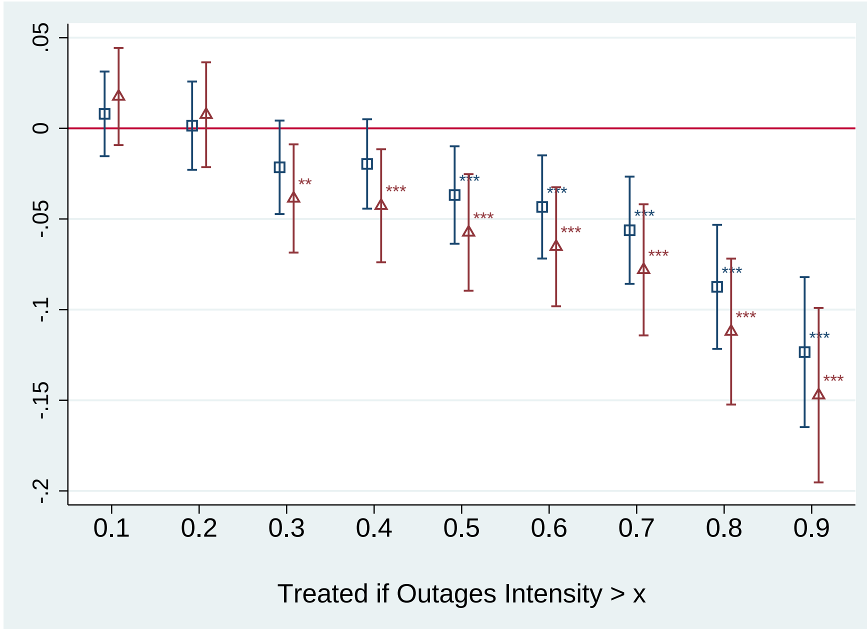
An additional complication arises due to the fact that our "treatment" (i.e. living in an EA frequently exposed to outages) is captured by the variable $\theta_{j,c,p}$, which does not take the values 0 and 1 only, but it rather takes values between 0 and 1 (in a continuous fashion). To overcome this issue, we estimate a rolling version of the propensity score-matching estimator, in which the "treatment" (exposure to frequent outages) is defined in an incremental way. By doing so, we capture the (incremental) effect on electrification decisions for different definitions of treated households.

In particular, we discretize the continuous variable $\theta_{j,c,p}$ in equal bins (or increments) of 0.1. That is, we define $\theta^x_{j,c,p}$ as those EAs whose value of $\theta_{j,c,p}$ is greater than or equal to x, where $x \in \{0.1, \cdots, 0.9\}$. Then, we estimate the effect of outages exposure on the probability of having a connection (relative to *control* households) by defining as *treated* the households that live in EAs with $\theta^x_{j,c,p}$ for every x.[34] That is, first we estimate the effect assuming that the treated households are those for which the outages-frequency variable is greater than (or equal to) 0.1; then, we estimate the effect assuming that the treated households are those for which the outages-frequency variable is greater than (or equal to) 0.2; and so on (until $\theta^{0.9}_{j,c,p}$). If this setup is correct, we should expect an increasing (more negative) effect on the probability of having a connection as x increases; that is, as we consider households that live in EAs that are increasingly affected by outages.

The results of this exercise are included in Figure 3. The blue squares and the red triangles capture the coefficients of the effect of frequent outages on the probability of having a connection for different definitions of *treated households* (as defined by $\theta^x_{j,c,p}$) using all the households in our sample (the vertical lines capture the 95% confidence intervals). To obtain the coefficients captured by the blue squares, we match households using the restricted set of household-level variables —those included in Column (2) in Table 2—, and to obtain the coefficients captured by the red triangles we match households using the full set of household-level variables. As expected, we find that for greater values of x, the impact of outages on households' electrification decisions becomes more negative and more significant. For instance, households that live in EAs where $\theta_{j,c,p}$ is greater than 0.9 are, on average, 13-15% less likely to have an electricity connection in comparison to similar

[34] In all cases, the *control* households are those that live in EAs for which $\theta_{j,c,p}$ is equal to 0.

Figure 3: Propensity-score matching estimator for different "treated" samples



Note: The figure shows the propensity-score matching coefficients for the outages-frequency variable ($\theta_{j,c,p}$) after matching *treated* and *control* households. From left to right, the households that live in EAs for which the outages-frequency variable is greater than $x \in \{0.1, \cdots, 0.9\}$ are considered as treated; and the control households are those that live in EAs for which the outages-frequency variable is equal to zero. Households are matched (using propensity-score matching) on *(i)* income, quality of the dwelling unit (proxied by "robust wall" dummy), head of household education, urban dummy and slum dummy (coefficients are displayed with a square); and on *(ii)* the full set of household-level controls included in our main regressions (coefficients are displayed with a triangle). Vertical bands represent ± 1.96 times the standard error for each point estimate, and the significance levels of the estimated coefficients are as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. These results are obtained using full set of households; very similar results are obtained using alternative sub-samples (see the appendix).

households that live in EAs where outages never occur.[35]

3.6.2 Falsification test

One could also argue that the frequency of outages ($\theta_{j,c,p}$) may be correlated with some other unobservable variables that may also explain whether households have an electricity connection at home. If this were the case, then the effect that we find would not be driven by $\theta_{j,c,p}$, but rather by these $\theta_{j,c,p}$-correlated (unobservable) variables (*hidden bias*) —Rosenbaum (2002).

To avoid this potential criticism, we perform a placebo-based falsification test. More precisely, we estimate equation 1 again but we use a *fake outcome* variable instead that is "similar" to our actual outcome variable (a dummy indicating whether a household has an electricity connection) but that is known to be unaffected by the frequency of outages. In particular, we use a dummy variable that indicates whether a household has/does not have a connection to piped water as our fake outcome. If there are no $\theta_{j,c,p}$-correlated unobserved variables that might potentially explain

[35]Very similar results are obtained if we use instead the sub-set of "permanent" households (see the appendix).

whether a household has a piped water/electricity connection at home, then we should expect the estimated coefficient of $\theta_{j,c,p}$ to be close to zero and not significant for the falsification test.

The results of this falsification test are included in Table 3.[36] In all the model specifications, the estimated coefficients of the frequency of outages ($\theta_{j,c,p}$) are close to zero and not significant. These coefficients, thus, suggest that our results on households' probability of having an electricity connection are unlikely to be driven by some other unobservable explanatory variables correlated with the frequency of outages at the EA-level.

Table 3: Impact of Outages on Electrification Decisions (falsification test)

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
Outages frequency	0.0255	0.0224	0.00562	0.0259	0.0272	0.00669
	(0.0320)	(0.0314)	(0.0303)	(0.0472)	(0.0467)	(0.0529)
Household controls		✓			✓	
Full set of controls			✓			✓
City & prov. dummies	✓	✓	✓	✓	✓	✓
Observations	11,276	11,202	8,604	3,618	3,596	2,102
R^2	0.172	0.201	0.182	0.202	0.211	0.133

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.6.3 Additional robustness checks

Finally, our results are also robust to alternative estimation techniques, different definitions of the main variable of interest, and a variety of sample selection criteria, which can be found in Appendix B. Here, we provide a summary of the extensive analysis performed in that Appendix.

First, we show that our results are robust if we estimate equation 1 using a probit model instead. In Appendix B.1 we provide both the estimated probit coefficients, and also the estimated *marginal effects*. These marginal effects are negative, significant, and very similar in magnitude to the coefficients included in Table 2. Second, our results do not significantly change either if we use alternative definitions of the outages-frequency variable. We check this fact in Appendix B.2

[36] The full table that includes the coefficients of the rest of the control variables is provided in the appendix.

by using two different alternative definitions of this variable: first, we build the outages-frequency variable using not only the outages reported by households that use electricity as the main source of lighting, but including all the households that have electricity (regardless of their primary source of lighting); second, we use households' reported data on the average hours per day that they get electricity. Finally, our results are also robust to alternative sample selection criteria. In particular, we check in Appendix B.3 that our results do not substantially change *(i)* if we exclude EAs in which there is at least one household aware of the existence of informal connections, *(ii)* if we exclude EAs in which at least one household owns a refrigerator, *(iii)* if we include households that own a solar panel, a battery, or a mini-generator, *(iv)* if we exclude EAs in which less than 10% use electricity as the main source of lighting, *(v)* if we exclude those whose electricity bill is included in the rent that they pay to their landlords, and *(vi)* if we exclude households that use prepaid cards.

4 Effect of Outages on Post-electrification Decisions

As explained above, access to electricity does not have an intrinsic value, but it is rather a necessary condition in order to provide a number of benefits to households, such as appliance access (among other things). In this section we focus on the households that have an electricity connection to study whether frequent outages do also have an impact on their subsequent decisions.

4.1 Empirical framework and potential threats to identification

The regression model that we estimate examines the effect of the frequency of outages on the probability with which a household owns certain large electric-powered appliances. In this regression model, the unit of observation is a household i that lives in EA j, city c, and province p. The effect of the frequency of outages in EA j on the probability that household i owns appliance r is measured in the following cross-sectional regression:

$$y^r_{i,j,c,p} = \beta_0 + \beta_1 \theta_{j,c,p} + \beta_2 \boldsymbol{X}_{i,j,c,p} + \beta_3 \boldsymbol{Z}_{j,c,p} + \lambda_p + \lambda_c + \varepsilon_{i,j,c,p} \tag{3}$$

where $y^r_{i,j,c,p}$ is a dummy variable that is equal to 1 if household i owns appliance r, and 0 otherwise; and $\theta_{j,c,p}$, $\boldsymbol{X}_{i,j,c,p}$, $\boldsymbol{Z}_{j,c,p}$, λ_p, λ_c and $\varepsilon_{i,j,c,p}$ are as defined in Section 3.1 for equation 1.

We estimate equation 3 by using OLS, with standard errors clustered at the EA level. In this regression, β_1 is the coefficient of interest, which measures the causal impact of the frequency of outages at the EA level on the probability with which households own electric-powered appliance r. Identification of this causal effect hinges on the assumption that, after controlling for a set of relevant characteristics that potentially explain whether a household owns appliance r, more frequent outages reduce the probability of having said appliance. However, this empirical strategy highlights at least one major threat to identification of which we should be aware.

In particular, one may be concerned that in neighborhoods (EAs) where a substantial number of households own large appliances, outages are more likely to occur. This is because local system overloads (which occur when the electrical load exceeds the supply) are a usual cause of blackouts in developing countries.[37] Therefore, one could hypothesize that causality also runs from appliance ownership to power outages at the EA level, leading thus to biased and inconsistent estimates.

To correct for this potential *reverse causality* problem, we estimate equation 3 using Instrumental Variables (IVs). In particular, we use the average number of informal electricity connections at the EA level as an instrument for power outages. The reason is that, as abundant literature has documented, informal electricity connections and thefts are typically associated with frequent outages —see, for instance, Lewis (2015) and Jamil (2013). Thus, the *exclusion restriction* is satisfied if we consider that the presence of a large number of illegal connections in an EA is uncorrelated with households' appliance choices, while it positively affects the frequency of outages on that EA.

Finally, one might also be also worried about unobserved characteristics correlated with the frequency of outages. For that reason, we also perform a falsification test for the appliances regressions, using as fake outcomes appliances that do not require an electricity connection.[38]

4.2 Data sources and samples

To estimate equation 3, we use the same data sources and variables that are extensively explained in Section 3.3. To this dataset, we add information on the households' ownership of appliances, obtained also from the "Cities Baseline Survey". As explained above, we focus on two of the most

[37] For the particular case of Kenya, Taneja (2017) documents that, while the more than half of the outages are due to lack of O&M, still about one in every three of them are caused by system overloads.

[38] The results of this falsification test together with some other robustness checks are included in Appendix C.

desired appliances among Kenyans, namely refrigerators and televisions —Lee et al. (2016).[39] In particular, we use the answers to the following questions: "does your household own a refrigerator?" and "does your household own a television?". Our dummy variables of interest are equal to 1 if the answers to these questions are "Yes", and are equal to zero if the answers are "No".

We also add additional control variables that are likely to explain whether a household owns a refrigerator or a television. For the refrigerator regression, we add a dummy variable indicating whether household i's dwelling unit has a kitchen, and a dummy indicating if there is a food shop in household i's neighborhood (within a 20 minute walk from household i's dwelling unit) — households that live far from a food shop may need the refrigerator to store food for longer periods. For the television regression, we add a dummy indicating whether there is a park in household i's neighborhood —a park is potentially a substitute for a television for households with kids.[40]

Table 4: Summary statistics (households with electricity only)

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Refrigerator	0.118	0.322	0	1	8,938
Television	0.675	0.468	0	1	8,938
Outages frequency	0.351	0.285	0	1	8,938
Slum	0.259	0.438	0	1	8,938
Urban	0.889	0.314	0	1	8,938
(log) Month expenditure	8.731	0.983	0	13.874	8,938
Robust/permanent wall	0.481	0.5	0	1	8,938
(log) Months in dwelling	3.376	1.315	0	6.951	8,917
Hh size	2.928	1.755	1	17	8,898
Head of hh primary educ	0.932	0.253	0	1	8,938
Kitchen in dwelling	0.329	0.47	0	1	8,937
Food shop nearby	0.971	0.166	0	1	8,937
Park nearby	0.146	0.353	0	1	8,933

Note: The table shows summary statistics for the subset of households that have an electricity connection at home. The first column shows means of the variables for all households, the second column includes the standard deviations, the third column includes the minima, the fourth column shows the maxima, and the last column includes the number of observations for each variable. Section 4.2 includes additional information on the data.

We restrict our sample to households that have an electricity connection at home. The reason is that the appliances that we consider (i.e., refrigerator and television) are useless for households that lack access to electricity. Hence, it is unlikely for these households to purchase them (irrespective of the frequency of outages in their EAs). Table 4 contains summary statistics for all the variables used to estimate equation 3 for the sample of households that have an electricity connection.

[39] Previous authors show that these appliances have a clear positive welfare, health and social impact among those that own them —Heller et al. (2005), Jensen and Oster (2009) and Inter-American Development Bank (2019).

[40] We exclude as controls some other variables that are unlikely to affect households' appliances choices (e.g. distance to power plants, street lights, etc.).

4.3 Empirical results

The main panel of Table 5 contains both the OLS and the (second-stage) IV estimation results for the appliances regressions, with standard errors clustered at the EA level in parenthesis. The bottom panel of Table 5 includes the first-stage estimation results for the IV regressions.

Columns (1)-(3) include the estimates for the refrigerator regressions. While column (1) contains the estimation results, including as controls the set of household characteristics, column (2) includes the full set of control variables considered. In both cases, we find a negative and significant (at the 1% level) effect of outages on refrigerators ownership. Our estimates suggest that households in EAs that experience frequent outages on a weekly basis are about 7%-11% less likely to own a refrigerator relative to households in EAs that experience no outages. However, the second stage of the IV regression suggests that the impact of outages on the ownership of refrigerators is actually larger. When using the instrument, we find that this probability increases to about 33%.

In columns (4)-(6) we include the estimates for the television regressions. The OLS coefficients suggest that households that live in an EA that frequently experiences outages are about 10% —if we include the household-level controls; column (4)— and 8% —if we include the full set of controls; column (5)— less likely to own a television relative to households that live in EAs with no exposure to outages. These results are again statistically significant at the 1% level. Once again, the IV regression —column (6)— suggests that this effect is potentially underestimated in the OLS regression. The coefficient of the second stage of the IV regression is about 5 times larger, suggesting that the probability of television ownership decreases by about 53%.

4.4 Robustness checks

The empirical results on the effect of outages on appliance ownership included in Table 5 are robust to a variety of robustness checks, which can be found in Appendix C. First, we show that our results do not significantly change if we use the alternative definitions of the outages-frequency variable that we discuss in Section 3.6.3 —see Appendix C.1. Second, our results are also robust to the same alternative sample selection criteria discussed above —see Appendix C.2. Finally, we also show that the coefficients are close to zero and not significant in a falsification test, in which we use appliances that do not require electricity as fake outcomes —see Appendix C.3.

Table 5: Impact of Outages on Appliances Ownership

	Refrigerator			*Television*		
	(1) OLS	(2) OLS	(3) IV	(4) OLS	(5) OLS	(6) IV
OLS and second stage IV estimates: Dependent variables is $y^r_{i,j,c,p}$						
Outages frequency	-0.109***	-0.0690***	-0.333***	-0.0954***	-0.0780***	-0.528**
	(0.0216)	(0.0187)	(0.124)	(0.0254)	(0.0251)	(0.229)
(log) Month expenditure	0.110***	0.0776***	0.0788***	0.127***	0.124***	0.124***
	(0.00843)	(0.00689)	(0.00688)	(0.00810)	(0.00805)	(0.00839)
Head of hh primary educ	0.0519***	0.0472***	0.0482***	0.146***	0.141***	0.143***
	(0.00950)	(0.00917)	(0.00959)	(0.0193)	(0.0188)	(0.0195)
Robust/permanent wall	0.0460***	0.0227**	0.0228**	0.0918***	0.0803***	0.0796***
	(0.0120)	(0.0105)	(0.0109)	(0.0163)	(0.0160)	(0.0163)
Hh size	-0.00401	-0.00666***	-0.00495**	0.0239***	0.0242***	0.0270***
	(0.00253)	(0.00220)	(0.00247)	(0.00305)	(0.00304)	(0.00352)
(log) Months in dwelling	0.0153***	0.00580**	0.00739**	0.0292***	0.0293***	0.0311***
	(0.00300)	(0.00291)	(0.00309)	(0.00411)	(0.00409)	(0.00428)
Kitchen in dwelling		0.216***	0.206***			
		(0.0130)	(0.0135)			
Food shop nearby		-0.00274	0.0119			
		(0.0195)	(0.0209)			
Slum		-0.0312***	-0.0108		-0.0716***	-0.0346
		(0.00991)	(0.0138)		(0.0159)	(0.0240)
Urban		0.0307***	0.0328**		0.0210	0.0249
		(0.0115)	(0.0140)		(0.0196)	(0.0232)
Park nearby					0.000403	0.00362
					(0.0149)	(0.0171)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	8,877	8,876	8,876	8,877	8,872	8,872
R^2	0.163	0.247	0.182	0.150	0.154	0.083
First stage IV estimates: Dependent variables is $\theta_{j,c,p}$						
Informal connections			0.265***			0.271***
			(0.0623)			(0.0625)

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5 Conclusions and Policy Implications

Over the past years, many governments (usually in cooperation with third parties) have allocated billions of dollars to expand the grid in developing countries (in general) and in SSA (in particular). However, the actual increase in the amount of households with an electricity connection at home has been paltry. This disconnection between the supply and demand of electricity infrastructure has recently motivated researchers —Bhatia and Angelou (2015) and United Nations (2019)— as well as practitioners in the field[41] to move beyond the binary metric of providing electricity access —defined as having an electric grid within reach— to incorporate additional dimensions of customer experience, with quality of supply being a key one. However, while there is substantial research concerning the impact of the reliability of power supply on firms' electrification decisions to support this new dimension, little existing work illuminates the extent to which power quality also affects households' electrification decisions.

Using survey data obtained from over 14,000 geo-localized households in Kenya, in combination with data on the Kenyan electricity infrastructure, we provide robust evidence that inadequate power supply quality limits consumption in two ways, namely, by reducing a household's likelihood of having an electricity connection and by discouraging connected ones from purchasing electrical appliances. Consistently with Jacome et al. (2019), our findings reveal that supply reliability (or a lack thereof) is a key factor in determining the benefits that electrification can deliver to households and, thus, directly contributes to explaining the so-called households' "energy access dividend" — that is, the quantification of the benefits of electrification (SEforALL and Power for All, 2017).

From an energy-policy perspective, our findings highlight some clear policy implications. First, contrary to the usual practice in the field, grid expansion projects should be evaluated not only on the basis of the number of additional households which would come within reach of an electric grid, nor just those that would have a connection at home: the reliability of the new infrastructure should be also tracked and, moreover, whether the new customers are able to use their connections on a daily basis to light their homes and run basic appliances (lamps, chargers, refrigerators, etc.) should also be monitored. These evaluation metrics will provide more precise information on the

[41] For instance, one of the *Project Development Objectives* of the "Kenya Electricity Expansion Project (KEEP)" (launched in October 2010, which provided over $400 million in credits and grants to this country), which was to "expand access to electricity in urban, peri-urban and rural areas", was replaced by "new consumers connected to the grid" —World Bank (2018).

actual impact of grid expansion projects on a household's well-being.

Moreover, our results suggest that government efforts should not be placed only in the expansion of the grid: indeed, this may not result in an optimal allocation of fundings if the supply is not reliable, as these grid expansion efforts will not be fully translated into actual household benefits. Thus, we suggest that policy-makers, institutions and donors should consider —either as stand-alone projects or as a complementary part of grid expansion projects— allocating some funds to improving the reliability of the grid. In fact, and as a final policy-implication, our results suggest that the benefits of doing so (i.e., improving the electricity grid) may have been previously underestimated inasmuch as previous scholars did not consider that greater reliability was directly associated with an increase in both the likelihood of a household having an electricity connection and in it using domestic electric-powered appliances that are directly linked to the welfare of said household.

References

AICD (2009). Kenya Electricity Transmission Network (release date: January 1, 2009; last updated: July 22, 2017). https://datacatalog.worldbank.org/dataset/kenya-electricity-transmission-network. Africa Infrastructure Country Diagnostics (AICD), World Bank Group.

Aidoo, K. and R. C. Briggs (2019). Underpowered: Rolling blackouts in Africa disproportionately hurt the poor. *African Studies Review 62*(3), 112–131.

Allcott, H., A. Collard-Wexler, and S. D. O'Connell (2016). How do electricity shortages affect industry? Evidence from India. *American Economic Review 106*(3), 587–624.

Allcott, H. and D. Taubinsky (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review 105*(8), 2501–38.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Bellemare, M. F., L. Novak, and T. L. Steinmetz (2015). All in the family: Explaining the persistence of female genital cutting in West Africa. *Journal of Development Economics 116*, 252–265.

Bernard, T. (2010). Impact analysis of rural electrification projects in sub-Saharan Africa. *The World Bank Research Observer 27*(1), 33–51.

Bhatia, M. and N. Angelou (2015). *Beyond connections: energy access redefined*. World Bank.

Blimpo, M., S. McRae, and J. Steinbuks (2018). Why are connection charges so high? an analysis of the electricity sector in Sub-Saharan Africa.

Blimpo, M. P. and M. Cosgrove-Davies (2019). *Electricity access in Sub-Saharan Africa: Uptake, reliability, and complementary factors for economic impact*. World Bank Publications.

Burgess, R., M. Greenstone, N. Ryan, and A. Sudarshan (2020). The consequences of treating electricity as a right. *Journal of Economic Perspectives 34*(1), 145–69.

Chakravorty, U., M. Pelli, and B. U. Marchand (2014). Does the quality of electricity matter? Evidence from rural India. *Journal of Economic Behavior & Organization 107*, 228–247.

Choi, S. J. and S. Kim (2012). Why Do Landlords Include Utilities in Rent? Evidence from the 2000 Housing Discrimination Study (HDS) and the 2002 American Housing Survey (AHS). *Journal of Housing Economics 21*(1), 28–40.

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.

De Nooij, M., C. Koopmans, and C. Bijvoet (2007). The value of supply security: The costs of power interruptions: Economic input for damage reduction and investment in networks. *Energy Economics 29*(2), 277–295.

Dehejia, R. H. and S. Wahba (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics 84*(1), 151–161.

Dzansi, J., S. L. Puller, B. Street, and B. Yebuah-Dwamena (2018). The Vicious Circle of Blackouts and Revenue Collection in Developing Economies: Evidence from Ghana.

Eberhard, A., V. Foster, C. Briceño-Garmendia, F. Ouedraogo, D. Camos, and M. Shkaratan (2008). Underpowered: the state of the power sector in Sub-Saharan Africa.

EIA (2019). Electric power monthly with data for december 2019. Technical report, U.S. Energy Information Administration.

Enerdata (2017). Review of the African Solar Market. public webinar, February 7.

Farquharson, D., P. Jaramillo, and C. Samaras (2018). Sustainability implications of electricity outages in sub-Saharan Africa. *Nature Sustainability 1*(10), 589–597.

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.

Greene, W. H. (2002). The behavior of the fixed effects estimator in nonlinear models.

Gulyani, S., W. Ayres, R. Struyk, and C. Zinnes (2012). Kenya State of the Cities Baseline Survey 2012-2013. Ref: KEN-2012-SCBL-v01-M. World Bank & NORC.

Gulyani, S., W. Ayres, R. Struyk, and C. Zinnes (2014). Kenya State of the Cities Baseline Survey: Overview Report. World Bank & NORC: Washington, DC, USA, 2014.

Heller, L., E. A. Colosimo, and C. M. Antunes (2005). Setting priorities for environmental sanitation interventions based on epidemiological criteria: a Brazilian study. *Journal of water and health 3*(3), 271–281.

Horrace, W. C. and R. L. Oaxaca (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters 90*(3), 321–327.

Huenteler, J., I. Dobozi, A. Balabanyan, and S. G. Banerjee (2017). Cost recovery and financial viability of the power sector in developing countries: A literature review.

IEA (2014). Africa Energy Outlook. Technical report, International Energy Agency.

IEA (2015). World Energy Outlook 2015. Technical report, International Energy Agency.

IEA (2017). Energy Access Outlook 2017: From Poverty to Prosperity (an executive summary). Technical report, International Energy Agency.

IEA (2018). World Energy Outlook 2018. Technical report, International Energy Agency.

IEG (2015). World Bank Group support to electricity access, FY2000-2014 : an independent evaluation: Main report (English). Washington, D.C. : World Bank Group. Technical report, Independent Evaluation Group.

Inter-American Development Bank (2019). The Energy Access Dividend in Honduras and Haiti. Technical report, Inter-American Development Bank Monograph 743, by Natacha C. Marzolf (director), Emily L. Pakhtigian, Eric Burton, Marc Jeuland, Subhrendu K. Pattanayak, Jonathan Phillips, Christine Eibs Singer, Hadley Taylor, Michelle Hallack, Javier Cuervo, Carlos Jacome.

Jacome, V., N. Klugman, C. Wolfram, B. Grunfeld, D. Callaway, and I. Ray (2019). Power quality and modern energy for all. *Proceedings of the National Academy of Sciences 116*(33), 16308–16313.

Jamil, F. (2013). On the electricity shortage, price and electricity theft nexus. *Energy policy 54*, 267–272.

Jensen, R. and E. Oster (2009). The power of TV: Cable television and women's status in India. *The Quarterly Journal of Economics 124*(3), 1057–1094.

Kennedy, R., A. Mahajan, and J. Urpelainen (2019). Quality of service predicts willingness to pay for household electricity connections in rural India. *Energy policy 129*, 319–326.

Lee, K., E. Brewer, C. Christiano, F. Meyo, E. Miguel, M. Podolsky, J. Rosa, and C. Wolfram (2016). Electrification for "under grid" households in rural Kenya. *Development Engineering 1*, 26–35.

Lee, K., E. Miguel, and C. Wolfram (2016). Appliance ownership and aspirations among electric grid and home solar households in rural Kenya. *American Economic Review 106*(5), 89–94.

Lee, K., E. Miguel, and C. Wolfram (2019). Experimental Evidence on the Economics of Rural Electrification. *Journal of Political Economy*.

Lewis, F. B. (2015). Costly 'throw-ups': Electricity theft and power disruptions. *The Electricity Journal 28*(7), 118–135.

Millien, A. (2017). Electricity supply reliability and households decision to connect to the grid. Technical report, Working Paper 192, Fondation pour les etudes et recherches sur le development international (FERDI).

Ministry of Energy and Petroleum, Government of Kenya (2016). Current Activities and Challenges to Scaling up Mini-grids in Kenya. Technical report, Ministry of Energy and Petroleum (Government of Kenya), World Bank and Energy Sector Management Assistance Program (ESMAP).

Ministry of Energy, Government of Kenya (2012). National Energy Policy. Technical report, Ministry of Energy, Government of Kenya.

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The review of Economics and Statistics*, 334–338.

Oseni, M. O. and M. G. Pollitt (2015). A firm-level analysis of outage loss differentials and self-generation: Evidence from african business enterprises. *Energy Economics 52*, 277–286.

Oyuke, A., P. H. Penar, and B. Howard (2016). Off-grid or'off-on': Lack of access, unreliable electricity supply still plague majority of africans.

Palit, D. and K. R. Bandyopadhyay (2016). Rural electricity access in South Asia: Is grid extension the remedy? A critical review. *Renewable and Sustainable Energy Reviews 60*, 1505–1515.

Rigobon, R. and T. M. Stoker (2007). Estimation with censored regressors: Basic issues. *International Economic Review 48*(4), 1441–1467.

Rosenbaum, P. (2002). Observational Studies, 2nd edn Springer. *New York, New York, USA*.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Rosenbaum, P. R. and D. B. Rubin (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association 79*(387), 516–524.

Salon, D. and S. Gulyani (2019). Commuting in Urban Kenya: Unpacking Travel Demand in Large and Small Kenyan Cities. *Sustainability 11*(14), 3823.

SEforALL and Power for All (2017). Why Wait? Seizing the Energy Access Dividend. Washington, D.C. License: NonCommercial—NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

Squires, T. (2015). The impact of access to electricity on education: evidence from Honduras. *Job Market Paper, Brown University*.

Steinbuks, J. and V. Foster (2010). When do firms generate? Evidence on in-house electricity supply in Africa. *Energy Economics 32*(3), 505–514.

Taneja, J. (2017). Measuring electricity reliability in kenya. Technical report, Working paper.

United Nations (2019). Tracking SDG 7.1 with the Multi-tier Framework Measuring Energy. Technical report, Policy Brief #15. Department of Economic and Social Affairs, United Nations.

Van de Walle, D., M. Ravallion, V. Mendiratta, and G. Koolwal (2013). *Long-term impacts of household electrification in rural India*. The World Bank.

WEC (2013). World Energy Resources: 2013 Survey. *Wold Energy Council (WEC), London*.

World Bank (2014). A curated list of datasets for the World Bank Negawatt Challenge competition in Accra and Nairobi cities. https://datahub.io/dataset/kenya-geolocalized-power-facilities-2014. Negawatt challenge. Data complied from the Kenya Power annual report 2014, the Kenyan Energy Regulatory Commission and Wikipedia for some geolocalizations (World Bank Group).

World Bank (2017). Annual Report 2016. Technical report, World Bank.

World Bank (2018). Implementation Completion and Results Report IDA-47430, IDA-58440 on credits and grants froms the global partnership on output based aid to the Republic of Kenya for an Electricity Expansion Project. Technical report, Energy & Extractives Global Practice, Africa Region, World Bank.

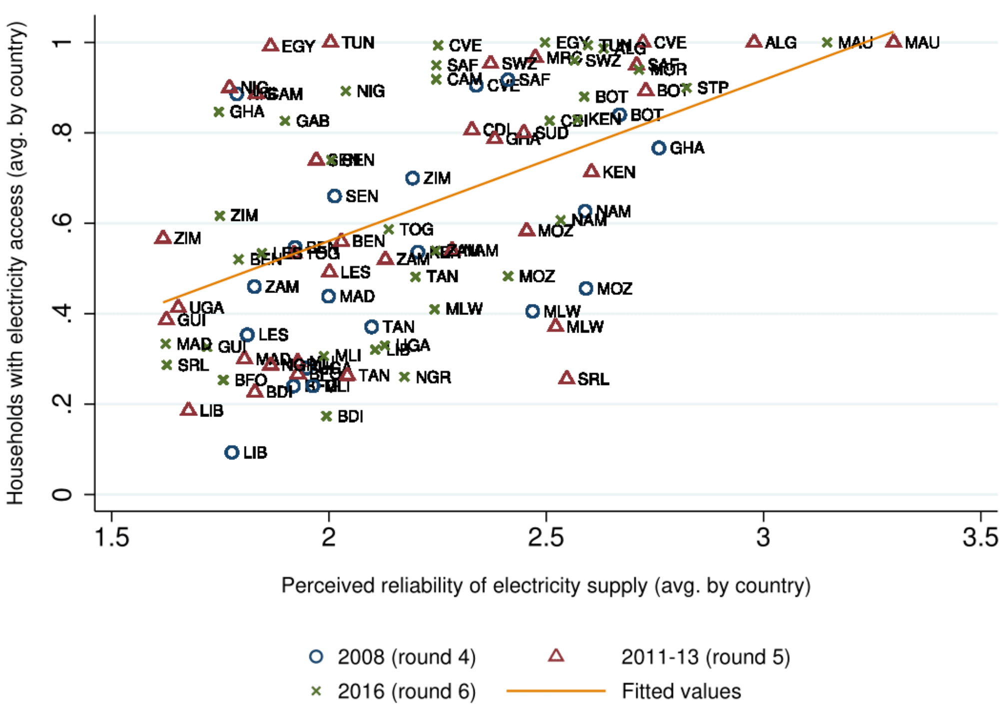
World Resources Institute (2018). Global Power Plant Database. https://datahub.io/dataset/kenya-geolocalized-power-facilities-2014. Global Energy Observatory, Google, KTH Royal Institute of Technology in Stockholm, Enipedia, World Resources Institute. Published on Resource Watch and Google Earth Engine; http://resourcewatch.org/ https://earthengine.google.com/.

Appendix A. Additional Figures and Tables

A.1 Preliminary evidence on the relationship between reliability and connections

In Figure A.1 we plot an index of perceived quality of supply against the average number of households with electricity for 37 selected African countries. The data was obtained from the Afrobarometer survey data —rounds 4, 5, and 6 (years 2008, 2011-13, and 2016 respectively). The raw aggregate data and the fitted values suggest a positive relationship between the perceived quality of supply and the average number of households with an electricity connection.

Figure A.1: Perceived reliability of electricity supply vs. access to electricity (Afrobarometer, 2008-2016)

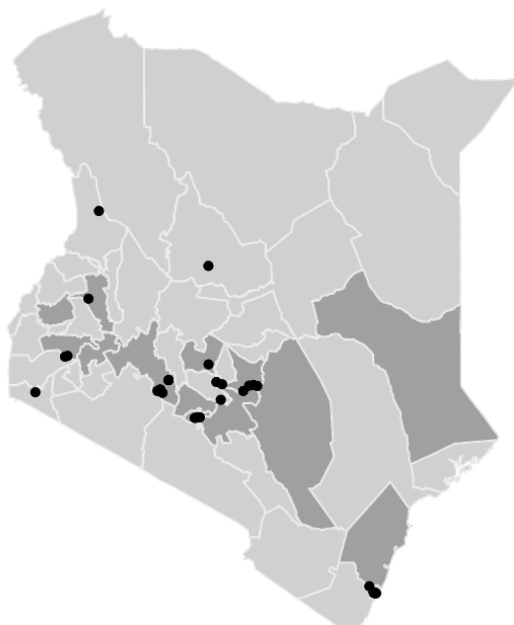


A.2 Additional details on the power plants considered in the analysis

As explained in the main text, the data on power-generating facilities was obtained from different sources. First, we include in our empirical analysis all the power-generating facilities included in

World Bank (2014)[42] However, the geographic coordinates of some of the power plants included in this database are missing. For that reason, we combine the information on power-generating facilities from World Bank (2014) with the geo-localization of the Kenyan power plants included in the World Resources Institute (2018) database.[43] Finally, we also add the geo-localization of five major hydroelectric facilities in Kenya that were commissioned before 2014 and that are not included in the previous databases. Figure A.2 contains a map of Kenya with the all the power-generating facilities included in this study (black dots). Table A.1 contains further details on the additional hydroelectric facilities considered that are not included in World Bank (2014) or in World Resources Institute (2018).

Figure A.2: Map of Kenya with the location of the power plants used in the empirical analysis



Note: The figure shows the map of the Kenyan power plants included in the empirical analysis. Kenyan counties for which there are households included in the survey are drawn in dark grey (counties for which there are no households included in the survey are drawn in light grey). Each dot on the map represents a power plant.

A.3 Distribution of permanent and non-permanent households

Figure A.3 includes the empirical cumulative distribution function (cdf) of the outages-frequency variable ($\theta_{j,c,p}$). We plot the cdf both for the sample of households that do not live on a permanent

[42] According to the information provided by the World Bank, this information was obtained from the Kenya Power and Lighting Company's 2014 annual report, and from the Kenyan Energy Regulatory Commission.

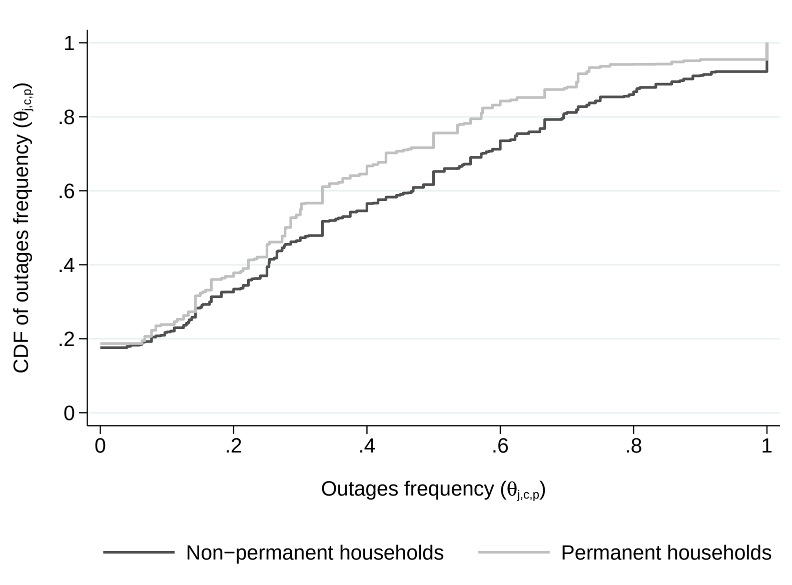
[43] We exclude the power plants that were commissioned after 2013 (when the survey was conducted).

Table A.1: Additional major hydroelectric power stations considered that are not included in World Bank (2014) or in World Resources Institute (2018)

Power plant name	**Installed capacity (MW)**	**City**	**County**	**Coordinates**
Wanjii Power Station	7.4	Kamuiru	Muranga	-0.74944, 37.17472
Ndula Power Station	2.0	Ngoliba	Kiambu	-1.02639, 37.24333
Gogo Power Station	2.0	Kajulu	Migori	-0.90932, 34.34919
Sagana Power Station	1.5	Mutathini	Nyeri	-0.47418, 37.05236
Sosian Power Station	0.4	Eldoret	Uasin Gishu	0.54936, 35.17855

basis in their dwelling units (black line) and for the sample of households that live on a permanent basis in their dwelling units (gray line). The distribution of $\theta_{j,c,p}$ is fairly similar for both subsets of households. The average of $\theta_{j,c,p}$ for "non-permanent" households is 0.39 (with standard deviation equal to 0.31), and the average of $\theta_{j,c,p}$ for "permanent" households is 0.33 (with standard deviation equal to 0.27).

Figure A.3: Empirical cdf of the distribution of the outages-frequency variable ($\theta_{j,c,p}$) across non-permanent and permanent households



Note: The figure shows the empirical cdf of the outages-frequency variable ($\theta_{j,c,p}$). The black line provides the cdf using only households that do not live on a permanent basis in their dwelling units. The gray line provides the cdf using only households that do live on a permanent basis in their dwelling units.

Appendix B. Effect of Outages on Electricity Connections: Robustness Checks

Although we consider that our empirical strategy is strong enough to support the causal link between the frequency of outages and the probability with which households have an electricity connection, we recognize there are alternative methodologies, definitions of the outcome variable, and subsamples that we could have been employed. Hence, in this appendix we engage in an extensive procedure of robustness checks, in order to verify that the results are not sensitive to the choices used throughout the paper.

B.1 Non-linear, binary regression model

In Section 3, we estimate the impact of the frequency of outages on the probability with which households have an electricity connection (equation 1) by using ordinary least squares (OLS). Given the binary nature of the outcome variable, the usage of OLS means that the estimated technique takes the form of a linear probability model (LPM). We did so because many authors have recently supported the usage of the LPM, especially *(i)* if the true probability model is not known, and *(ii)* in the presence of fixed effects —see, among others, Greene (2002), Angrist and Pischke (2008), Bellemare et al. (2015), and Allcott and Taubinsky (2015). However, some other authors have previously argued that the LMP might also present some bias and inconsistency problems in comparison to the usual probit (or logit) regression model —Horrace and Oaxaca (2006).

To double-check that our results are not significantly affected by the model choice, in Table B.1 (Panel A) we include the estimated coefficients of equation 1 using instead a probit regression model.[44] In this case, the coefficient of interest is still negative and significant at the 1% level in all the sub-samples considered.

Notice, however, that the estimated coefficients using the probit regression model are substantially different from those included in Table 2. The reason is that the coefficients obtained in a probit regression do not have a straightforward interpretation. For that reason, in Table B.1 (Panel B) we include the estimated marginal effects, which can be easily compared with those obtained

[44] Similar results are obtained using a logit regression model. Results are available upon request.

Table B.1: Impact of Outages on Electrification Decisions (probit regression)

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Probit coefficients			
Outages frequency	-0.229**	-0.235**	-0.235**	-0.370**	-0.396**	-0.397*
	(0.0997)	(0.0978)	(0.108)	(0.178)	(0.176)	(0.204)
			Panel B: Probit marginal effects			
Outages frequency	-0.0809**	-0.0796**	-0.0816**	-0.0791**	-0.0786**	-0.0796*
	(0.0351)	(0.0329)	(0.0373)	(0.0384)	(0.0352)	(0.0412)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	13,101	13,015	10,121	4,176	4,152	2,603
Pseudo R^2	0.039	0.182	0.199	0.037	0.101	0.142

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

using the LPM model. The estimated marginal effects are negative, significant and very similar in magnitude to those obtained in Table 2.

B.2 Alternative definitions of frequency of outages

In Section 3.3, we build the frequency of outages for each EA ($\theta_{j,c,p}$) using data from households that use electricity as their primary source of lighting. We now provide evidence that our main empirical results are similar to the empirical results included in the main text if we use alternative definitions of the outages-frequency variable.

First, we build an alternative outage frequency variable using information reported not just by the subset of households that use electricity as their primary source of lighting, but using the information reported by the full subset of households that have electricity at home. That is,

$$\theta_{j,c,p} = \frac{\sum_{i=1}^{n_{j,c,p}} \mathbb{1}\{\text{Outages at least once a week}\}_{i,j,c,p}}{n_{j,c,p}} \tag{B.1}$$

where $n_{j,c,p}$ is the total number of households that are connected to electricity for each EA. The empirical results using this alternative definition of the frequency of outages are included in Table

B.2. The estimated coefficients are extremely similar to those obtained in Table 2.

Table B.2: Impact of Outages on Electrification Decisions (alternative outages-frequency)

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
Outages frequency (alt.)	-0.0777**	-0.0569*	-0.0578*	-0.0854*	-0.0819*	-0.0910*
	(0.0351)	(0.0295)	(0.0330)	(0.0455)	(0.0436)	(0.0515)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	13,130	13,044	10,126	4,192	4,168	2,605
R^2	0.047	0.208	0.228	0.030	0.087	0.124

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Second, we use households' reported data on the average number of hours per day that they get electricity in their dwelling units. Using this information included in the survey, we build the following alternative measure of the outages-frequency variable:

$$\theta_{j,c,p} = \frac{\sum_{i=1}^{n_{j,c,p}} \{\#\text{ hours of electricity per day}\}_{i,j,c,p}}{n_{j,c,p}} \tag{B.2}$$

where $n_{j,c,p}$ is the total number of households that have an electricity connection for each EA.

Unfortunately, the variable that captures this alternative definition of outages-frequency is "censored"; that is, there is a disproportionate number of EAs for which the average number of hours per day of power supply is equal to 24, which is also the upper bound of this variable. More precisely, among all the households considered, over 40% of them live in an EA for which this variable is equal to 24. This feature of our regressor of interest poses potential problems to identification and estimation, as discussed by Rigobon and Stoker (2007). Therefore, following these authors, we propose the following alternative estimation strategy. We create two variables; the first one is a dummy variable that is equal to 1 if $\theta_{j,c,p}$ is equal to 24, and 0 otherwise. The second one is a continuous variable that is equal to the actual value of $\theta_{j,c,p}$ if it is not equal to 24, and 0 otherwise. The coefficients of these two variables are expected to be positive and significant. That is, for the former variable (the dummy one), EAs that receive electricity 24 hours per day (on average) are expected to increase the probability with which a household that lives in such an EA has an electricity connection. For the latter variable (the continuous one), having an additional hour of electricity per day at the EA level is also expected to increase a household's probability of having a connection.

The empirical results using these two variables are included in Table B.3. Consistently, we find that the coefficients of both variables are positive. These coefficients are statistically different to zero in all the model specifications using the full sample of households, and also statistically different to zero using the subsample of "permanent" households in all the model specifications, except in the last one —column (7). In any case, as Rigobon and Stoker (2007) explain, these coefficients should to be "interpreted with care".

Table B.3: Impact of Outages on Electrification Decisions (using avg. hours of power)

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
Avg. hours with power (<24)	0.0228***	0.0142***	0.0110**	0.0168***	0.0145***	0.00909
	(0.00431)	(0.00374)	(0.00434)	(0.00554)	(0.00543)	(0.00660)
Avg. hours with power (=24)	0.413***	0.229***	0.171*	0.294**	0.247**	0.126
	(0.0933)	(0.0823)	(0.0946)	(0.121)	(0.118)	(0.142)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	13,130	13,044	10,126	4,192	4,168	2,605
R^2	0.058	0.215	0.232	0.039	0.094	0.129

Standard errors clustered at the Enumeration Area (EA) in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B.3 Alternative sample selection

As discussed above, one potential threat to identification relates to the concern that, in some particular neighborhoods, a *marginal household* connected to electricity is likely to cause additional outages by overloading the local system. In particular, this is the case in neighborhoods where informal connections are relatively common (as informal connections are typically associated with frequent outages), and in neighborhoods where households are likely to use electricity to plug-in high-power appliances (such as refrigerators). These neighborhoods are, thus, a potential source of *endogeneity* ("reverse causality") that bias our OLS coefficients. Therefore, in this appendix we estimate equation 1 after excluding from the sample the EAs in which this issue is a major concern.

Table B.4 contains the estimated coefficients after removing from our sample EAs in which at least one household is aware of the existence of informal connections. Table B.5 includes the coefficients after removing from our sample EAs in which at least one household owns a refrigerator (which is a high-power domestic appliance). In both cases, the coefficients remain negative and similar in magnitude as those obtained in Table 2. They are also statistically significant in all the model specifications if we use the full set of households; and also statistically significant if we use just the subset of "permanent" households in our sample in all the model specifications except in the one included in column (7) (i.e., using full set of controls). This is potentially due to the substantial decrease in the number of observations.

Table B.4: Impact of Outages on Electrification Decisions (drop EAs with informal connections)

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
Outages frequency	-0.0708*	-0.0846***	-0.0901***	-0.0862*	-0.0821*	-0.0929
	(0.0379)	(0.0315)	(0.0335)	(0.0473)	(0.0467)	(0.0603)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	6,612	6,572	5,426	2,119	2,107	1,439
R^2	0.066	0.239	0.257	0.024	0.093	0.117

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.5: Impact of Outages on Electrification Decisions (drop EAs with refrigerators)

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
Outages frequency	-0.0635*	-0.0591*	-0.0621*	-0.0801*	-0.0833*	-0.0844
	(0.0357)	(0.0308)	(0.0344)	(0.0486)	(0.0470)	(0.0555)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	12,040	11,959	9,326	3,598	3,576	2,236
R^2	0.052	0.197	0.215	0.035	0.088	0.122

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.6: Impact of Outages on Electrification Decisions (alternative samples)

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Keep households with solar panel/battery/mini-generator					
Outages frequency	-0.0794**	-0.0577**	-0.0585*	-0.0904**	-0.0862**	-0.0908*
	(0.0348)	(0.0294)	(0.0328)	(0.0445)	(0.0428)	(0.0508)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	13,181	13,095	10,175	4,211	4,187	2,618
R^2	0.048	0.210	0.228	0.033	0.089	0.124
	Panel B: Drop if $<$ 10% of household use light					
Outages frequency	-0.0943***	-0.0727***	-0.0788***	-0.0921**	-0.0879**	-0.0902*
	(0.0333)	(0.0279)	(0.0297)	(0.0447)	(0.0431)	(0.0509)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	13,020	12,935	10,040	4,177	4,153	2,603
R^2	0.050	0.210	0.230	0.032	0.088	0.124
	Panel C: Drop if electricity bill included in monthly rent					
Outages frequency	-0.144***	-0.0832*	-0.0797*	-0.234***	-0.213***	-0.189***
	(0.0517)	(0.0426)	(0.0447)	(0.0639)	(0.0594)	(0.0624)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	6,709	6,667	5,491	2,092	2,080	1,622
R^2	0.073	0.253	0.277	0.073	0.154	0.168
	Panel D: Drop households with prepaid contracts					
Outages frequency	-0.0728**	-0.0569*	-0.0556*	-0.0891*	-0.0869**	-0.0885*
	(0.0353)	(0.0297)	(0.0330)	(0.0454)	(0.0436)	(0.0514)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	12,937	12,855	9,992	4,104	4,081	2,552
R^2	0.047	0.209	0.228	0.031	0.088	0.124

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Finally, we check that our empirical results do not substantially change if we use alternative sample selection criteria. The estimated coefficients for the outages-frequency variable are included in Table B.6. First, in Panel A, we include in our sample the subset of households that own either a solar panel, a battery, or a mini-generator. Next, in Panel B, we use our main sample but we exclude households that live in EAs where less than 10% use electricity as the primary source of lighting. In Panel C, we present the coefficients after removing from our main sample households whose electricity bill is included in the rent that they pay to their landlords. Finally, in Panel D, we drop from our sample households that do not pay directly to the utility company either, but that rather use prepaid cards. In all these cases, the coefficient of interest remains negative, significant, and similar in magnitude to those included in Table 2.

B.4 Falsification test (full set of coefficients)

In this appendix we provide the full set of coefficients of the falsification test included in Section 3.6.2 in the main text. These coefficients are included in Table B.7.

B.5 Additional propensity score-matching estimations

Finally, in this appendix we present additional results of the rolling version of the propensity score-matching estimator —see Section 3.6.1.

First, Figure B.1 includes the results of the propensity score-matching estimator using just the subset of "permanent" households. As is the case in Figure 3, we find that for greater values of x, the impact of outages on households' electrification decisions becomes more negative and significant (due to the decrease in the number of observations, we lose some precision in these estimates).

Finally, Figure B.2 contains additional results of the propensity score-matching estimator using the alternative subsamples suggested in Appendix B.3 (Table B.6). That is, we use our main sample but (a) including the subset of households that own either a solar panel, a battery, or a mini-generator; (b) excluding households that live in EAs where less than 10% use electricity as the primary source of lighting; (c) removing households whose electricity bill is included in the rent that they pay to their landlords; and (d) dropping households that do not pay directly to the utility company either, but that rather use prepaid cards. In all cases, the result do not substantially change relative to the estimates included in Figure 3.

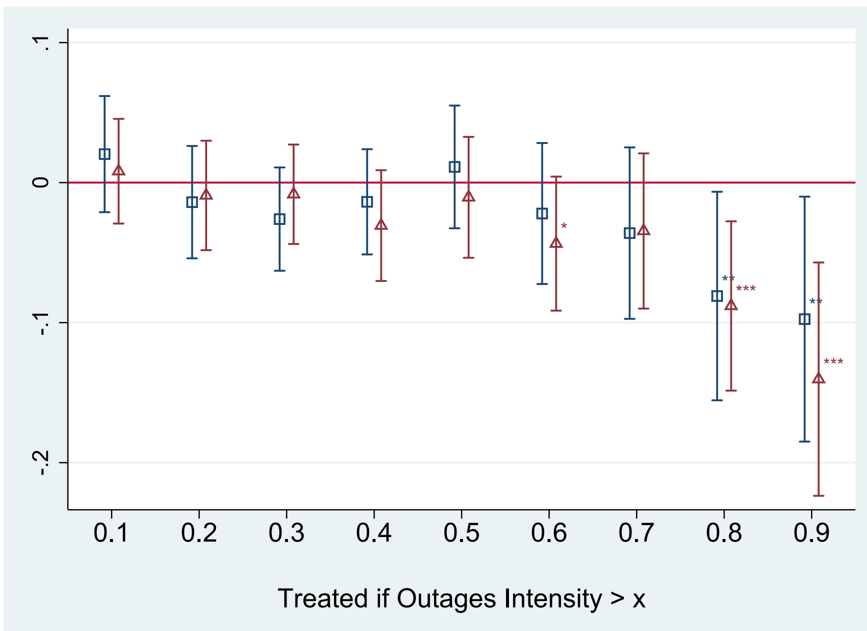
Table B.7: Impact of Outages on Electrification Decisions (falsification test; full set of coefficients)

	Full sample			*Permanent households*		
	(1)	(2)	(3)	(4)	(5)	(6)
Outages frequency	0.0255	0.0224	0.00562	0.0259	0.0272	0.00669
	(0.0320)	(0.0314)	(0.0303)	(0.0472)	(0.0467)	(0.0529)
(log) Month expenditure		0.0351***	0.0369***		-0.00856	0.00000821
		(0.00769)	(0.00796)		(0.0113)	(0.0128)
Head of hh primary educ		0.103***	0.0870***		0.0908***	0.0935**
		(0.0156)	(0.0177)		(0.0302)	(0.0450)
Robust/permanent wall		0.145***	0.102***		0.0778***	0.102***
		(0.0186)	(0.0167)		(0.0265)	(0.0340)
Hh size		-0.0160***	-0.0156***		-0.0104**	-0.0134*
		(0.00309)	(0.00340)		(0.00502)	(0.00687)
(log) Months in dwelling		-0.000295	-0.00470		-0.000812	-0.0143
		(0.00394)	(0.00460)		(0.00611)	(0.00933)
Distance grid (no Garissa)			-0.00488			-0.00942
			(0.00358)			(0.00724)
Informal connections			-0.0584**			0.0433
			(0.0252)			(0.0427)
Street lights			0.0289			0.0556*
			(0.0218)			(0.0289)
Distance plant (no Garissa)			0.00139			0.00715
			(0.00219)			(0.00586)
Slum			-0.137***			-0.103**
			(0.0254)			(0.0417)
Urban			-0.0179			-0.0754
			(0.0283)			(0.0645)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	11,276	11,202	8,604	3,618	3,596	2,102
R^2	0.172	0.201	0.182	0.202	0.211	0.133

Standard errors clustered at the Enumeration Area (EA) in parentheses

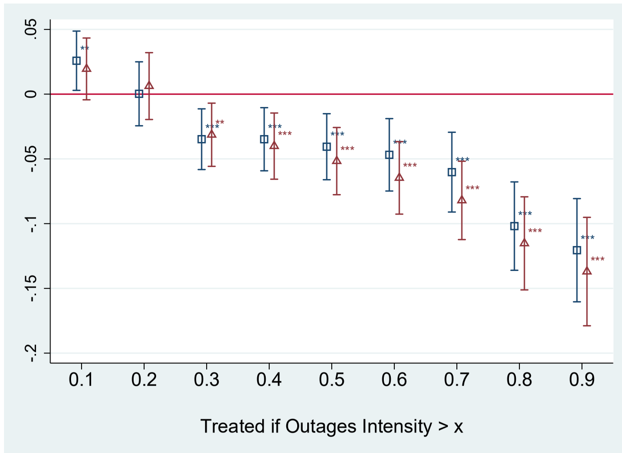
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure B.1: Propensity-score matching estimator for different “treated” samples (“permanent” households only)

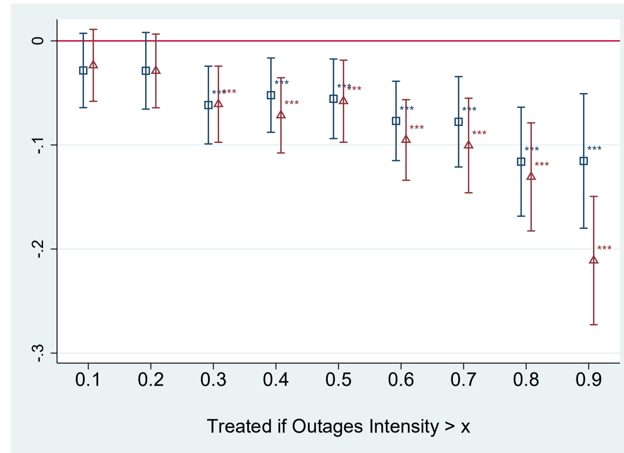


Note: The figure shows the propensity-score matching coefficients for the outages-frequency variable ($\theta_{j,c,p}$) after matching *treated* and *control* households. From left to right, the households that live in EAs for which the outages-frequency variable is greater than $x \in \{0.1, \cdots, 0.9\}$ are considered as treated; and the control households are those that live in EAs for which the outages-frequency variable is equal to zero. Households are matched (using propensity-score matching) on *(i)* income, quality of the dwelling unit (proxied by “robust wall” dummy), head of household education, urban dummy and slum dummy (coefficients are displayed with a square); and on *(ii)* the full set of household-level controls included in our main regressions (coefficients are displayed with a triangle). Vertical bands represent ± 1.96 times the standard error for each point estimate, and the significance levels of the estimated coefficients are as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. These results are obtained using the sub-set of “permanent” households.

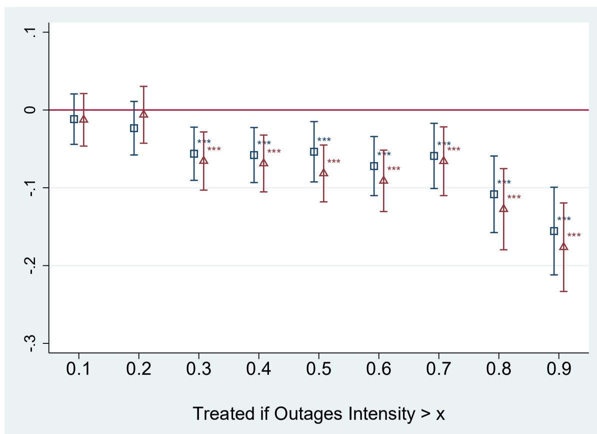
Figure B.2: Propensity score matching estimator for different “treated” samples (alternative samples)



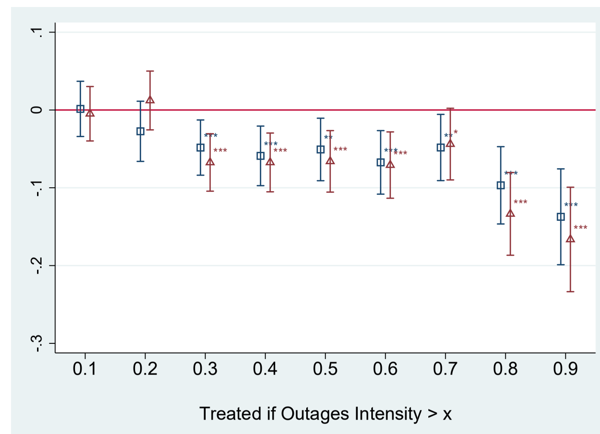
(a) Full sample, keep if hh has solar panel or similar



(b) Full sample, drop if < 10% hh use light



(c) Full sample, drop if bill included in rent



(d) Full sample, drop hh with prepaid contract

Note: The figures show the propensity-score matching coefficients for the outages-frequency variable ($\theta_{j,c,p}$) after matching *treated* and *control* households. In each subfigure, from left to right, the households that live in EAs for which the outages-frequency variable is greater than $x \in \{0.1, \cdots, 0.9\}$ are considered as treated; and the control households are those that live in EAs for which the outages-frequency variable is equal to zero. Households are matched (using propensity-score matching) on *(i)* income, quality of the dwelling unit (proxied by “robust wall” dummy), head of household education, urban dummy and slum dummy (coefficients are displayed with a square); and on *(ii)* the full set of household-level controls included in our main regressions (coefficients are displayed with a triangle). Vertical bands represent ± 1.96 times the standard error for each point estimate, and the significance levels of the estimated coefficients are as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. These results are obtained using different subsamples of the full set of households. Subfigure B.2a displays the results using all the households in the main sample but including also households that own either a solar panel, a battery, or a mini-generator; in Subfigure B.2b we exclude from our main sample households that live in EAs where less than 10% of them use electricity as the primary source of lighting; in Subfigure B.2c we remove from our main sample households whose electricity bill is included in the rent that they pay to their landlords; finally, in Subfigure B.2d we drop from our main sample households that do not pay directly to the utility company either, but that rather use prepaid cards.

Appendix C. Effect of Outages on Post-electrification Decisions: Robustness Checks

Although we consider that our empirical strategy is strong enough to support the causal link between the frequency of outages and the probability with which a household owns certain large electric-powered appliances, we recognize there are different alternative definitions of the outcome variable and subsamples that we could have been employed. Hence, in this appendix we engage in an extensive procedure of robustness checks, in order to verify that the results are not sensitive to the choices used throughout the paper.

C.1 Alternative definitions of frequency of outages

In Section 3.3, we build the frequency of outages for each EA ($\theta_{j,c,p}$) using data from households that use electricity as their primary source of lighting. We now provide evidence that our main empirical results are similar to the empirical results in the main text if we use the alternative definitions of the outages-frequency variable that we use in appendix B.2.

First, we use the alternative definition of outages-frequency included in equation B.1 (i.e., using the full subset of households that have electricity at home). The estimated coefficients for the appliances regressions are included in Table C.1. These coefficients are extremely similar in magnitude to those included in Table 5.

Second, we use households' reported data on the average number of hours per day that they get electricity in their dwelling units —see equation B.2. As done in Appendix B.2, we use two variables: the first one is a dummy variable that is equal to 1 if $\theta_{j,c,p}$ is equal to 24, and 0 otherwise. The second one is a continuous variable that is equal to the actual value of $\theta_{j,c,p}$ if it is not equal to 24, and 0 otherwise. The coefficients of these two variables are expected to be positive and significant. That is, for the former variable (the dummy one), EAs that receive electricity 24 hours per day (on average) are expected to increase the probability with which a household that lives in such an EA owns a refrigerator/television. For the latter variable (the continuous one), having an additional hour of electricity per day at the EA level is also expected to increase a household's probability of purchasing these appliances.

Table C.1: Impact of Outages on Appliances Ownership (alternative outages-frequency)

	Refrigerator			*Television*		
	(1) OLS	(2) OLS	(3) IV	(4) OLS	(5) OLS	(6) IV
Outages frequency (alt.)	-0.109***	-0.0697***	-0.337***	-0.0901***	-0.0727***	-0.525**
	(0.0217)	(0.0188)	(0.126)	(0.0257)	(0.0254)	(0.230)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	8,883	8,882	8,882	8,883	8,878	8,878
R^2	0.163	0.246	0.181	0.149	0.153	0.083
First stage IV estimates: Dependent variables is $\theta_{j,c,p}$						
Informal connections			0.263***			0.270***
			(0.0625)			(0.0626)

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.2: Impact of Outages on Appliances Ownership (using avg. hours of power)

	Refrigerator		*Television*	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Avg. hours with power (<24)	0.00502*	0.00298	0.0128***	0.0107***
	(0.00270)	(0.00276)	(0.00347)	(0.00352)
Avg. hours with power (=24)	0.150**	0.105*	0.318***	0.266***
	(0.0583)	(0.0591)	(0.0748)	(0.0760)
City dummies	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓
Observations	8,883	8,882	8,883	8,878
R^2	0.161	0.167	0.151	0.154

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Consistently, in Table C.2 we see that both coefficients are positive and significant.[45] These coefficients suggest both that households that receive power 24/7, and that households that receive an *extra* hour of electricity are more likely to own these appliances.

C.2 Alternative sample selection

Next, we check that the coefficients obtained for the appliances regressions are also robust to the alternative sample selection criteria that we use in Appendix B. The estimated coefficients are included in Table C.3, in which we consider the following inclusion/exclusion of households. First, in Panel A, we include in our main sample the subset of households that own either a solar panel, a battery or a mini-generator. In Panel B, we exclude from our sample households that live in EAs where less than 10% use electricity as the main source of lighting. In Panel C, we present the coefficients after removing households whose electricity bill is included in the monthly rent. Finally, in Panel D, we exclude households that use prepaid cards. In all these cases, the coefficient of interest remains negative, significant, and similar in magnitude to those included in Table 5.

C.3 Falsification test

Finally, we perform placebo-based falsification tests for the appliances regressions. We estimate equation 3 again but we use two *fake outcome* variables instead that are "similar" to our actual outcome variables (a dummy indicating whether a household owns a refrigerator/television) but that are known to be unaffected by outages. In particular, we use a dummy variable that indicates whether a household has/does not have a sewing machine as our fake outcome for the refrigerator regressions, and a dummy variable that indicates whether a household has/does not have a radio as our fake outcome for the television regressions. Bearing in mind that sewing machines are manually powered, and that radios typically use portable batteries, if there are no $\theta_{j,c,p}$-correlated unobserved variables that might potentially explain whether a household has a refrigerator/television, then we should expect the coefficients of $\theta_{j,c,p}$ to be close to zero and not significant.

The results of these falsification tests are included in Table C.4. Consistently, we find that the estimated coefficients for $\theta_{j,c,p}$ are close to zero and not significant.

[45] The IV regressions were not included in Table C.2 because we would need an additional instrument to estimate them.

Table C.3: Impact of Outages on Appliances Ownership (alternative samples)

	Refrigerator			*Television*		
	(1) OLS	(2) OLS	(3) IV	(4) OLS	(5) OLS	(6) IV
	Panel A: Keep households with solar panel/battery/mini-generator					
Outages frequency	-0.114***	-0.0726***	-0.381***	-0.0939***	-0.0767***	-0.508**
	(0.0221)	(0.0191)	(0.130)	(0.0253)	(0.0249)	(0.224)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	8,941	8,940	8,940	8,941	8,936	8,936
R^2	0.175	0.258	0.181	0.151	0.155	0.088
	Panel B: Drop if < 10% of household use light					
Outages frequency	-0.109***	-0.0693***	-0.334***	-0.0959***	-0.0784***	-0.532**
	(0.0216)	(0.0187)	(0.125)	(0.0255)	(0.0252)	(0.230)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	8,873	8,872	8,872	8,873	8,868	8,868
R^2	0.164	0.247	0.182	0.150	0.154	0.082
	Panel C: Drop if electricity bill included in monthly rent					
Outages frequency	-0.166***	-0.110***	-0.248**	-0.0908**	-0.0604	-0.533**
	(0.0317)	(0.0282)	(0.110)	(0.0376)	(0.0377)	(0.233)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	3,980	3,980	3,980	3,980	3,978	3,978
R^2	0.177	0.288	0.255	0.137	0.145	0.077
	Panel D: Drop households with prepaid contracts					
Outages frequency	-0.0995***	-0.0621***	-0.292**	-0.0991***	-0.0821***	-0.531**
	(0.0213)	(0.0185)	(0.120)	(0.0255)	(0.0252)	(0.238)
City dummies	✓	✓	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓	✓	✓
Observations	8,716	8,715	8,715	8,716	8,711	8,711
R^2	0.157	0.240	0.185	0.150	0.153	0.084

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.4: Impact of Outages on Appliances Ownership (falsification test)

	Sewing machine		*Radio*	
	(1)	(2)	(3)	(4)
	OLS	OLS	OLS	OLS
Outages frequency	-0.00613	-0.00535	-0.0183	-0.0149
	(0.00647)	(0.00654)	(0.0223)	(0.0227)
(log) Month expenditure	0.00795***	0.00678***	0.0404***	0.0394***
	(0.00177)	(0.00177)	(0.00533)	(0.00531)
Head of hh primary educ	0.0101**	0.0102**	0.0412**	0.0392**
	(0.00441)	(0.00441)	(0.0178)	(0.0178)
Robust/permanent wall	0.00916**	0.00859**	-0.00108	-0.00331
	(0.00419)	(0.00436)	(0.0122)	(0.0121)
Hh size	0.00332***	0.00317***	-0.00211	-0.00195
	(0.00112)	(0.00112)	(0.00274)	(0.00275)
(log) Months in dwelling	0.00351***	0.00310**	0.0339***	0.0336***
	(0.00135)	(0.00142)	(0.00355)	(0.00357)
Kitchen in dwelling		0.00883**		
		(0.00391)		
Food shop nearby		0.00714		
		(0.00769)		
Slum		0.00145		-0.0134
		(0.00393)		(0.0122)
Urban		-0.00101		-0.00641
		(0.00561)		(0.0157)
Park nearby				0.0116
				(0.0139)
City dummies	✓	✓	✓	✓
Province dummies	✓	✓	✓	✓
Observations	8,877	8,876	8,877	8,872
R^2	0.020	0.020	0.056	0.056

Standard errors clustered at the Enumeration Area (EA) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$