Electricity Consumption: The role of grid reliability in appliance ownership in Rwanda

Joel Mugyenyi
jm5352@columbia.edu

Gabriel Gonzalez Sutil
gg2718@columbia.edu

Vijay Modi
modi@columbia.edu

January 01, 2024

Abstract

This study delves into household-level electricity consumption in Rwanda, examining the intricate relationship between electricity reliability and appliance ownership. It focuses on both the total number and specific types of appliances owned by households. Leveraging a unique dataset and instrumental variables, we explore how households adapt to low grid quality by adjusting their appliance stock. The results highlight the significant influence of reliability on the likelihood of owning specific appliances, including smart-phones, TVs, and decoders. Higher-income households tend to prefer alternatives like music systems, while lower-income households gravitate toward appliances less reliant on the grid, such as sewing machines. Furthermore, the study reveals that reliability has a nuanced impact on ownership, shaping the types of appliances owned rather than the overall count. Additionally, our observations indicate that reliability does not affect consumption for households already possessing appliances. The study also investigates the impact of other income and non-income household characteristics on appliance ownership and usage. These findings contribute valuable insights to the literature on households' responses to electricity reliability improvements, offering guidance for targeted policies aimed at increasing appliance ownership and, consequently, enhancing residential electricity consumption.

Keywords: *Rwanda; Electricity reliability ; Appliance ownership; Energy demand*

1 Introduction

Access to electricity can significantly enhance the well-being of individuals only when electric appliances are not only acquired but also actively utilized within households. (Lenz, Munyehirwe, Peters, & Sievert, 2017; Richmond & Urpelainen, 2019). However, the unfortunate reality is that households in Sub-Saharan Africa own few and a limited variety of appliances, even years after electrification efforts (Adesina et al., 2020; Lenz et al., 2017). Moreover, deficiencies in appliance uptake contribute to low residential electricity consumption (Auffhammer & Wolfram, 2014; Dubin & McFadden, 1984; Nielsen, 1993), affecting the financial sustainability of distribution utilities[1] (Blimpo & Cosgrove-Davies, 2019). Achieving sustained progress in electricity access necessitates not only expanding

[1]This reality, as a consequence, diminishes the incentives to connect households and invest in grid expansion. In Rwanda, the main principle adopted for financing transmission lines was an "80-10-10" shared financing policy. Under this policy, 80% of the capital requirements would be sourced from the government and the development partners; 10% from the utility's retained earnings, and 10% from customer connection charges.

access but also increasing the number and diversity of appliances owned and actively used by households.

While low income is a key variable explaining low appliance ownership, economic theory suggests that an unreliable electricity supply reduces the incentives from acquiring new appliances (Hashemi, 2022; McRae, 2010; Meeks, Omuraliev, Isaev, & Wang, 2023). The current state of the electricity sector in most African countries is characterized by pervasive reliability challenges (Blimpo & Cosgrove-Davies, 2019; Day, 2020; IEA, 2022). Consequently, households are likely to respond substantially to an unreliable service by refraining from purchasing certain appliances, anticipating that frequent outages will hinder their regular use.[2]. However, previous literature has found that consumers might be boundedly rational[3]; in an extreme, consumers might not be forward-looking agents solving a dynamic programming problem as theory predicts (Himarios, 2000). Moreover, adoption patterns are less straightforward in poorer and rural contexts. For instance, appliance ownership might be motivated by social status (Ramakrishnan, Kalkuhl, Ahmad, & Creutzig, 2020), and low reliability might decrease ownership indirectly through lower household incomes, rather than having a direct effect (Dang & La, 2019). However, little empirical evidence which refutes or corroborates these theories exists. The central question remains: How does reliability affect appliance ownership in Sub-Saharan Africa, and consequently, how can investments in reliability influence electricity consumption?

In order to contribute to the literature we assess how the state of electricity reliability impacts ownership of a wide range of appliances in Rwanda using a unique data set and state-of-the-art instrumental variables for reliability. According to the Rwanda electricity distribution plan (REG, 2021), the distribution network suffers from poor reliability and quality of supply which is attributed to under-investment. We follow previous literature and evaluate two outcome variables: a count of electric appliances (size) and the ownership of specific appliances (composition) (see Richmond and Urpelainen (2019) and Matsumoto (2016a)). We use conditional fixed-effects Poisson models and linear fixed-effects probability models to investigate the household appliance stock. Appliance data at the household-level was obtained from the Integrated Household Living Conditions Surveys (EICV). Our key explanatory variable is grid reliability which we measure with the frequency of outages per day. We use administrative reliability data from the Rwanda Energy Group (REG) which we link with house locations using non-public GPS location of interviewed household, accessible via an agreement with the National Institute of Statistics of Rwanda (NISR). Additionally, we instrument our reliability measures with lightning activity, specifically, lightning radiance and strikes frequency. Our models also include household control variables and fixed effects in an additive approach.

Our results demonstrate that households are forward-looking and adapt to low reliability levels in Rwanda. The frequency of outages per day is associated with a decreased probability of households owning certain appliances, such as smart-phones, TVs, and decoders. Notably, higher-income households demonstrate an increased likelihood of owning appliances like music systems, while lower-income households have a higher probability of owning items such as sewing machines. These appliances rely less on electricity from the grid. Interestingly, the frequency of outages does not significantly impact the total number of appliances owned by the household. This underscores that the influence of reliability on the ownership of key appliances primarily affects the type of appliance

[2] An outage is a complete stoppage within the distribution system, preventing end users' consumption of electricity services. Planned outages are either for regular repairs and maintenance, which are typically of limited duration and scheduled for off-peak months. Unplanned outages are typically due to infrastructure breakage, malfunction, and overloaded distribution systems.

[3] For some reason, consumers are not able to account for the future or form expectations with available information (Himarios, 2000)

rather than the overall quantity. Additionally, we highlight that reliability does not influence appliance usage for households that already own appliances. Therefore, improved grid reliability has limited impact in augmenting household-level consumption for those already possessing appliances.

We contribute to a small but growing literature on households' response to electricity reliability improvements. Despite increased electricity access in the 21st century, many developing countries still face challenges in ensuring satisfactory service quality (Blimpo & Cosgrove-Davies, 2019; Burgess, Greenstone, Ryan, & Sudarshan, 2020; Meeks et al., 2023). In this sense, understanding residential consumers' responses to experiencing changes in electricity quality has attracted attention by researchers. Meeks et al. (2023) explores appliance ownership and reliability in Nepal, Hashemi (2022) investigates the same in Kyrgyz Republic, and McRae (2010) in Colombia. These studies indicate that households in middle-income countries significantly respond to unreliable services by refraining from purchasing certain appliances. Our study extends this inquiry to low-income countries, specifically analyzing appliance ownership in Rwanda. The adoption patterns in low-income rural settings are nuanced. Additionally, our study sets itself apart by leveraging novel administrative data and instrumental variables to enhance identification.

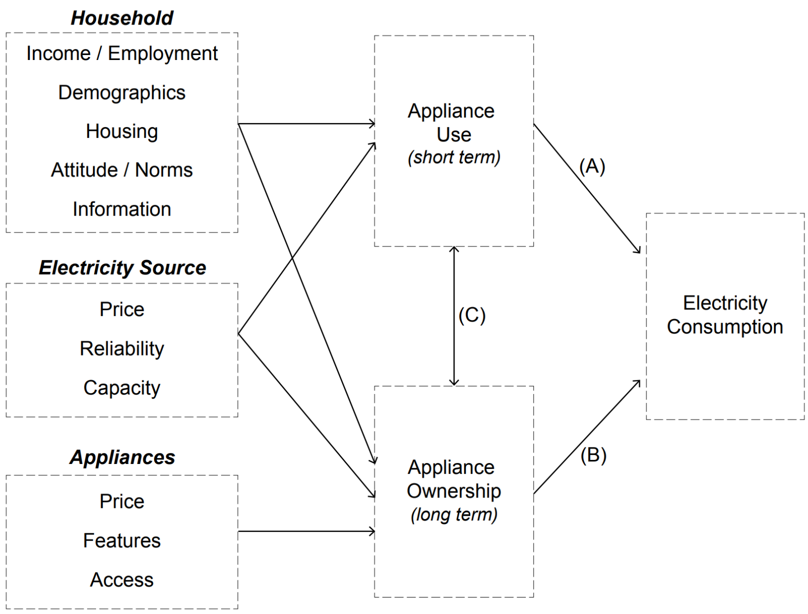
The existing literature has predominantly examined the correlation between household income and the adoption of specific appliances, anticipating their role in driving household electricity demand growth (see Auffhammer and Wolfram (2014) and Gertler, Shelef, Wolfram, and Fuchs (2016)). Nonetheless, considerable variability in appliance ownership persists across income levels, and non-income factors play a crucial role (Debnath, Bardhan, & Sunikka-Blank, 2019; Rao & Ummel, 2017). Our findings not only shed light on the influence of reliability on appliance ownership but also present descriptive evidence highlighting the impact of non-income drivers, such as gender, household composition, education, and dwelling characteristics.

The subsequent sections of the paper are organized as follows. In the next section, we delve into the existing literature. Section 3 provides an overview of the data employed in our analysis of appliance ownership. In section 4, we outline our methodology and present the ensuing results. Section 5 delves into the examination of the impact of reliability and other factors on appliance-level electricity consumption, elucidating key implications for policy. Finally, in section 6, we present our conclusions.

2 Appliances and Residential Electricity Consumption

The conceptual framework of this paper, summarized in Figure 1, follows Dubin and McFadden (1984) and Nielsen (1993) in which residential electricity consumption is determined by ownership of electrical appliances and the intensity of use of these devices. In other words, electricity does not enter directly in households' utility function, but it is the input for electric appliances to produce energy services (Atkeson & Kehoe, 1999; Sievert & Steinbuks, 2020). Households own appliances and utilize these appliance at an intensity level that provides the "necessary" service (Nielsen, 1993). This results in a electricity consumption represented by arrow (A) in Figure 1. Moreover, appliances continue to use energy and drain power even it is not used, an effect known as energy vampire consumption. We represent such indirect consumption with arrow (B) in the bottom-right of Figure 1. Consequently, residential electricity consumption is conceptualized as a direct implication of appliance ownership and usage. Thus, understanding the patterns of appliance acquisition and usage is important for both policymakers and utilities aiming to increase electricity consumption.

Figure 1: Determinants of household's electricity consumption



This paper focus mainly on households' long-term decision, in particular, the demand for appliances. Appliance usage is conditional on ownership: households cannot use appliances if they do not own, or have access to them -arrow (c) in Figure 1-. Indeed, household first acquire new appliances; afterwards, households use these appliance (short-term choice) in order to achieve the "necessary" service level demanded from an appliance Nielsen (1993). In this context, theory and empirical studies, discussed below, suggest that households appliance demand can depend on socioeconomic and several other factors.

2.1 Demand for electric appliances

This section explores the literature studying households' demand for appliances. Economic theory of the demand for durable goods suggests that such demand arises from the flow of services provided by durables ownership where utility is best characterized as indirect (Dubin & McFadden, 1984). This utility is assumed to depend on several factors including household characteristics, appliance factors, among others. The optimization problem posed is thus quite complex: the household, in the spirit of the theory, must weigh the alternatives of each appliance against expectations of future use, future energy prices, among others. In this sense, appliance ownership can be determined by a wide range of factors in a complex interrelationship, and several variables factor into households decision-making process for purchasing appliances (Lenz et al., 2017).

One such factor is household income. Early work (see Farrell (1954)) assumed an S-shaped relationship between income and the share of households who own appliances in a model based on a log-normal distribution of "acquisition thresholds". A household must save to acquire the appliance, and this delays the appliance acquisition to a higher income. Past a certain income threshold, households become much more likely to acquire appliances[4]. Gertler et al. (2016) show that a rise in income has a linear relationship with ownership of low-level (i.e. low cost) appliances, and the non-linear relationship is only for major appliances as for example refrigerators (i.e high costs). This suggests that

[4]Low-income households do not allocate additional income to acquire energy-using assets

households wait until they have enough income to purchase high cost and long lifespan appliances which are not replaced frequently, but not low cost appliances. This intuition is similar to the one introduced by Khandker, Barnes, and Samad (2009), where households are more likely to adopt simple electric lighting appliances initially and invest in more energy-intensive appliances over time as households are able to save up for appliance purchases.

While most studies examine the impact of total household income on appliance ownership, Matsumoto (2016a) analyses households' income structure and its effect on appliance ownership. The author finds that in double-income households, non-labor income, such as pension income, lowers the likelihood of dishwasher ownership; yet, labor income raises it. Moreover, wives' income, and non-labor income, increase the number of televisions while a husband's income decreases it. Finally, (Wolfram, Shelef, & Gertler, 2012) explains that households at very low levels of income are less able to self-finance the appliance, and credit constraints is hence important barrier to appliance adoption at lower income levels.

Income is a key predictor of appliance take-up, but non-income drivers might matter as well. Recent evidence from other studies suggests that appliance diffusion can remain low despite rising incomes (Debnath et al., 2019), and that non-income drivers can be important determinants of appliance choice (see Richmond and Urpelainen (2019) for a review). In this context, understanding these drivers can be helpful to identify barriers to appliance uptake and residential energy demand growth within countries. Yet, empirical evidence of these drivers are limited, and since adoption patterns can be less straightforward among relatively poorer rural households, empirical evidence on appliance ownership is important to guide policy-makers and utilities, specially, in Sub-Saharan Africa..

First, ownership of key appliances might be limited due to housing conditions. Matsumoto (2016a) finds that households owning a detached house have more appliances, excluding PCs and cellular phones. In addition, the authors find that home ownership has a positive impact on the ownership of appliances in Japan. In a similar way, O'Doherty, Lyons, and Tol (2008) found that homeowners are more likely to have more appliances. In certain situation, adoption of appliances might depend on the context and exogenous variables to the household (McNeil & Letschert, 2010). For example, the usefulness of an fan or air conditioner is climate dependent. That is, adoption might depend on climate and geographic variables. Yet, dwelling size and structure has not been extensively studied.

Second, household willingness to adopt appliances is not necessarily straightforward if they wrongly perceive the benefit of using the appliance and have limited information about them. Bos, Chaplin, and Mamun (2018) reviews several electrification programs over different period of times and found that sometimes it takes several years for household to internalize the value of appliance usage. However, households can be persuaded by those individuals already using appliances (Hanna & Oliva, 2015). This effect is know as the "demonstration effect" (Bos et al., 2018). In this context, education might affect ownership of key appliances as more educated household might have more information. Dhanaraj, Mahambare, and Munjal (2018a) recently found that refrigerator ownership was higher among more educated households.

Finally, the evidence of how gender and households' demographics affect appliance ownership is scant on the literature. Rao and Ummel (2017) found that race and religion together, among other household characteristics, help explain the heterogeneity in appliance ownership at lower income levels in Brazil and South Africa. However, religion was not found significant at all by Richmond and Urpelainen (2019) who studies appliance uptake in India. In addition, the author finds that gender of the decision maker is a significant factor affecting appliance choice.

In this context, theory suggests that a principle barrier limiting ownership of key durable goods might an unreliable electricity service (McRae, 2010). The demand for appliances is expected to arise from the flow of services provided by ownership of the appliances (Dubin & McFadden, 1984). Hence, a low quality electricity service limits their usefulness, and reduces the demand for them (McRae, 2010). Previous evidence for middle income countries corroborate this claim (see Meeks et al. (2023), Hashemi (2022), and McRae (2010)). However, such patterns might not be the same for Sub-Saharan African countries. First, low service quality could decrease the ownership of electricity appliances only through lower household's incomes as affordability is a main constraint (Dang & La, 2019). Moreover, previous literature has found that there is a large share of rule-of-thumb, myopic, and bounded rational consumers in general (Himarios, 2000), and hence, one can expect that adoption patterns are not the same in low-income countries. Indeed, Ramakrishnan et al. (2020) shows that end-users base their consumption decisions not only on available budget and direct use value, but also on their social environment. They find that while income and household demographics are predominant drivers of appliance take-up, household's perception of status emerges as a key social dimension influencing the take-up in India. The main goal of this paper is to provide empirical evidence of the role of reliability on households' appliance demand in low-income settings. We do this by studying appliance ownership in Rwanda leveraging a combination of household level survey data and electricity utility data.

3 Data Description

Estimating the relationship between electricity service quality and household outcomes is typically challenging. Measuring electricity relability is difficult due to common data limitations: utilities may not record outages and, if they do, they may lack incentives to share such data (Meeks et al., 2023). As a result, most prior economics research on electricity quality has either employed data on self-reported electricity quality, which is prone to misreporting, or used electricity shortages as a proxy for outages[5] (Meeks et al., 2023). We overcome these data challenge through a novel dataset which combines public data with proprietary data on electricity outages at the feeder level obtained directly from the Rwanda Energy Group (REG) [6].

The proprietary reliability dataset comprises information on electricity outages for feeder lines from 2016 to 2020[7], encompassing details such as outage duration, date of occurrence, underlying cause of the outage, and the substation and feeder line implicated. The electricity network consists of many long radial feeder lines[8] in Rwanda; in extreme cases, these lines are longer than 300 km. Faults on such feeders would result in wide-spread outages affecting many households. Complementing this information, the Ministry of Infrastructure collects data for a collection of 52,418 low voltage electricity conductor lines. Each of these lines is characterized by attributes including its parent substation and feeder line, as well as its inherent nature (whether it is underground or overhead), length, and voltage

[5] A power outage entails the temporary disruption of electricity supply, affecting either a portion or the entirety of a power grid. Various factors, such as an imbalance between demand and supply, can contribute to the occurrence of a power outage. In contrast, a power shortage occurs when the existing electricity supply falls short of meeting the overall demand.

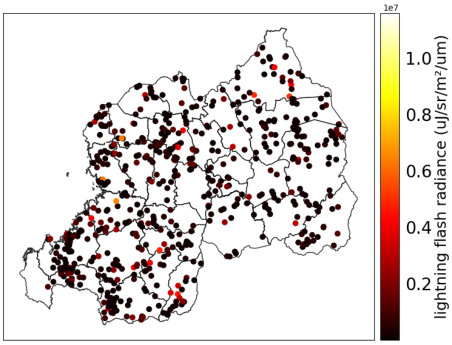
[6] Rwanda Energy Group Limited (REG), is a government-owned holding company responsible for the import, export, procurement, generation, transmission, distribution and sale of electricity in Rwanda

[7] The number of unique feeders in each year are as follows: 53 in 2016, 63 in 2017, 73 in 2018, 92 in 2019 and 77 in 2020. The reason for these changes is that grids are updated every year meaning that new feeders are getting added, or long feeders are being divided.

[8] An electricity feeder line is a power line that carriers electricity from a substation to individual customers or smaller substations.

capacity. By leveraging the corresponding substation and feeder line designations, we link outages at the feeder lever with low voltage conductor lines which we use to estimate outages for each individual low voltage line. The geographical distribution of these lines is visually depicted in Figure 2.

Figure 2: Distribution Feeder lines



Note: In the plot, the lines sharing the same feeder affiliation are being represented by a the same color and a bounding box is drawn around them to show the area of coverage of each feeder. The areas of coverage of feeders further from the capital Kigali, which is at the center of the country, tend to be larger than those closer to Kigali. The smallest feeder has an area of 1.25km^2 while the largest feeder covers an area of 3320km^2.

In the last step, we assign the grid reliability statistics to each household. For this, a collaborative effort was initiated in conjunction with the National Institute of Statistics of Rwanda (NISR)[9]. The primary objective of this collaborative endeavor was to establish a closest-distance association between the geo- graphical locations of households and the network of low voltage lines. To realize this objective, a meticulous process was undertaken wherein household GPS coordinates were matched to the nearest low voltage distribution line, within a proximity range of 800 meters. This specific range was determined in alignment with the prevailing connection policy enforced by the utility company, which delineates that low voltage connections are confined to within an 800-meter radius of the nearest distribution transformer (REG, 2020). In case the GPS locations of a low voltage line is missing, we use the closest medium voltage line that would supply a low voltage line connection to a household. Subsequent to this proximity-based matching process, each individual household was duly assigned the pertinent outage statistics attributed to the matched feeder line.

We combine this data with public information on household characteristics and appliance ownership. The Integrated Household Living Condition Survey (EICV) is a nationwide cross section survey that was initiated in 2000, and it is conducted every five years. We use the most recent survey which was completed in 2016/2017. The data was gathered using questionnaires conducted over a period 12-month cycle from October 2016 to October 2017. At the national level there were 1,260 sample villages and 14,580 sample households[10]. In the urban strata there are 245 sample villages and 2,526 sample households, and in the rural strata there are 1,015 sample villages and 12,054 sample households.

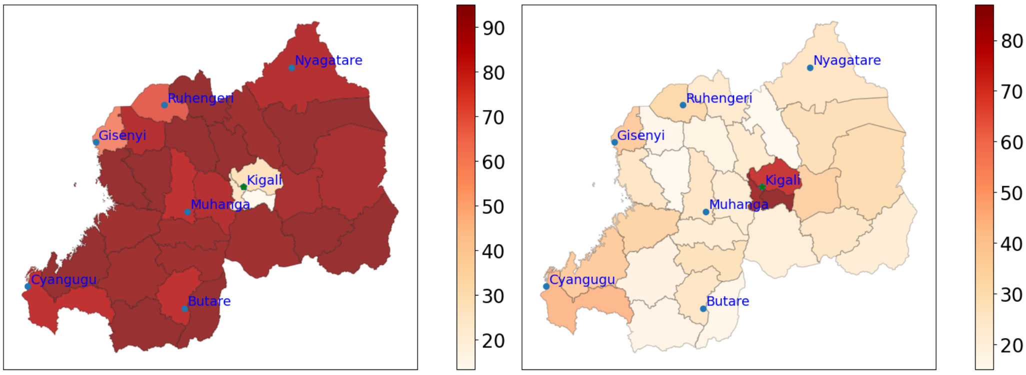
[9]It is noteworthy that the National Institute of Statistics of Rwanda (NISR) does not disclose household GPS coordinates in the public domain. Therefore, a dedicated visit to their headquarters in Kigali, Rwanda was essential to securely execute this alignment procedure within their internal systems

[10]The 2023 census has found more than 3 million household in Rwanda. The collection of the data was divided into 10 cycles in order to represent seasonality in the income and consumption data. The survey is filled by enumerators, and it is conducted all over the country to representative households. Households are chosen based on socioeconomic characteristics to obtain high correlation between households within a cluster. In Kigali Province, only 9 households are interviewed in each cluster, and for the rest of the clusters

540 households are sampled from the 3 districts that make up the capital city Kigali. From the remaining 27 districts in the country, 480 households are sampled.

The EICV inquiries surveyed households about the following appliances: radio, mobile phones, televisions, satellite dishes, decoder, music systems, computers, printers, laundry machines, electric fans, fridges, hotplates and cookers. We study appliance ownership for grid electrified households. Electricity access is defined as household which has a grid connection or access to another technology, including solar, batteries, etc. In the survey, there are 9,775 household which do not have access to electricity and 4,799 household with access to electricity (33% of surveyed households). From these, 3,600 have a connection to the electricity grid. Over 70% of sampled households in Kigali report having access to an electricity source, but the electricity access rate is much lower in the rest of the country as shown in Figure 3.

Figure 3: Spatial distribution of survey



(a) Share rural households

(b) Share electrified households

Note: Figure (a) presents the concentration of rural households by district and figure (b) shows the electrification rate of the sampled households by district. The dots show major cities in the country, the capital city Kigali is at the center of the country. Households sampled from the three districts that make up the capital Kigali are predominately urban households (Kicukiro, Gasabo and Nyarugenge). Rubavu district which hosts Gisenyi, a major commercial hub bordering the Democratic Republic of Congo has the next highest concentration of urban households outside of the capital Kigali. The rest of the districts in the country are predominately rural.

Using this data we constructed several appliance ownership measures following Richmond and Urpelainen (2019) and Matsumoto (2016a). First, we calculated the total number of appliances owned by each household. Second, appliances were categorized according to the service provided (i.e household usage), capital cost, and wattage level[11]. The categorizations are shown in Table 1. We calculated the number of appliances per tier for each household. Note that these categories are also associated with electricity consumption. Appliances in category 4 are expected to consume more electricity due to higher wattage requirements. Finally, we generated two distinct variables to capture aspects of household appliance ownership: a binary variable indicating whether a household possesses a particular appliance and a numeric variable representing the quantity of each type of appliance owned by the household. The utilization of the binary variable is motivated by the understanding that ownership of at least one unit of each appliance type reflects the household's

12 household.

[11]Wattage is a measurement of energy over a period of time and it indicates how much electrical energy they require to run.

inclination to invest in that specific category. Concurrently, the second variable quantifies the extent of ownership for each appliance within the household.

Table 1: Appliance Categories

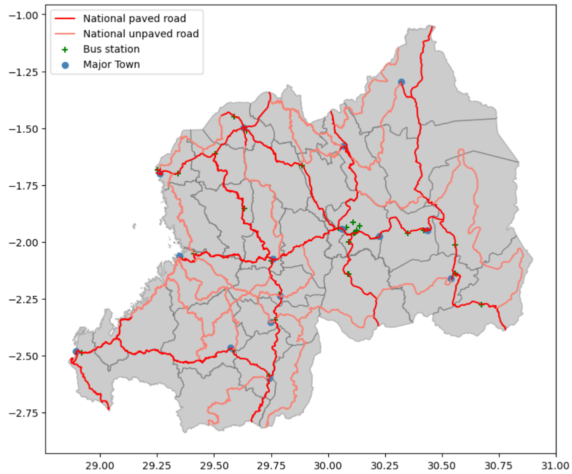
Category	Appliance(s)		Cost	Use	Wattage
0	No appliances		-	-	-
1	Analogue phones Radio	Smart phones	Low	Communication	Low
2	TVs Satellite dish Music system	Decoders DVD player Camera	Medium	Entertainment	Low
3	Computer Sewing Machine	Printer	Medium	Productivity	Medium
4	Fridge Hotplate Fan	Laundry machine Cooker Water Filter	High	Convenience	High

The EICV survey, finally, contains information on households well-being and expenditure including living conditions, education, health, housing conditions, consumption, deposits and loans, among others. We use this information to construct income and non-income drivers at the household level including household composition (number of females, children, seniors, and age and nationality of the head of household), house ownership (type of house, years in the dwelling and whether the household is the owner or not of the house), as well as several income variables including expenditure, savings, and job stability.

Simultaneously, we use information on house locations, in particular, access to markets and major cities. The key issue in accessibility measurement is the definition of the cost distance which employs the geographic principle of "*friction of distance*", which posits that there is a cost associated with traversing any location, and this cost correlates with distance. To operationalize this concept, we compute the distance from each household to the nearest major city and market. This calculation is facilitated by geospatial data pertaining to economic infrastructure, encompassing commercial centers and major cities, as illustrated in Figure 4, sourced from the Ministry of Infrastructure of Rwanda.

In our analysis, we opt for the Euclidean distance, representing an unconstrained straight line, as opposed to alternative measures like Geodesic distance, where travel is constrained to the surface of a sphere. This choice is deliberate, as these alternative measures exhibit a high degree of correlation with the Euclidean distance and convey an equivalent amount of information.

Figure 4: Transport and major towns in Rwanda



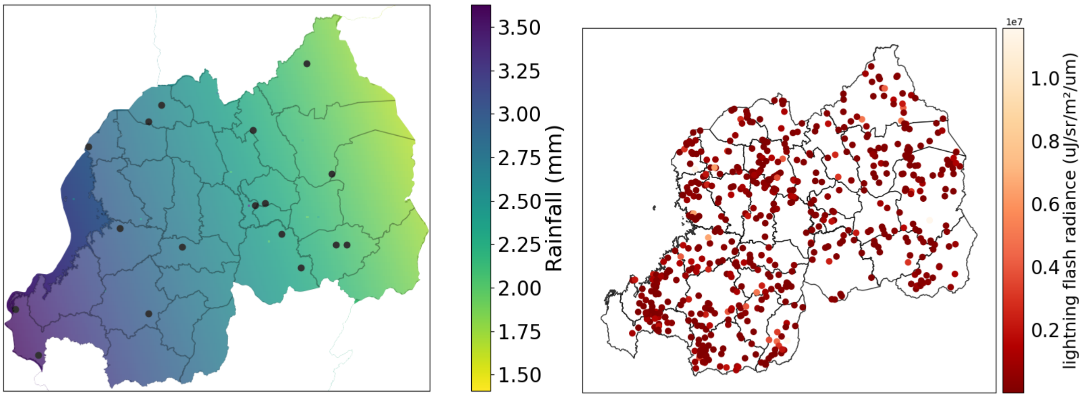
Finally, we use rainfall and lightning data to characterize weather and lightning activity in different regions of the country. We obtained rainfall data from the Rwanda Meteorology Agency which monitors and maintains datasets on weather and climate patterns in Rwanda. The agency has a data request portal through which researchers and other academics can request datasets spanning the last forty years. Through the data request portal, we obtained forty years of daily rainfall data from eighteen rainfall stations spread across the across. We calculated the average rainfall at each location of the country using geostatistical interpolation technique explained in Appendix I. The locations of the rainfall stations and the resulting distribution of average rainfall is presented in Figure 5 (a). We observe that the east is drier relative to the the west where rainforest is more common. These result are consistent with the values reported by the the Rwanda Meteorological Agency[12].

Instrumental variables for assessing reliability were derived from lightning data, utilizing the Tropical Rainfall Measuring Mission (TRMM) Lightning Imaging Sensor (LIS) satellite dataset. Specifically designed to observe and analyze lightning activity in tropical regions, the TRMM mission, a collaborative effort between NASA and the Japan Aerospace Exploration Agency (JAXA), was operational from 1998 to 2015 (Blakeslee, 1998). For this study, we use that last three years of lightning data from 2013 to 2015 to obtain the frequency, location and intensity of lightning events across Rwanda.

Our dataset contains 592 lightning events which includes all the flashes (strikes) recorded by the imaging sensor. Figure 5 (b) visually presents the distribution of lightning strikes across Rwanda during the period 2013-2015, illustrating the widespread occurrence of strikes throughout the entire country. Leveraging this data, we allocated lightning event statistics to the feeder region using bounding boxes as depicted in Figure 2. We calculated the average yearly number of strikes in the area and the average radiance (intensity) of all the flashes in the area. We then assign these values to each household in the feeder region, and hence, our lightning data measures lightning activity in the grid area which serves electricity to the household.

[12]https://www.meteorwanda.gov.rw/index.php?id=30

Figure 5: Rainfall and lightning activity



(a) Mean annual rainfall (mm)

(b) Lightning Events (2013-2015)

Note: Figure (a) presents the mean rainfall in Rwanda. Each dot is one weather station recording data on rainfall. Figure (b) presents all the lightning strikes in Rwanda. Each dot is a unique lightning flash recorder by the satellite TMMR, while the color shows the intensity of the strike.

3.1 A Closer Look

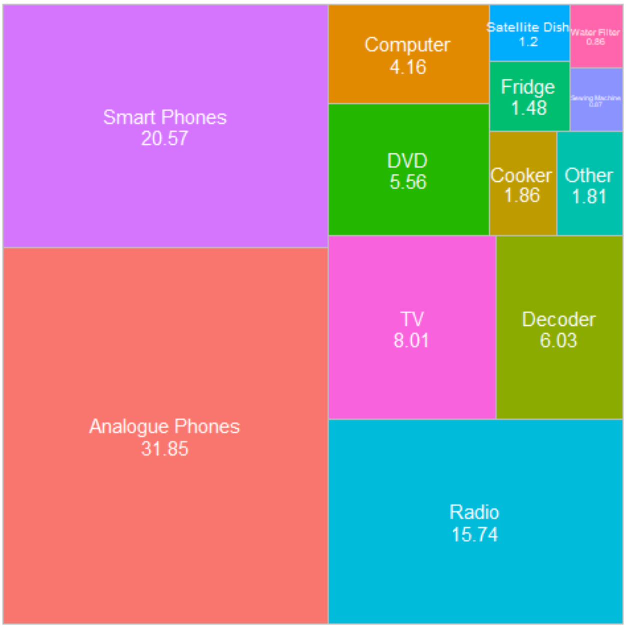
For the period October 2016 to October 2017, our sample of 3,600 grid electrified household owned a total of 15,510 appliances recorded in the data. Figure 6 shows the composition of this stock of appliances. In order to create this plot, we categorize mobile phones with internet access as smart phones and those without as analogue phones. Figure 6 shows that the total stock of appliances in our sample is mainly composed of mobile phones, radios, and televisions(TV). This observation is consistent with previous literature for Rwanda and Sub-Sahara Africa (Bos et al., 2018; Lenz et al., 2017; Muza & Debnath, 2021).

Figure 6 does not provide any insights on households appliance ownership. For this reason, Table 2 presents the number of grid electrified households which own at least one appliance. The table shows that almost 4% of the households do not own any appliances, signifying a noteworthy yet modest fraction reliant solely on electricity for lighting in their residences.

Among households with at least one appliance, the data indicates an average ownership of 4 appliances, with a maximum of 27 appliances owned by a single household. Remarkably, 86.70% of these households possess more than one appliance. The table further delineates the penetration rates and the quantity of units owned for various appliances identified in the EICV dataset. While radios and mobile phones are prevalent possessions among households, other categories exhibit lower ownership rates. Additionally, for households owning a particular appliance type, the average ownership is typically one unit, except for mobile phones, where the average ownership is nearly 2 per household. This observation will be important in determining the approach to modeling appliance ownership.

Table 2 shows that appliance ownership is limited both in quantity and variety in Rwanda. In order to understand this reality, Figure 7 presents the composition of appliances owned by households across the different appliance categories defined in Table 1. Figure 7(a) shows the share of household who own at least one appliance for each category; Figure 7(b) presents the distribution of units per category for those households who own at least one appliance for each category. Both plots illustrate a clear dominance of communication and entertainment appliances over other categories. This suggests that a limited

Figure 6: Composition of appliance stock Rwanda



Note: This figure represents the share of appliances owned by a sample of 3,600 grid electrified households in Rwanda. "Other" appliances include printers, camera, electric fans, hotplates, music systems and laundry machines.

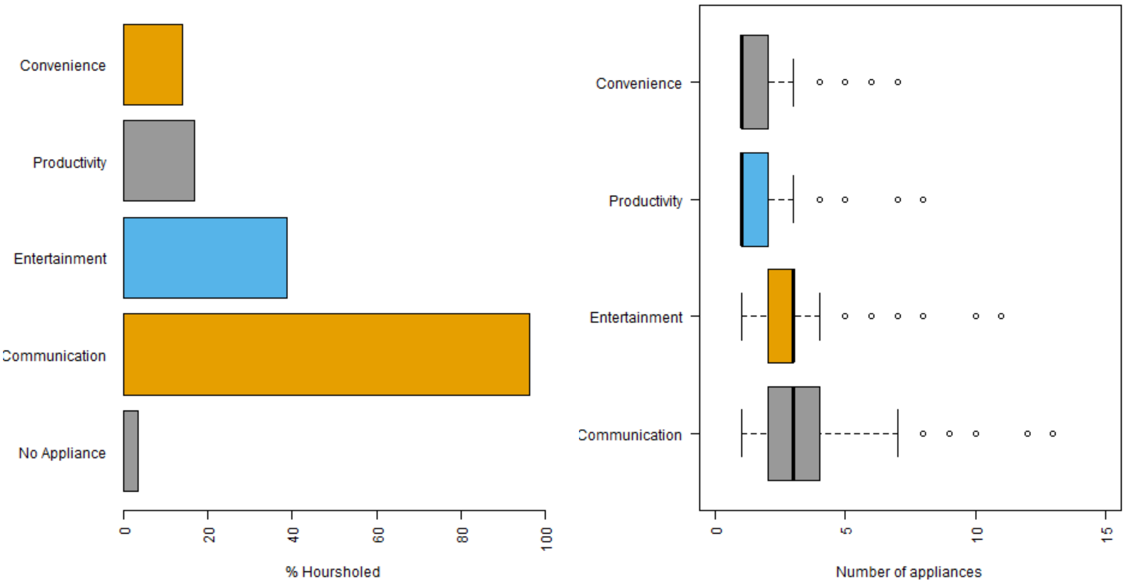
Table 2: Appliance distribution among electrified households

	Penetration		Households who own the appliance				
	Number	%	Mean	St. Dev.	Min.	Max.	HHs > 1 (%)
Any	3,480	96.670	4.440	3.330	1	27	86.70
Analogue Phones	2,955	82.080	1.840	1.100	1	11	52.83
Radio	2,324	64.560	1.150	0.430	1	5	13.08
Smart Phones	1,976	54.890	1.770	1.100	1	10	49.04
TV	1,322	36.720	1.030	0.190	1	3	3.03
Decoder	982	27.280	1.050	0.230	1	4	4.28
DVD player	897	24.920	1.060	0.400	1	10	3.90
Computer	520	14.440	1.360	0.710	1	6	26.35
Cooker	301	8.360	1.050	0.270	1	4	4.65
Fridge	242	6.720	1.040	0.240	1	3	3.31
Satellite Dish	198	5.500	1.040	0.190	1	2	3.54
Water Filter	146	4.060	1.000	0.000	1	1	0.00
Sewing Machine	109	3.030	1.370	1.020	1	8	18.35
Camera	79	2.190	1.130	0.430	1	4	10.13
Hotplate	73	2.030	1.030	0.160	1	2	2.74
Music System	71	1.970	1.060	0.290	1	3	4.23
Electric Fan	30	0.830	1.000	0.000	1	1	0.00
Printer	24	0.670	1.120	0.340	1	2	12.50
Laundry Machine	11	0.310	1.090	0.300	1	2	9.09

Note: The values in this table were calculated using the 3,600 grid electrified households in the EICV sample. The number of appliances was only calculated for the households who own at least 1 appliance of each category.

number of households have ascended the appliance ladder towards more expensive and electricity-consuming devices. Not only do fewer households possess convenience and productivity appliances, but the count of units owned by those households is also lower for these two categories compared to communication and entertainment appliances.

Figure 7: Ownership by appliance category



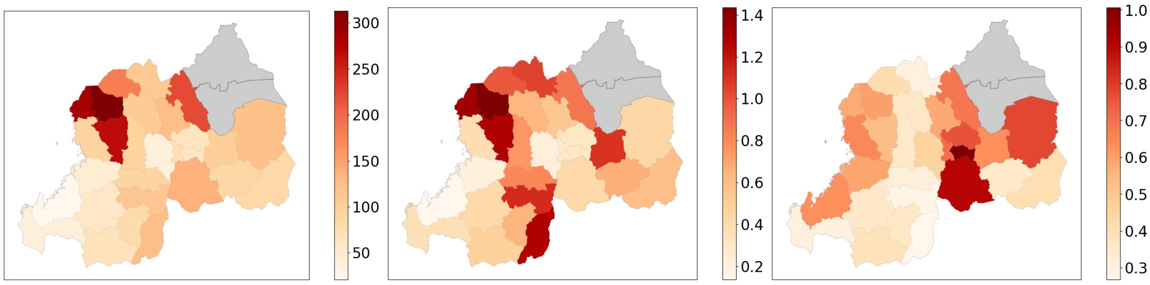
(a) Share of households (b) Number of units

The aforementioned adoption patterns suggest that limited affordability plays a pivotal role in constraining households from acquiring more expensive appliances. This outcome is unsurprising, given the prevalence of extreme poverty among several households in Rwanda(Sievert & Steinbuks, 2020). To contextualize, the average prices for cookers and ovens stood at 410 USD in 2018, while refrigerators and washing machines commanded average prices of 540 USD and 420 USD, respectively. (Source: Statista Market Insights[13]). Considering that the average annual income in 2018 was 780 USD (Sally Smith & Prates, 2020), a substantial portion of the monthly income is necessitated to acquire these appliances. For some households, this financial constraint translates into the limited use of electricity solely for lighting purposes. Consequently, it stands to reason that higher incomes and improved access to credit could potentially enhance appliance ownership in the country, subsequently contributing to increased electricity consumption.

However, previous literature suggests that non-income factors might be important in explaining residential appliance take-up (see Debnath et al. (2019), Blimpo and Cosgrove-Davies (2019), among others). The current state of the electricity sector in most African countries is characterized by pervasive reliability challenges (Blimpo & Cosgrove-Davies, 2019). Despite having better measures than other African countries, low electricity grid reliability could contribute to limited adoption patterns of certain appliances in Rwanda. Figure 8 present three statistics characterizing grid reliability at the district level in Rwanda. These variables are the total annual outage time in hours, the average daily outage frequency, and the outage occurrence time in hours. Note that figure 8 presents district averages for the exposition, but there is significant within district variation which we exploit in our statistical models.

[13]https://www.statista.com/outlook/cmo/household-appliances/major-appliances/rwanda#price

Figure 8: Average Reliability metrics per district - 2017



(a) Annual outage time (hrs) (b) Frequency (number per day) (c) Duration per outage (hrs)

The figure above reveals large geographic heterogeneity in reliability quality. Districts located at the nation's geographical center tended to experience a higher availability of electricity by 2017, coupled with a reduced frequency of outages in this particular region. On the other hand, households in the southern and northern districts suffered the most frequent outages while households in the central western part of the country suffered the longest outages in 2017. We are missing reliability data for the two most northern districts in our data set. There are 144 grid connected households in these two districts that we drop from our regression analysis as a consequence of this missing data.

In this context, it can be expected that improvements in reliability would increase incentives to buy appliances, and hence, electricity consumption and welfare impacts from electricity access would increase. Indeed, households could be adapting to low reliability by buying less appliances in total. Moreover, the stock of appliances owned by households might depend on the reliability level in the area where the household lives; for example, household in low reliability ares could favor appliances which rely less on the grid over other type of appliances. Yet, limited evidence exists on the role of reliability on appliance ownership. Understanding these relationships is important to predict how how households would respond to improvements in electricity reliability. Our goal is to determine empirically to what extent has reliability affected appliance adoption and usage patterns in Rwanda.

4 Appliance Ownership and Reliability

Our empirical goal is to studying the role of reliability on appliance ownership. Ideally, we would model how the probability of buying a given appliance changes as the grid reliability faced by each household changes over time (see Meeks et al. (2023) for an example). Regrettably, the nature of our available data precludes the execution of such a longitudinal study. Instead, we conduct a cross-sectional analysis in which we compare appliance ownership across regions with different reliability levels. We also analyze other drivers including income, demographics, education, gender, and dwelling characteristics.

4.1 Research Design

Our research design follows Richmond and Urpelainen (2019) and (Matsumoto, 2016a). It takes into careful consideration the intricacies of our particular case, appliance ownership in Rwanda, characterized by its limited prevalence, resulting in certain appliances being owned by only a few households and consequently generating an abundance of zeros in our dataset. This section explains our empirical strategy.

First, we analyze the role of reliability and other non-economic drivers on the total number of appliances owned by household by studying the intensity at which households invest in appliances with a conditional fixed-effects Poisson models. Let y_{ij} be the total number of appliances owned by household i in district j. Under the Poisson assumption, the probability of owning y_{ij} units of appliances is given by

$$Pr(Y = y_{ij}|X_{ij}, Z_{ij}, \alpha_j) = \frac{(E[Y|X_{ij}, Z_{ij}, \alpha_j])^{y_{ij}} \cdot e^{-E[Y|X_{ij}, Z_{ij}, \alpha_j]}}{y_{ij}!} \tag{1}$$

where $E[Y|X_{ij}, Z_{ij}, \alpha_j] = exp\{X_{ij}\beta + Z'_{ij}\Gamma + \alpha_j\}$ represents the anticipated number of appliances, dependent on a series of variables. Here, X_{ij} is the grid reliability; Z_{ij} is a vector of control variables which includes both income and non-income drivers; α_j are district fixed effect to capture common characteristics for households within the district. We do not use village fixed effects since villages are generally very small and therefore we won't have enough within village variation in our data. Note that even though Poisson models are inherently nonlinear, the use of the linear index and the exponential link function lead to multiplicative separability which allows us to estimate the model with fixed effects. Consequently, we employ the conditional maximum likelihood methodology proposed by Hausman, Hall, and Griliches (1984) to estimate this model.

We derive our measure of grid reliability, referred to as outage frequency[14] , from the average number of outages over the 24-month period spanning 2016 and 2017[15]. Unfortunately, we lack an extensive time series predating 2016/2017, prompting us to rely on the average for these two years to construct variables that encapsulate grid reliability. Note that the stability of reliability metrics across time is a hallmark, given that outage occurrences predominantly hinge on factors like weather conditions, vegetation interference, animal disruptions, feeder length, and various other determinants that exhibit minimal temporal variability.

In this context, our parameter of interest is the vector β which measures the change in the log of the expected number of appliances owned by a household when reliability improves by one unit. Using the point estimates we also calculate the incidence ratio rate, which measures the increase in the expected number of appliances owned by the household as reliability improves by 1 unit, by exponentiating the regression coefficient.

Equation 1 includes several controls variables aiming to control for households characteristics. First, we control for households income using the logarithm of the monthly expenses. Data on expenditures are more reliable than data on income, particularly for households at the low end of the distribution who may have substantial informal and non-monetary income. Expenditure data is highly correlated with income. Additionally, we include a control for the total savings owned by the household, recognizing that savings play a crucial role in determining the household's capacity to finance appliances and signify varying levels of wealth among households. The EICV survey contains information on households employment details such as nature of contracts, payment frequency, among others. We also incorporate control for the average turnover of jobs across household members, recognizing that households with frequent job changes may experience heightened uncertainty in their income sources. Previous authors suggest that uncertainty of future income might affect households' electricity uptake and consumption (Blimpo &

[14]We rely on outage frequency and not outage duration because outage frequency is highly correlated with the total hours a household does not have access to electricity in a given year. In Rwanda, the average duration of an outage was 20 minutes in the period 2016/2017 with a standard deviation of 7.4 minutes.

[15]Within our dataset, these years have the worst reliability performance. Indeed, there is a substantial reduction in grid reliability in the proceeding years (i.e. 2018,2019 and 2020). Note that the data for the year 2020 is incomplete which can explain part of the variation.

Cosgrove-Davies, 2019).

Next, we control for several demographic variables at the household level. First, we control for the gender of the decision makers with a dummy variable which takes value 1 if the head of household is female. Households with female decision-makers may exhibit distinct intra-household dynamics, influenced by historical roles where women have traditionally been primary caretakers with limited involvement in other facets of Rwandan society (Izabiliza, 2003). Additionally, women tend to exhibit lower levels of expenditure which are likely to impact and curtail appliance adoption (Richmond & Urpelainen, 2019). We also control for the number of children, women, and senior members in the household. Given the direct health impacts of indoor air pollution from cooking, lighting, and heating in rural contexts, seniors, women, and children are particularly vulnerable. Acknowledging their heightened susceptibility, we recognize that realizing the benefits of improved indoor air quality may require overcoming financial and informational barriers (Richmond & Urpelainen, 2019).

The level of education and skills at the household level can be an important driver of appliance ownership. (Dhanaraj, Mahambare, & Munjal, 2018b) recently found that refrigerator ownership was higher among more educated households. Enhanced education levels may afford individuals greater knowledge about appliances and their applications. Given that these variables can exhibit noteworthy geographical variations, potentially correlated with reliability, we introduce a dummy variable that takes the value of 1 if the head of the household has attended school. Additionally, we include two supplementary dummy variables: one indicating the presence of a business within the household and another signifying the involvement of household members in high-skill occupations.

We also control by the house characteristics and access to major towns and markets. Limited access to major towns and markets may result in reduced exposure to household appliances and commerce (Richmond & Urpelainen, 2019), making it more likely that households with such access are inclined to acquire certain appliances. Firstly, we introduce a dummy variable, taking the value of 1 if the household is located in a rural area, as these areas typically exhibit smaller exposure and access to appliances compared to urban areas. Secondly, we incorporate the Euclidean distance to the nearest major town and trade center. Furthermore, we address house characteristics. Initially, we include the number of rooms, recognizing that larger households might possess the same appliance multiple times. Next, we introduce a dummy variable for houses with multiple buildings, reflecting the potential for duplicated appliances. Considering the dynamic nature of appliance ownership, with household owners being more likely to acquire appliances that are less easily moved, we incorporate a dummy variable, taking the value of 1 if the household is the owner of the house. Lastly, we introduce a variable controlling for the number of years the household has resided in the current location.

We incorporate controls for the mean rainfall in the locality where the household resides, recognizing that household members' time utilization can be influenced by the local climatology Sakah, du Can, Diawuo, Sedzro, and Kuhn (2019a). Certain appliances may hold varying levels of value under specific weather conditions[16] (Cabeza, Ürge-Vorsatz, Ürge, Palacios, & Barreneche, 2018). Rwanda, being a small tropical country with consistently warm temperatures throughout the year, exhibits four primary climatic regions: eastern plains, central plateau, highlands, and regions around Lake Kivu. The central plateau region maintains a mean annual temperature between 18°C and 20°C, the eastern plains have a mean annual temperature oscillating between 20°C and 22°C. The highlands are colder with temperature between 10°C and 18°C. Lake Kivu and Bugarama plains

[16]For examples, fans are more valuable in regions were hot weather is more often. Some appliances, on the contrary, suffer from some weather conditions; for example, TV satellite dishes are affected by rainfall.

have annual mean temperatures between 18°C and 22°C. Although temperature variations within districts are limited due to the country's small size, such variations are adequately captured by the district fixed effects.In contrast, rainfall patterns exhibit considerable diversity, ranging from 1,000 to 1,400 millimeters (40 to 55 inches) annually, depending on the area, and presenting within-district variation[17]. To account for this, we introduce a control for mean rainfall in the household's vicinity, defined as the average across a 5kmx5km grid.

Finally, we introduce a control for the utility's capacity to restore service, gauged by the average duration of outages in the area. Anticipated to exhibit heterogeneity across regions and correlation with outage frequency, the utility's proficiency in service restoration is a crucial factor. This proficiency, linked to the duration of outages, may influence households' preferences for using certain appliances at specific times (McRae, 2010). Utilizing the average duration of outages from 2016 to 2017, we incorporate this control variable to assess the utility's response time[18]. While the cause of a power outage can influence the restoration duration[19], the average duration of outages serves as a reliable metric for evaluating the utility's overall responsiveness to outages.

The variables used in the analysis are summarized in Table 3. Specifically, our regression models incorporate data from 2,706 grid-connected households for which complete variable information is available.

Table 3: Summary statistics (Number of Obs = 2,706)

Variable	Mean	St. Dev.	Min.	Max.
Reliability				
Average duration without electricity (hrs/year)	116.920	92.750	17.260	496.500
Average Frequency (outages/day)	1.069	0.837	0.064	2.91
Average Outage Duration (min/outage)	20.040	7.410	9.670	67.110
Income and employment				
Expenditure (log RWF month)	11.570	0.930	8.790	14.580
Savings (million RWF)	0.250	2.320	0	99
Has business (dummy)	0.450	0.500	0	1
Job instability (number of jobs/member)	1.480	0.620	1	7
Involves in high skill occupation (dummy)	0.190	0.370	0	1
Demographics				
Female (number)	2.310	1.580	0	13
Children (number)	1.860	1.640	0	9
Seniors (number)	0.140	0.410	0	3
Head of Household				
Female (dummy)	0.190	0.390	0	1
Below 35 years old (dummy)	0.430	0.490	0	1
Rwandese (dummy)	0.990	0.100	0	1
Attended School (dummy)	0.250	0.440	0	1
Dwelling and ownership				
Number of rooms (count)	3.760	1.610	1	10
Multiples houses (dummy)	0.180	0.380	0	1
Multiples households (dummy)	0.280	0.450	0	1
Number of years in house (count)	6.850	8.850	0	63
Own house (dummy)	0.560	0.500	0	1
Location				
Rural (dummy)	0.450	0.500	0	1
Distance to major town (km)	9.600	7.900	0.050	43.860
Distance to trade center (km)	1.780	1.540	0.010	9.990
Mean rainfall (mm)	2.600	0.370	1.850	3.690

[17] The wet season months in Rwanda are from March to May and from September to December.

[18] Unfortunately, we were not able to find good instruments for the duration, and hence, we are cautious in analyzing the coefficient for this variable. We expect this variable to be endogenous.

[19] Utilities can fix a minor incident quickly, but when the cause of a blackout is a natural disaster, you can expect to be out for several days and even months in some extreme cases.

In the second part of the analysis, we study empirically the effect of reliability on the composition of appliances owned by the household, that is, the ownership of key appliances. A challenge on studying individual appliances in our setting is that the penetration rate of most appliances is small, and hence, the data has an abundance of zeros, surpassing the accommodation capacity of a typical count distribution.[20] (Hilbe, 2014). Indeed, the penetration rate for most appliances is below 30% (see Table 2). In this context, we conceptualize ownership of key appliances as a hurdle model[21] (see Feng (2021) for a discussion). In this model, appliance ownership is viewed as a two-step decision: households initially determine whether to acquire an appliance, and among those willing to invest, they subsequently decide on the quantity to procure.

To model the household's willingness to invest in different type of appliances we let y^{ℓ}_{ij} be the number of appliance ℓ owned by the household, and we define $q^{\ell}_{ij} = 1[y^{\ell}_{ij} > 0]$ as dummy variable which takes value 1 if appliance ℓ exists at the household. While non-linear models are commonly employed for discrete choice outcome variables, we opt for linear models in this context. This choice mitigates the risk of incidental parameters, which may arise due to fixed effects and clustering[22], and enhances the interpretability of the results. Although conditional logit models could be used to estimate these models, addressing panel fixed effects through a likelihood function transformation (see Chamberlain (1980)), certain appliances in our dataset are owned by very few households (see Table 2), constituting rare events in the literature. This rarity can introduce bias in binary nonlinear models (see (King & Zeng, 2001)). As our primary interest lies in utilizing statistical models to understand directional and relative relationships between variables, we are less concerned about bias in our linear models with binary outcome variables. Consequently, our fixed-effects models for each appliance type ℓ are expressed as follows:

$$q^{\ell}_{ij} = X_{ij}\beta_{\ell} + Z'_{ij}\Gamma_{\ell} + \alpha_j + \varepsilon_{ij} \tag{2}$$

where X_{ij}, Z_{ij}, and α_j are defined as in Equation 1. In this specification, each coefficient is allowed to vary by appliance, and we estimate this model as seemingly unrelated equations. We also run an alternative model for each category $c = \{1, 2, 3, 4\}$ defined in Table 1. In this case, the dependent variable is a dummy variable for each category c and takes value 1 if the household owns at least one appliance of category c. Again, we allow the coefficients to vary by appliance type by estimating individual equations for each category.

Apart from the question of the presence or absence of a particular appliance in the household, we are also interested in how many appliances of a particular type the household possesses, conditional on having invested in the appliance. Following the Poisson distribution, conditional on having invested at least once in appliance ℓ, the probability of owning y^{ℓ}_{ij} units of appliance ℓ is given by

$$Pr(Y = y^{\ell}_{ij}|X_{ij}, Z_{ij}, \alpha_j, q^{\ell}_{ij} = 1) = \frac{(E[Y|X_{ij}, Z_{ij}, \alpha_j, q^{\ell}_{ij} = 1])^{y^{\ell}_{ij}} \cdot e^{-E[Y|X_{ij}, Z_{ij}, \alpha_j, q^{\ell}_{ij}=1]}}{y^{\ell}_{ij}!} \tag{3}$$

where $E[Y^{\ell}|X_{ij}, Z_{ij}, \alpha_j, q^{\ell}_{ij} = 1] = exp\{X_{ij}\beta_{\ell} + Z'_{ij}\Gamma_{\ell} + \alpha_j\}$.

[20] Excess of zero counts are described as zero-inflated in the statistics literature (Hilbe, 2014).

[21] Such a model, originally proposed by Mullahy (1986), overcomes this problem by not constraining the intensive and extensive problem to be the same. Moreover, this model deals with the many zeros situation.

[22] Linear panel data models use the linear additivity of the fixed effects to difference them out and circumvent the incidental parameter problem present in nonlinear models as the fixed-effect ordered logit model Richmond and Urpelainen (2019).

We then fit a standard fixed-effects count model on the sub-sample of those who have invested in the appliance under analysis. There are two challenges when estimating equation 3. Firstly, in our sample, households typically own few appliances more than once(see Table 2), implying that a limited number of appliances can be adequately modeled following Equation 3[23]. Secondly, as previously mentioned, the majority of appliances are possessed by only a small fraction of households. Given this scenario, the subset of households that have invested in more than one appliance is generally too modest to effectively estimate Equation 3. Consequently, our examination of key appliance ownership primarily revolves around the binary decision of whether the household owns a particular appliance or not. We exclusively present estimates of Equation 3 for those appliances where we have sufficient observations for estimation and where we anticipate a Poisson distribution namely, phones, computers, and radios.

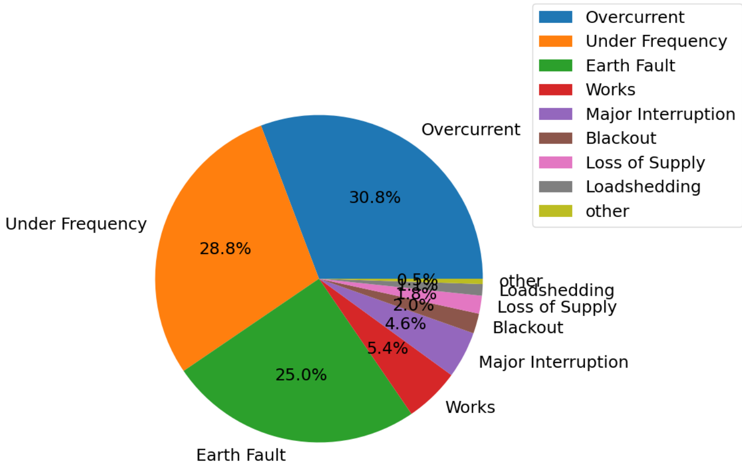
4.1.1 Identification

Estimating the relationship between electricity reliability and household outcomes is typically challenging. Service quality is often endogenous and correlated with household characteristics. Two key factors contribute to this complexity. Firstly, the non-random nature of household locations, influenced by regional factors such as weather and economic activity (Pawar & Jha, 2023; Sinha, Caulkins, & Cropper, 2018). These factors have an important role determining reliability levels of an electric system. Figure 9 shows that overcurrents, under-frequency and earth faults are the predominant causes of outages in our data, and these are the consequence of regional factors such as weather, vegetation, and electricity demand. Secondly, grid reliability is contingent on utility decisions, including maintenance and grid design, showcasing substantial regional variations that correlate with household characteristics (Meeks et al., 2023). Indeed, the distribution network design typically adopts radial feeders for rural areas in Rwanda, in contrast to networked feeders commonly seen in urban locales[24] (REG, 2021). Radial networks usually have lengthy feeder lines rendering them more susceptible to outages.

[23]Following our data, these appliances are phones, computers, sewing machines, printers, cameras and radios. Modeling as count data other appliances might not be accurate since the data does not follow a Poisson distribution. Maximum likelihood models require fully-specified models and misspecification leads to violations of the information matrix equality

[24]The distribution network can have a radial or networked configurations. Radial networks lack interconnections with alternative supply points, while networked networks boast multiple connections to diverse supply sources. Radial networks are used in rural Rwanda due to the isolated nature of rural loads, rendering the use of networked feeders economically less feasible (REG, 2021).

Figure 9: Main causes of outages



Inevitable measurement errors in power infrastructure quality exist (Chen, Jin, Wang, Guo, & Wu, 2023). Despite possessing novel data on reliability, our dataset may not precisely align with localized outages, which often go unnoticed in utility tracking. Achieving a comprehensive match between these measured outages and household-level outages proves challenging, particularly given extensive feeder lines that stretch over considerable distances and branch into multiple distribution spurs serving smaller communities. This inherent limitation poses a methodological challenge. Our use of feeder outages serves as a proxy to characterize the "standard" service quality experienced by households. As acknowledged in econometrics literature, measurement errors in the independent variable result in attenuation bias (Bollen, 1989; Wooldridge, 2010).

In this context, fixed effects can eliminate the time-invariant effects, but the endogeneity problem still cannot be solved. Similarly, our control variables fail to capture for unobserved heterogeneity of, for instance, differences in financing schemes within districts (across sectors), or other unobserved variables at the household level which might be correlated with household location. Consequently, our identification strategy relies on the use of two instrumental variables[25] which characterize the lightning activity in the different parts of the country: average radiance of lightning strikes[26] and number of lightning strikes.

Lightning disturbances are usually a significant issue for electricity networks and service interruptions (Rezinkina, Babak, Gryb, Zaporozhets, & Rezinkin, 2022). For example, lightning damage accounts for about 65% of distribution network failures in South Africa (McDonald et al., 2011). The energy carried by a single lightning bolt is immense, averaging around 1 gigavolt with a typical current from 10,000 to 30,000 amperes (Gunther, 2023). The heat produced can also be substantial, reaching temperatures five times hotter than the surface of the sun (Rezinkina et al., 2022). Hence, when a lightning strike hits close to the electricity network the following events may occur: line overvoltages which exceed the insulation capabilities; transient currents that propagate through the network[27]; damages in transformers, poles, conductors, insulators, substation, and transformers due to high currents, voltages, and intense heat; and temporary disruptions in the grid due

[25] The rank condition establishes that we need at least 2 valid instruments for the identification of the model.

[26] Radiance is used to characterize diffuse emission and reflection of electromagnetic radiation, and to quantify emission of neutrinos and other particles.

[27] Transients create high-frequency harmonics and voltage spikes into the power system. They can disrupt sensitive equipment, leading to malfunctions, faults, or even tripping of protective devices

to line tripping, automatic reclosing, or protection system operations. Finally, lightning strikes can induce high electromagnetic fields that can affect the operation of the grid.

We instrument outage frequency with the average frequency of lightning strikes. Previous evidence has found that in areas with high lightning density, the frequency of power outages is higher (Chisholm and Cummins, 2006). Additionally, we introduce a second instrument: average intensity of the lightnings in the region measured by the average radiance of the lightning flashes. The probability of observing a grid failure is directly associated with the intensity of the lightning strikes. Summary statistics are presented in Table 4.

Table 4: Instrumental Variables (Number of Obs = 2,706)

Variable	Mean	St. Dev.	Min.	Max.
Average Lightning Radiance (millon uJ/sr/m^2/um)	0.500	0.260	0.100	2.120
Frequency Lightning (count/year)	7.170	8.590	0.330	27

Our reduced-form equation for reliability is given by

$$X_{ij} = W'_{ij}\Pi + Z'_{ij}\Lambda + \alpha_j + \varepsilon_{ij} \tag{4}$$

where X_{ij} is the outage frequency, W'_{ij} are our instruments, Z_{ij} are the control variables from our structural equation, and α_j district fixed-effects. For the instruments to be valid, $\Pi \neq 0$. Table 5 presents the results from our first-stage reduced-form regression. As observed in the table, the coefficients are positive and significant, meaning that the average number of outages increases with the frequency of lightning strikes and the intensity of the strikes. Furthermore, the results affirm the instruments' relevance, substantiated by both the F-statistics[28] and the Cragg-Donald Wald-F statistic[29].

[28] Staiger and Stock (1997) establish the rule-of-thumb for this test: if the F-statistic is less than 10, the instruments are weak and no valid statistical inference can be made. Hence, we use a value of 10 for the F-statistic as the threshold for the relevance test because we want the IVs to be strongly significant (not just significant). Moreover, we want the first-stage F-statisitc to be above 10 so that the relative bias of 2SLS, relative to OLS, is less than 10% (using the instruments have a real advantage).

[29] Critical values are presented at the bottom of the table

Table 5: First-Stage Results

Dependent variable: Frequency of Outages (number/day)	
Lightning Radiance	0.463**
	(0.202)
Lightning Frequency	0.069***
	(0.006)
Relevance and Weak-IV Test	
F-statistic	401.89
Cragg-Donald Wald-F statistic[a]	1477.14
Overidentification	
J-Statistic	0.918
$\chi^2(1)$ p-value	0.3379
Observations	2,706
Number of district	26
Mean observations per group	104.1

Note: All the exogenous variables from the structural equation are including in the first-stage regression, including the district-fixed effects. Clustered standard error at the district level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1
[a]: Stock-Yogo (2005) weak IV F–test critical values for single endogenous regressor: 19.93 (10% maximal IV size); 11.59 (15% maximal IV size);8.75 (20% maximal IV size)

In this context, our identification assumption is that conditional on the control variables, our instruments are exogenous and uncorrelated with the structural error term. We believe this a plausible assumption given that lightning is a random act of nature that can strike anywhere and at any time of the year (Oceanic & Administration), 2020). Three factors affect the formation of lightning strikes. Regions were lightning is more probable often have ideal meteorological conditions conducive to thunderstorm development, including warm temperatures and strong solar radiation, which can drive moist air into convection currents capable of lifting negative charges higher in the atmosphere[30] (Gunther, 2023). Yet, a storm is not a sufficient condition for a lightning strike (Oceanic & Administration), 2020). Lightning strikes only happen if regions of excess positive and negative charge develop between clouds and the ground. Finally, the intensity of a lightning strike depends on the electric charges inside the clouds. These factors are difficult to be correlated to the economic and social characteristics of the region (Chen et al., 2023). In fact, lightning strikes occur all around the world, so every region has potential risk (Gunther, 2023). Moreover, solar radiation and temperature are not expected to have important differences across Rwanda as the country is small. To support our assumption, we present the J-statistic and the p-value for the over-identification test in Table 5. The results show that we fail to reject the null hypothesis that the instruments are exogenous, meaning that they are good instruments.

It's essential to note that, for our linear models, we implement 2SLS fixed effects models. However, the 2SLS approach is not universally valid for nonlinear models and may not produce a consistent estimate. In cases where the second-stage equation involves nonlinearity, as seen in our Poisson models, the predicted endogenous variable from the first-stage regression can become correlated with the residuals. To address this challenge, we opt for the Control Function Approach (CFA) in a two-step process to enhance our preliminary results. For nonlinear models, especially those like Poisson models, we apply a

[30]This allows positive charges below to attract them, creating powerful discharges of electricity known as lightning.

two-step Control Function Approach (CFA), wherein we incorporate the predicted residuals from the first stage into the second stage (Wooldridge, 1997; Rivers and Vuong, 1988). In this approach, bootstrap standard errors are employed to accommodate the uncertainty stemming from the first stage.

4.2 Empirical Results

This section presents our empirical results. First, we discuss the role of reliability on appliance ownership. Then, we provide descriptive evidence of other drivers of household's demand for appliances. In this second part of the analysis, we are cautious on how we interpret the regression results as coefficients cannot be understood as casual but a correlation.

4.2.1 The Role of Reliability in Ownership

Table 6 presents the regression outcomes for the households' total appliance ownership, specifically highlighting the estimated relationship between overall appliance possession and reliability. We employ conditional fixed-effects Poisson models to analyze the total number of appliances. Each column presents a different specification, and we present the estimated coefficients and the incidence-rate ratios obtained by exponentiating the coefficients. Additionally, we explore alternative model versions using conditional fixed-effects, negative binomial regression and a linear fixed-effects. The assumptions of the Poisson regression are quite restrictive: the mean is assumed to be the same as the variance. Hence, we relax these assumptions in the last columns of Table 6. Appendix 2 presents all the regression results for all the variables used in the empirical study.

Table 6: Reliability and Total Number of Appliances

	Conditional Fixed-Effects Poisson						FE + Instrumental Variables		
	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	Mod. 6	Poisson	Neg. Bin.	2SLS
Frequency of Outages (number/day)									
Point Estimate	0.0001	0.010	0.003	0.001	-0.010	-0.028	-0.055	-0.056	-0.232**
	(0.052)	(0.037)	(0.033)	(0.031)	(0.023)	(0.019)	(0.054)	(0.055)	(0.068)
Incidence Ratio	1.000	1.010	1.002	1.001	0.999	0.971	0.946	0.946	-
Control Variables									
Income and employment		Y	Y	Y	Y	Y	Y	Y	Y
Demographics			Y	Y	Y	Y	Y	Y	Y
Head of Household				Y	Y	Y	Y	Y	Y
Dwelling and ownership					Y	Y	Y	Y	Y
Location						Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Wald	0.00	3568.56	10076.80	7987.79	25062.76	35460.09	83673.86	34979.94	
Log pseudolikelihood	-7109.56	-5468.70	-5358.95	-5300.72	-5205.33	-5193.05	-5192.44	-5193.49	
F-statistic									16476.82
R^2									0.6336
Observations	2,706	2,706	2,706	2,706	2,706	2,706	2,706	2,706	2,706
Number of district	26	26	26	26	26	26	26	26	26
Mean observations per group	104.1	104.1	104.1	104.1	104.1	104.1	104.1	104.1	104.1

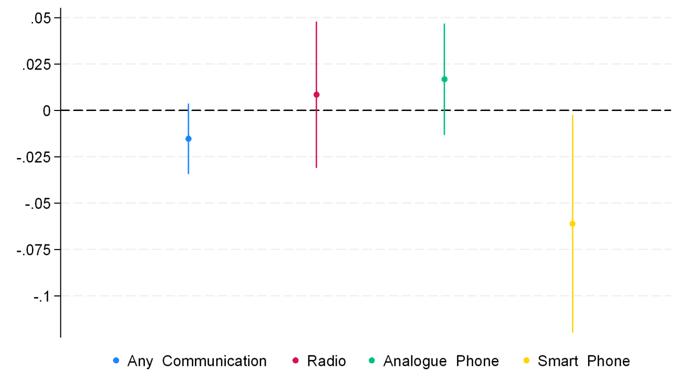
Note: Clustered standard error at the district level in parenthesis for linear model as well as the conditional fixed-effects poisson model. For the two-step negative binomial and poisson model, the standard errors are bootstrapped.
Total number of appliances include radios, phones, TV, decoder, satellite dishes, cookers, fridges, DVDs, music systems, computers and printers, cameras, hotplates, electric fans, laundry machines, water filters and sewing machines. *** p<0.01, ** p<0.05, * p<0.1

The results in Table 6 underscore the importance of accounting for household characteristics when studying how reliability affects household outcomes. In fact, we can observe that the coefficient for frequency of outages becomes negative as we include additional control variables. Similarly, controlling by the residual of the first-stage regression reduces the magnitude of the coefficients towards zero, providing indication that unobservables would bias the coefficient if we do not account for the endogeneity of reliability. In our preferred model, the two-step conditional fixed-effects Poisson model, we observe a negative relationship between the total number of appliances owned by households in Rwanda and our variable of interest, although it is not statistically significant. This outcome aligns with the results from other models that encompass all control variables, excluding the 2SLS model. While the 2SLS model yields a statistically significant coefficient at a 5% confidence level, it's worth noting that the magnitude of the coefficient is very small.

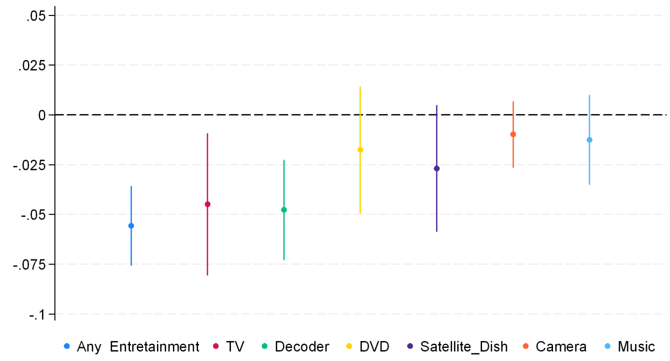
Our results suggest that the frequency of outages does not exert a significant impact on the number of appliances owned by households in Rwanda. One possible explanation is that households might lack information about grid quality or perceive the true system quality differently. If this were the case, we wouldn't observe any effect of reliability on the composition of the appliance stock. Alternatively, the results in Table 6 can be a consequence of affordability constraints. Given the low median income in Rwanda relative to appliance costs, households may experience limited demand for appliances due to budget constraints, and variations in reliability may not significantly influence the overall number of appliances they own. Nevertheless, households can still adapt by substituting between appliances, thereby affecting the composition of their appliance stock.

Figure 10 presents the probability of households investing in different appliance categories, including key appliances, based on differences in the frequency of outages in their respective areas. The plotted coefficients quantify the disparity in the willingness to invest in each category for households residing in areas with varying outage frequencies. For a comprehensive overview, Appendix 2 provides all regression results for the variables used in the empirical study.

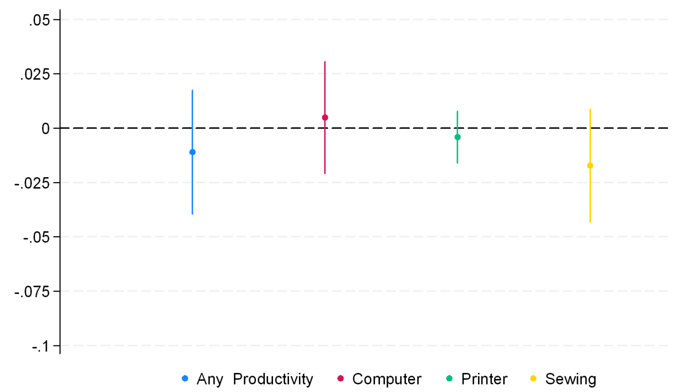
Figure 10: Reliability and Willigness to Invest in Appliances



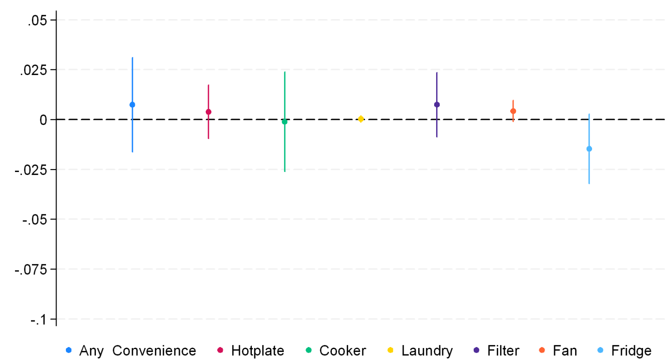
(a) Communication



(b) Entertainment



(c) Productivity



(d) Convenience

As depicted in the figure, the likelihood of a household investing in a television or decoder diminishes as outage frequency increases. Specifically, our results show that one additional outage per day reduces the probability of owning a television by 4%, and for decoders the change is nearly 5%[31]. The absence of electricity impedes households from watching television, and since these appliances typically serve as the primary entertainment sources, their ownership is negatively influenced by reliability. The scenario is intriguing for communication appliances, as indicated by Figure 10 (a). In areas with robust reliability, households do not exhibit a notably higher probability of investing in these appliances compared to those in areas with poor reliability. However, larger frequency of outages is associated with a lower share of household owning a smart phone -i.e. one extra outage per day decreases the probability of a household investing in this type of phone by 6%. This phenomenon can be attributed to the fact that smartphones are more energy intensive than analogue phones and therefore would require more frequent charging, which becomes challenging in areas with poorer reliability.

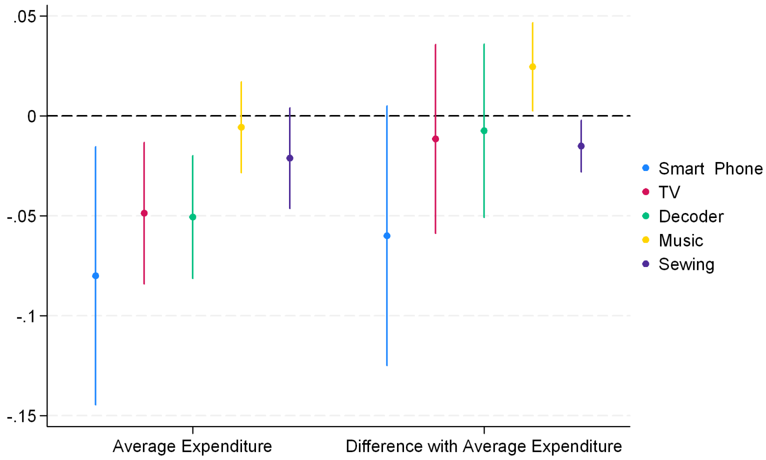
To complement these results, Figure 11 presents the estimated coefficients obtained by interacting our reliability variable with the centered logarithm of expenditure[32]. The coefficients on the left side of the plot depict the changes in the probability of investing in key appliances when reliability is worse by one unit at the average expenditure level. It's

[31]The larger estimated coefficient for decoders than for televisions can be explained by the fact that televisions, in general, are necessary for decoders.

[32]We re-run the first stage by using a non-linear function of our instruments given our interaction term to avoid the ”forbidden” regression

important to note that these coefficients are consistent with those in Figure 10. On the right side, the additional effect of outage frequency is presented, accounting for the deviation of expenditure from the mean value.

Figure 11: Expenditure, Income, and Willigness to Invest in Appliances



The figure highlights nuanced patterns in household appliance investments based on income levels. Specifically, the probability of investing in smart phones, televisions, and decoders decreases for households with average expenditure, while the probability of investing in music equipment rises for those with above-average expenditure. Conversely, households with below-average expenditure are more likely to invest in sewing machines.

These findings suggest adaptive behavior among households, with a tendency to steer away from appliances heavily reliant on the electricity grid (e.g., smart phones, televisions, and decoders). Instead, wealthier households lean toward music equipment, often equipped with independent power sources, while less affluent households opt for treadle sewing machines, known for their minimal electricity consumption.

Finally, Table 7 presents the impact of reliability on the number of units owned by households for key appliances, focusing solely on appliances with sufficient data for estimating Equation 3. The coefficients for all appliances are negative, with the exception of analogue phones; however, none of the coefficients achieves statistical significance. This implies that the investment intensity in these appliances is not substantially influenced by reliability.

Table 7: Number of Key Appliances Owned by Households

	Radio	Phone: Analogue	Phone: Smart	Computer
Frequency of Outages (number/day)				
Point Estimate	-0.053	0.086	-0.014	-0.043
	(0.048)	(0.066)	(0.081)	(0.187)
Incidence Ratio	0.948	1.089	0.985	0.958
Control Variables				
Income and employment	Y	Y	Y	Y
Demographics	Y	Y	Y	Y
Head of Household	Y	Y	Y	Y
Dwelling and ownership	Y	Y	Y	Y
Location	Y	Y	Y	Y
District FE	Y	Y	Y	Y
Wald	2684.00	78688.24	48160.64	1012.43
Log pseudolikelihood	-652.00	-1067.03	-711.15	-142.91
Observations	656	829	579	150
Number of district	26	26	26	18
Mean observations per group	25.1	31.9	22.3	8.3

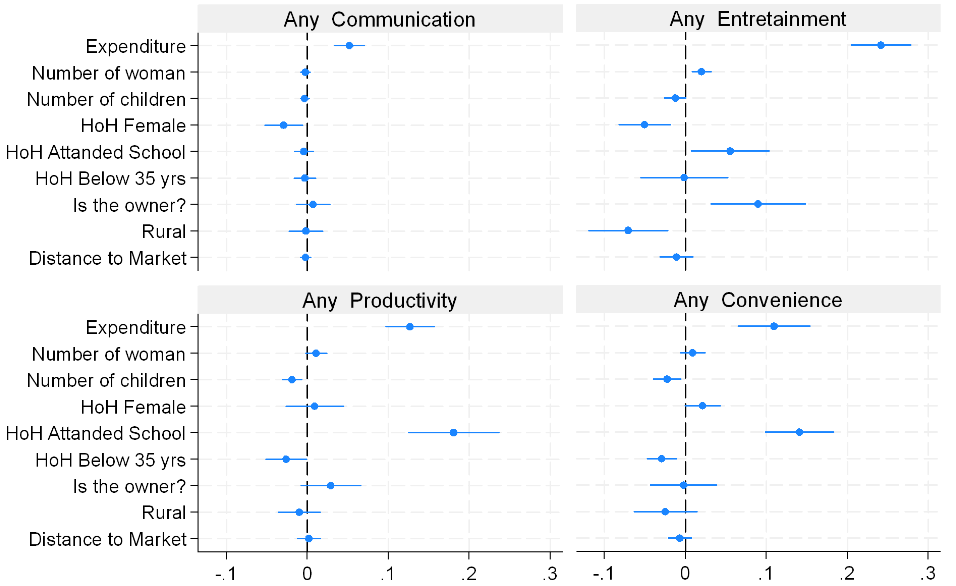
Note: Two-step conditional fixed-effects models were used to estimate these models. Standard errors are boostrapped. *** p<0.01, ** p<0.05, * p<0.1

Overall, our findings indicate that households in Rwanda exhibit forward-looking behavior, adapting to low reliability levels by fine-tuning the composition of their appliance stock rather than altering the total number of units owned. Given the prevalent affordability challenges in Rwanda, households may face limitations in acquiring additional appliances. Wealthier households may pivot towards music devices, while less affluent households to sewing machines. This adaptive behavior underscores households' strategic optimization of expected utility, factoring reliability into their appliance selection process. Thus, households navigate low reliability by adjusting their appliance mix within similar cost ranges.

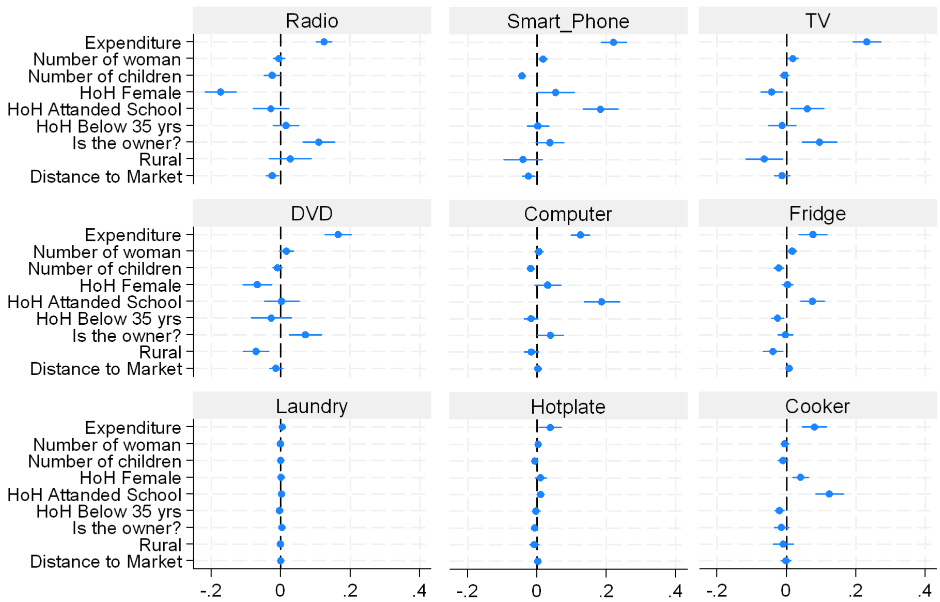
4.2.2 Other Factors Affecting Appliance Ownership

This section provides descriptive evidence of other drivers of appliance ownership. Note that these results are not casual and should be interpreted carefully. Appendix 3 presents the regression tables. We summarized the results for appliance ownership of key appliances in Figure 12.

Figure 12: Other Drivers of Appliance Ownership



(a) Categories



(b) Key Appliances

Figure 12 shows a substantial positive correlation between household financial variables and appliance ownership, with the exception of laundry machines, which could be attributed to their high cost or cultural considerations. As shown in the regression tables presented in Appendix 3, households experiencing high job turnover among members tend to possess fewer appliances, indicating the impact of financial uncertainty on household appliance ownership. These findings suggest that programs offering subsidies for appliances could have a profound impact on increasing appliance ownership. However, it is noteworthy that not only income levels but also the associated uncertainty play a crucial role.

Demographic characteristics of the households also shows a significant relationship with the dependent variables. Historical gender roles, where women traditionally take on caretaking responsibilities in Rwandan homes (Izabiliza, 2003), could explain why households with more females tend to invest in more appliances, possibly for entertainment, support with chores, or home-based productive activities. Homes with many children on the other hand have a negative significative relationship with number of appliances in the home as well as the probability of investing in smart-phones, fridges, computers, among others. It might be expected that these households would devote a larger portion of their expenditure towards their children's needs like education and health care.

The age and gender of the head of the household play significant roles in determining the household's appliance ownership. The regression table in Appendix 3 highlights a noteworthy pattern: female heads of households and those below the age of 35 exhibit a negative and statistically significant relationship with the total number of appliances owned by the household. This finding may be attributed to existing socioeconomic disparities between genders, where female heads of households could face lower incomes and limited access to resources compared to their male counterparts. Younger heads of households might also experience lower financial stability, contributing to their reduced ownership of appliances. Additionally, the results presented in Figure 12 shed light on distinct demographic based expenditure patterns. Specifically, households led by females are inclined to own more convenience appliances but fewer entertainment appliances. Notably, female-headed households show a higher likelihood of owning cookers, possibly reflecting traditional gender roles in Rwandan society where women are primarily responsible for household tasks, including cooking.

Education levels of the head of the household significantly influence the types of appliances owned, revealing important insights into household appliance ownership patterns. Specifically, households led by individuals with higher education levels tend to own a greater number of appliances. Moreover, the likelihood of a household investing in specific categories of appliances, such as entertainment (e.g., TVs), productivity (e.g., computers), and convenience (e.g., cookers and fridges), increases with the education of the head of the household.This positive correlation can be attributed to several factors. Educated heads of households are likely to possess more information about the benefits of various appliances, enabling them to make informed decisions about their utility. Additionally, their higher level of education may equip them with the necessary skills to operate certain appliances, particularly those in the productivity category, such as computers.

Lastly, dwelling characteristics, such as the number of rooms in a house, exhibit a positive and significant relationship with the dependent variable. Larger houses are associated with a higher likelihood of owning appliances, particularly items like televisions, suggesting that the demand for appliances is influenced by the spatial requirements of a household. Additionally, the stability brought about by long-term homeownership may encourage households to acquire more appliances over time. In contrast, rural households and multiple families sharing a home are less likely to have an extensive collection of appliances.

5 Linking Appliance Ownership to Electricity Consumption

We extend the analysis by assessing how residential electricity consumption depends on ownership of key appliances. Our empirical results suggest that the composition of the stock of appliances owned by households is affected by low reliability, and therefore, we quantify these empirical results in terms of electricity consumption by estimating the appliance-specific electricity consumption. We follow the conditional demand model proposed by Larsen and Nesbakken (2004) and Matsumoto (2016b). Household electricity consumption depends on both appliance ownership and how appliances are used once they are owned (see Figure 1). Merely regressing electricity consumption on appliance ownership variables would yield biased results. While we could control for factors influencing appliance use in the demand equation[33], doing so would result in biased and inconsistent estimates due to the simultaneous nature of consumption decisions and appliance ownership(Dubin & McFadden, 1984). Addressing this simultaneity problem requires identifying instrumental variables that influence the purchase decision but not the usage decision. However, finding a valid instrument for each appliance is impractical given the diversity of appliances. The conditional demand model offers a solution by allowing us to estimate appliance-specific consumption for the average household while accounting for appliance use.

5.1 The conditional demand model

Assume a household can own $\ell \in L$ different type of appliances. We follow the hurdle model explained in our empirical section, in which household first invest in each appliance and then decide the amount of units. Let D_i^ℓ be a dummy which takes value 1 if household i owns appliances ℓ, and let $K^\ell > 0$ be the number of units of that appliance owned by the household. We assign each household owning $K^\ell > 0$ units of appliances ℓ to group K^ℓ and we estimate the intensity of the use of appliance ℓ within group K^ℓ. For this, assume that electricity consumption for the k^{th} appliance ℓ for household i is observed through direct metering. The appliance-usage equation is then

$$y_{ik}^\ell = \alpha_\ell + \sum_{m=1}^{M} \gamma_{\ell,m}(C_{i,m} - \bar{C}_{K^\ell,m}) + \varepsilon_{ik}^\ell \tag{5}$$

where the parameter α_ℓ measures the electricity required for an appliance of type ℓ for the mean household, and ε_{ik}^ℓ is an independent and identically distributed error term. The parameter $\gamma_{\ell,m}$ measures the effect of the m^{th} observable characteristic $C_{i,m}$ on the use of appliance ℓ. This variable can be the household socioeconomic characteristics as well as other factors. In this model, $\bar{C}_{K^\ell,m}$ is the mean characteristic for household in group K^ℓ. Therefore, the second term is the adjustment to appliance consumption due to usage on account of other variables. This equation enables us to investigate, for instance, whether high-income households utilize each appliance ℓ more intensively than their low-income counterparts and whether households in areas with low reliability use certain appliances less intensively than those in areas with good reliability.

Given that each household owns K_i^ℓ units of the appliance, we assume each unit has the same energy requirements, and the effect of household characteristics on appliance usage is the same for all K_i^ℓ units. Therefore, the total electricity consumption of appliance ℓ is

$$y_i^\ell = y_{ij}^\ell \cdot K_i^\ell = \alpha_\ell \cdot K_i^\ell + \sum_{m=1}^{M} \gamma_{\ell,m}(C_{i,m} - \bar{C}_{K^\ell,m}) \cdot K_i^\ell + \omega_i^\ell \tag{6}$$

where $\omega_i^\ell = K_i^\ell \cdot \varepsilon_{ik}^\ell$. Given that there are L varieties of appliances, the total electricity consumption of household i becomes

[33]Several factors might affect households' ability to use appliances, including, but not limited to, service characteristics (Blimpo & Cosgrove-Davies, 2019) and household characteristics (Matsumoto, 2016b).

$$y_i = \sum_{\ell=1}^{L} y_i^\ell \cdot D_i^\ell = \sum_{\ell=1}^{L} \alpha_\ell \cdot (K_i^\ell \cdot D_i^\ell) + \sum_{\ell=1}^{L}\sum_{m=1}^{M} \gamma_{\ell,m}(C_{i,m} - \bar{C}_{K^\ell,m}) \cdot (K_i^\ell \cdot D_i^\ell) + \mu_i \tag{7}$$

where $\mu_i = \tau + \omega_i^\ell \cdot D_i^\ell$, and τ is the consumption due to unobserved appliances. Since all the variables in Equation 7 are observed we can estimate it by least squares.

In this model, the parameters of interest are α_ℓ and $\gamma_{\ell,m}$. The parameter α_ℓ represents the electricity consumption associated to one unit of appliance ℓ for the mean household. That is, this variable measures how much electricity of a unit of appliance ℓ is expected to consume at the mean household. On the other hand, the parameter $\gamma_{\ell,m}$ are the deviations in consumption from the mean due to usage differences across households. In other words, this method allows us to explain intensity of appliance usage in terms of *variations* in the different household-level characteristics, for example, income and reliability. Hence, we can also estimate how appliance use is expected to change due to reliability changes. Finally, these estimates can be used to calculate the expected electricity consumption from K_i^ℓ units of appliances ℓ for any household i as

$$E\left[y_i^\ell | K_i^\ell, C_{i,m}\right] = \begin{cases} 0 & \text{if } K_i^\ell = 0 \\ K_i^\ell \left(\hat{\alpha}_\ell + \sum_{m=1}^{M} \hat{\gamma}_{\ell,m}(C_{i,m} - \bar{C}_{K^\ell,m})\right) & \text{if } K_i^\ell > 0 \end{cases} \tag{8}$$

Equation 8 provides a framework to estimate residential electricity consumption for various appliances, considering household-specific characteristics. We leverage this equation to quantify our empirical findings from the preceding section and draw policy implications.

5.1.1 Electricity Consumption Data

To estimate the conditional demand model we use the EICV data explained in section 3. However, the EICV data presents significant challenges for studying electricity consumption. Firstly, the data does not directly provide household electricity consumption in kilowatt-hours (kWh); instead, it reports monthly electricity expenditure. We converted expenditure values into consumption quantities for each household using the country's tariff, as explained in detail in Appendix 3. Secondly, the data is susceptible to misreporting and measurement errors, impacting inference and potentially introducing bias (Bruckmeier, Riphahn, & Wiemers, 2019; Meyer, Román-Palacios, & Wiens, 2018). To address these issues, we conducted a validation procedure using proprietary data from the Rwanda Energy Group (REG). The REG dataset comprises customer prepaid electricity transaction[34] data collected from 2013 to 2019 for 777,023 unique meter IDs, covering almost 1 million meters installed between 1996 and 2020. Access to this data was obtained through a data sharing agreement with the Rwanda Energy Group[35].

[34]In Rwanda, a majority of electricity purchases in Rwanda and carried out using the prepaid electricity framework. Under the prepaid framework, customers purchase electricity units through a mobile telecom network typically using a mobile or web application if purchasing through the internet, otherwise, USSD quick codes can be used for offline customers (Mwaura, 2012). REG maintains a record of each customer transaction with corresponding customer details such as the customer's name, consumer category, transaction timestamps, corresponding taxes and fees, customer's meter location details such as; administrative district, GPS coordinates. Our data includes 85% residential households while the remainder are non-residential (it does not include large industrial meters). The data for residential consumers is used to validate the reported consumption data as explained below.

[35]This agreement was signed by e-Guide which is a collaboration between engineering research groups at 5 Universities. The following e-GUIDE professors have provided as access to this data: Vijay Modi (Columbia

In order to validate the EICV consumption data, we first matched the survey data and REG data sets to quantify the discrepancy between a respondent's reported electricity expenditure and their actual electricity consumption. Lastly, we studying the relationship of this measurement error and different covariates at the household level. Under the classical measurement errors assumption, a noise measure of the dependent variable will increase the noise of the residual but won't bias the estimates. However, measurement error in the dependent variable that is correlated with the dependent variables (non-classical measurement error) usually does lead to biased estimates[36] (see Bound, Brown, Duncan, and Rodgers (n.d.) for a general framework on the topic). The details of this validation procedure is presented in the Appendix 3 together with the regression results. Our analysis concludes that the variance of the error term, which affects our inference, is expected to increase significantly as the variance of the measurement error is 447.60 KWh. Moreover, we find that the measurement error is not random and correlated to ownership status and the number of laundry machines, phones, satellite dishes and printers owned by the home.

In this context, there is a trade-off between the EICV data and the administrative data. The reported data present measurement error which will affect our hypothesis tests and might lead to bias estimates. On the other hand, the administrative data does not suffer from measurement error but there is sample selection bias since we are able to match a sub-sample of the data. One solution to the non-classical measurement error problem in the survey data would be to use the administrative data for those household we were able to match between dataset. Unfortunately, only 693 households are matched, and hence, we are concern by sample selection bias[37] Appendix 3 also presents the regressions for the sample selection analysis. Unfortunately, selection is not random under our matching process. Consequently, we estimate the conditional demand using analysis on both data sets and compare the results. In addition, we conduct a two-step Heckman selection correction to our model when using the administrative data.

5.1.2 Empirical Estimates

Table 8 presents the results of our residential electricity consumption analysis. Models 1, 2 and 3 use the reported consumption data in EICV, while model 4 uses the administrative data[38]. Model 1 encompasses solely appliance ownership variables, while models 2 and 3 incorporate usage drivers. Our favored choice is model 3. Robust standard errors are enclosed in parentheses.

University), Nathan Williams (RIT), Barry Rawn (CMU-Africa) and Jay Taneja (UMass Amherst). We are thankful for their support and collaboration.

[36]The bias, in more general cases, can always be thought in terms of regression coefficients from regressing the measurement errors on the mis-measured covariates.

[37]If the sample is truncated in a nonrandom way, then OLS suffers from selection bias.

[38]Most of the coefficients are robust when comparing between the administrative data and the reported consumption which can suggest that the attenuation bias is not a big concern. Yet, there are some important differences across the models which can be explained by the different samples we use to estimate the models.

Table 8: Electricity Consumption Analysis

	Reported Consumption (KWh/month)			Administrative
	Model 1	Model 2	Model 3	Data
Communication				
Radio	-0.179 (0.466)	0.257 (0.413)	0.130 (0.395)	-2.071* (1.152)
Analogue phones	1.655*** (0.374)	0.637** (0.271)	0.595** (0.256)	0.656 (0.693)
# Outage freq. (number/day)		-0.003** (0.001)	-0.003** (0.001)	0.0002 (0.003)
# Expenses (log RWF)		0.620 (0.435)	0.792*(0.437)	2.337** (0.916)
# Children (number)			-0.186 (0.138)	-0.348 (0.342)
Smart phones	2.60*** (0.456)	1.897*** (0.407)	2.148*** (0.383)	0.810 (0.934)
# Outage freq. (number/day)		-0.003* (0.002)	-0.004*** (0.002)	-0.003 (0.003)
# Expenditure (log RWF)		1.400*** (0.422)	1.390*** (0.448)	1.753* (0.983)
# Children (number)			0.178 (0.189)	0.301 (0.487)
Entertainment				
TV	6.172*** (0.665)	5.460*** (0.838)	5.209*** (0.829)	6.184** (2.703)
# Outage freq. (number/day)		0.003 (0.004)	0.002 (0.003)	-0.001 (0.009)
# Expenditure (log RWF)		2.460** (1.004)	2.144** (0.985)	0.758 (3.013)
# Children (number)			-0.328 (0.348)	-0.568 (1.733)
# Seniors (number)			4.005*** (1.501)	5.875** (2.955)
Music system	0.633 (2.591)	0.964 (2.092)	1.022 (2.019)	19.176 (11.995)
Camera	1.098 (4.007)	-1.150 (3.609)	-0.965 (3.618)	-11.917* (6.321)
Productivity				
Computer	4.827*** (1.528)	2.445 (1.534)	2.151 (1.417)	12.549*** (2.471)
# Outage freq. (number/day)			0.014* (0.007)	0.006 (0.015)
# Expenditure (log RWF)		1.022 (1.619)	2.069 (1.772)	6.438* (3.286)
# Members (number)			-0.607 (0.760)	-3.938*** (1.170)
# Children (number)			0.676 (1.329)	0.805 (1.855)
Sewing Machine	-0.451 (0.568)	-0.472 (0.520)	-0.589 (0.544)	3.633 (3.661)
Convenience				
Hotplate	24.894*** (7.181)	16.444*** (6.280)	15.601** (6.248)	-5.248 (11.891)
# Outage freq. (number/day)		-0.023 (0.097)	-0.017 (0.092)	0.151 (0.200)
# Expenditure (log RWF)		28.147** (11.555)	29.402** (12.593)	28.395 (26.385)
# Members (number)			-2.243 (3.425)	-8.282 (8.582)
# Females (number)			0.438 (5.911)	19.179 (16.112)
Cooker	3.510* (2.039)	1.945 (1.721)	2.023 (1.822)	0.729 (2.732)
Fridge	20.419*** (3.186)	19.582*** (3.213)	19.585*** (3.180)	23.665*** (4.885)
# Outage freq. (number/day)		0.057 (0.039)	0.047 (0.035)	-0.003 (0.029)
# Expenditure (log RWF)		10.658*** (2.716)	11.079*** (2.607)	9.444 (5.926)
Laundry machine	57.440*** (20.267)	38.404** (15.563)	36.675** (15.781)	5.409 (7.876)
# Outage freq. (number/day)		-0.202 (0.411)	-0.221 (0.411)	-2.463*** (0.802)
# Expenditure (log RWF)		43.679* (23.832)	43.082* (24.529)	224.090*** (82.272)
Water Filter	-0.379 (3.000)	0.588 (2.536)	0.774 (2.582)	-0.859 (8.401)
Number of rooms	0.568** (0.232)	0.669*** (0.200)	0.713*** (0.241)	0.540 (0.689)
Constant	3.536*** (0.812)	4.628*** (0.623)	4.363*** (0.776)	1.478 (3.686)
Inverse Mill's Ratio				Y
Observations	2,906	2,906	2,906	496
R^2	0.523	0.603	0.610	0.697
Adjusted R^2	0.520	0.599	0.605	0.672
F Statistic	226.074*** (df = 14; 2891)	161.962*** (df = 27; 2878)	124.684*** (df = 36; 2869)	28.416*** (df = 37; 458)

Note: White's robust standard error in parenthesis. For models 3 and 4, we control for the inverse Mill's Ratio to account for sample selection. In our first step, we regress a dummy which takes value 1 if the household survived the matching procedure on a set of variables which predict selection. These variables are number of members, gender of the head of household, number of children, a dummy which takes value of 1 if the household is rural, number of years in the dwelling, a dummy which takes value 1 if there are multiple households in multiple houses, total savings, distance to major towns, and dummies for the district.
*p<0.1; **p<0.05; ***p<0.01

In model 3, we observe positive coefficients for appliance ownership, with the exception of sewing machine and camera. However, model 1, which does not account for intensity of usage, shows negative coefficients for water filters and radios. It's noteworthy that model 4 also indicates a negative coefficient for radios. The results highlight significant electricity consumption for certain appliances, notably hotplates, fridges, and laundry machines. Specifically, our findings suggest that, on average, households consume 16.44 kWh per month on hotplates, 19.58 kWh on fridges, and 38.40 kWh on laundry machines. Comparatively, these estimates for fridges are lower than reported figures for other countries; for instance, research in Japan indicates consumption ranging from 49.33 to 72.08 kWh per month for fridge usage Matsumoto (2016b). Similarly, a study in Ghana reports an average consumption of 31 kWh for refrigerators (Sakah, du Can, Diawuo, Sedzro, & Kuhn, 2019b). Ghana and Japan are admittedly much wealthier than Rwanda.

On average, a household consumes 5.46 kWh of electricity per month from using a television. Notably, smart phones are associated with higher electricity consumption compared to analogue phones. Specifically, a typical household consumes 1.90 kWh per month for a smart phone and 0.64 kWh per month for an analogue phone. This discrepancy suggests that smart phones require more frequent recharging due to their higher energy demands. It's worth mentioning that the consumption from the remaining appliances is not statistically significant, indicating that only a few households utilize these appliances.

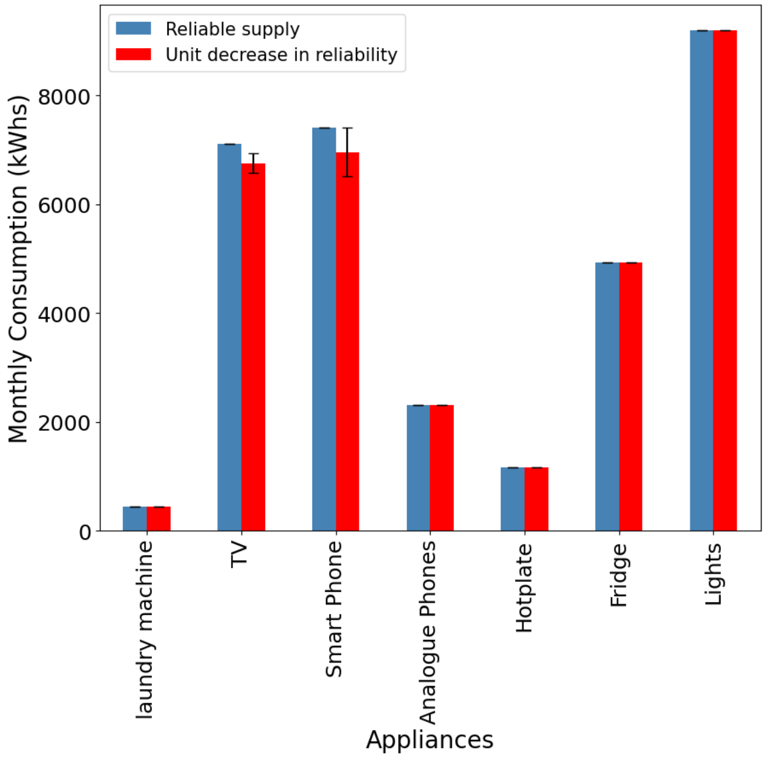
Models 2 and 3 in Table 8 underscore the significance of usage variations among households in determining residential electricity consumption. The adjusted R^2 for these models is higher compared to that of model 1, indicating that incorporating controls for appliance use enhances the explanatory capability of the empirical model. Given the constraints of space in this paper, we concentrate on two variables—household income and reliability. Our findings suggest that the impact of reliability on appliance use is negligible. Specifically, model 3 reveals that monthly electricity consumption from both analogue and smart phones decreases by 0.003 kWh and 0.004 kWh, respectively, for each additional outage.

While the role of reliability on appliance use is negligible, model 3 in table 8 shows that higher-income households use televisions and phones more intensively. This finding could be attributed to two possible explanations. Firstly, high-income households may have the financial capacity to purchase more electricity at a given tariff level compared to low-income households. The second explanation is straightforward: high-income households may allocate their time differently than low-income households. Notably, the income effect is more pronounced for smart-phones than for analogue phones, with smart phones requiring higher electricity consumption and being more significantly impacted by additional income.

The impact of income is notably significant for convenience appliances. Model 3 reveals that a typical household consumes 15.601 kWh of electricity per month from a hotplate, and consumption increases by 29.40 kWh for each additional unit of expenditure. Additionally, higher income is associated with increased electricity consumption from fridges and washing machines, possibly because lower income households may struggle to afford the electricity required to operate such appliances and use them sparingly.

5.2 Policy Discussion

Figure 13: Aggregate household consumption



Note: The above figure depicts the consequences of a one-unit reduction in reliability on aggregate household consumption across all grid connected surveyed households. The combined household consumption is presented at the appliance level. The blue bars represent consumption levels under a reliable electricity supply, while the red bars depict consumption levels following a one-unit decrease in reliability.

In figure 13 household consumption at the appliance level is presented for all grid-connected households. Notably, the top three consuming appliances are lights[39], TVs, and smartphones, with ownership rates of 36.7%, 54.8%, for TVs and smartphones, respectively(refer to Table 2). The substantial ownership rates of TVs and smartphones, coupled with their average consumption levels (Table 8), explain their significant shares relative to other household appliances. Conversely, despite an 82% ownership rate, analogue phones contribute minimally to aggregated appliance-level household consumption due to their low consumption levels. Although laundry machines, hotplates, and fridges are classified as demand-intensive appliances, their consumption levels remain minimal owing to their low levels of household ownership.

From Figure 10 in the Results section, we observed that a decrease in reliability led to reduced ownership rates for TVs and smartphones. Given that these two appliances, along with lights, are the highest contributors to aggregate household-level consumption (Figure 13), a decrease in their ownership results in an average decrease in aggregate household consumption of approximately 800 kWh. According to Joel Mugyenyi and Modi (2022), households in Rwanda consumed an average of 22 kWh per month in 2016-2017. As such, a unit decrease in reliability results in a reduction in aggregate consumption equivalent to

[39]Please note that number of rooms in a household is used as a proxy to determine the number of lights in the household.

the aggregate monthly consumption of 36 households. In concise terms, a unit decrease in reliability would have the equivalent impact as the loss of consumption for 1% of the grid-connected households surveyed in EICV2016/2017.

Hence, it is reasonable to conclude that investments in grid reliability alone may have a limited impact on enhancing household consumption levels. If policymakers aim to boost household consumption, a more effective strategy would involve increasing ownership rates of demand-intensive appliances, such as fridges, laundry machines, and hotplates. Making these appliances more accessible, perhaps through credit schemes, and reducing the operational costs through more affordable tariffs could significantly contribute to elevating household consumption levels.

6 Conclusions

This paper delves into the impact of electricity reliability on household appliance ownership in Rwanda, aiming to overcome the acknowledged barrier of low appliance ownership in Sub-Saharan Africa. Our model scrutinizes both the total number of appliances owned by households and the ownership of key appliances, leveraging a distinctive dataset and employing instrumental variables—specifically, lightning frequency and radiance. The study capitalizes on rare access to administrative reliability data linked to house locations, providing a unique opportunity to explore how reliability influences appliance ownership and usage. However, investigating the role of reliability on household outcomes entails navigating empirical challenges stemming from endogeneity and measurement error. Our results underscore that reliability shapes the composition of the appliance stock owned by households in Rwanda. Households exhibit a demand for energy services and adapt to low grid quality by adjusting the types of appliances they own. Notably, the likelihood of owning smart-phones, decoders, and televisions diminishes in low-reliability areas. In such settings, we observe a higher probability of owning music systems for higher income households and sewing machines for lower income households.

We also note that, while reliability has a modest impact on the ownership of certain appliances, it doesn't influence the usage of appliances for households that already own them. The predominant factor affecting electricity consumption is income, emphasizing that the primary barrier to higher consumption levels is affordability rather than grid reliability. To achieve higher levels of residential consumption on the grid, efforts should be directed towards making electricity more affordable for the average household.

It's worth highlighting that reliability does not significantly affect the ownership of high-consumption appliances like washing machines and fridges. However, it does diminish the ownership of lower-consumption and less expensive appliances such as televisions. Therefore, a strategy aimed at enhancing grid reliability may have only a modest impact on household electricity consumption. A more effective strategy to boost household electricity consumption would involve making appliances more affordable, possibly through improved access to credit, while simultaneously making electricity more affordable, perhaps through targeted subsidies. This dual approach addresses both sides of the affordability equation and holds promise for fostering increased electricity usage among households.

This paper studies the short-run effects of low reliability on residential consumption. However, we do not study the long-run effects of reliability. Low reliability is expected to affect household's ability to engage in productive activities which would increase their income in the medium and long-run. Moreover, reliability is expected to affect household location, and it will have effects on migration patterns within the country. Finally, we do not study the role of reliability on commercial and industrial electricity consumption. We leave these questions for future research.

References

Adesina, A., Akinwumi abd Amin, Birol, F., Bitature, P., Blair, T., Cheikhrouhou, H., Dayal, A., et al. (2020). *Global comission to end energy poverty: 2020 report electricity access* (Tech. Rep.). Global Commission to End Energy Poverty.

Atkeson, A., & Kehoe, P. J. (1999). Models of energy use: Putty-putty versus putty-clay. *American Economic Review*, *89*(4), 1028–1043.

Auffhammer, M., & Wolfram, C. D. (2014). Powering up china: Income distributions and residential electricity consumption. *American Economic Review*, *104*(5), 575–580.

Blakeslee, R. (1998). Lightning imaging sensor (lis). *Dataset available online from the NASA EOSDIS Global Hydrology Resource Center Distributed Active Archive Center, Huntsville, Alabama, USA doi: http://dx. doi. org/10.5067/LIS/LIS/DATA201*.

Blimpo, M., & Cosgrove-Davies, M. (2019). *Electricity access in sub-saharan africa: Uptake, reliability, and complementary factors for economic impact*. World Bank Publications.

Bos, K., Chaplin, D., & Mamun, A. (2018). Benefits and challenges of expanding grid electricity in africa: A review of rigorous evidence on household impacts in developing countries. *Energy for sustainable development*, *44*, 64–77.

Bound, J., Brown, C., Duncan, G. J., & Rodgers, W. L. (n.d.). *Measurement error in cross-sectional and longitudinal labor market surveys: Results from two validation studies* (Working Paper). National Bureau of Economic Research. doi: 10.3386/w2884

Bruckmeier, K., Riphahn, R. T., & Wiemers, J. (2019). *Benefit underreporting in survey data and its consequences for measuring non-take-up: new evidence from linked administrative and survey data* (Tech. Rep.). IAB-Discussion Paper.

Burgess, R., Greenstone, M., Ryan, N., & Sudarshan, A. (2020). The consequences of treating electricity as a right. *Journal of Economic Perspectives*, *34*(1), 145–169.

Cabeza, L. F., Ürge-Vorsatz, D., Ürge, D., Palacios, A., & Barreneche, C. (2018). Household appliances penetration and ownership trends in residential buildings. *Renewable and Sustainable Energy Reviews*, *98*, 1–8.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The review of economic studies*, *47*(1), 225–238.

Chen, H., Jin, L., Wang, M., Guo, L., & Wu, J. (2023). How will power outages affect the national economic growth: Evidence from 152 countries. *Energy Economics*, *126*, 107055.

Dang, D. A., & La, H. A. (2019). Does electricity reliability matter? evidence from rural viet nam. *Energy Policy*, *131*, 399–409.

Day, E. (2020). *Moving beyond energy access–the challenge and impact of unreliable electricity in emerging economies* (Tech. Rep.). Technical Report, Energy and Economic Growth Energy Insight.

Debnath, R., Bardhan, R., & Sunikka-Blank, M. (2019). How does slum rehabilitation influence appliance ownership? a structural model of non-income drivers. *Energy Policy*, *132*, 418–428.

Dhanaraj, S., Mahambare, V., & Munjal, P. (2018a). From income to household welfare: lessons from refrigerator ownership in india. *Journal of Quantitative Economics*, *16*, 573–588.

Dhanaraj, S., Mahambare, V., & Munjal, P. (2018b). From income to household welfare: lessons from refrigerator ownership in india. *Journal of Quantitative Economics*, *16*, 573–588.

Dubin, J., & McFadden, D. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, *52*.

Farrell, M. J. (1954). The demand for motor-cars in the united states. *Journal of the Royal*

Statistical Society. Series A (General), *117*(2), 171–201.

Feng, C. X. (2021). A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. *Journal of Statistical Distributions and Applications*, *8*(8).

Gertler, P. J., Shelef, O., Wolfram, C. D., & Fuchs, A. (2016). The demand for energy-using assets among the world's rising middle classes. *American Economic Review*, *106*(6), 1366–1401.

Gunther, R. (2023). *Illuminating the dark side: Lightning's assault on power networks.* Retrieved from https://clouglobal.com/illuminating-the-dark-side-lightnings-assault-on-power-networks/#:~:text=Lightning%20strikes%20can%20induce%20high%20electromagnetic%20fields%20that,data%20transmission%2C%20affecting%20grid%20monitoring%20and%20control%20operations. (Accessed: 2023-11-29)

Hanna, R., & Oliva, P. (2015). Moving up the energy ladder: the effect of an increase in economic well-being on the fuel consumption choices of the poor in india. *American Economic Review*, *105*(5), 242–246.

Hashemi, M. (2022). How would residential electricity consumers respond to reductions in power outages? *Energy for Sustainable Development*, *69*, 1–10.

Hausman, J. A., Hall, B. H., & Griliches, Z. (1984). *Econometric models for count data with an application to the patents-r&d relationship.* national bureau of economic research Cambridge, Mass., USA.

Hilbe, J. M. (2014). *Modeling count data*. Cambridge University Press.

Himarios, D. (2000). How forward looking are consumers? further evidence for the united states. *Southern Economic Journal*, *66*(4), 991–1000.

IEA. (2022). *Africa energy outlook 2022.* Retrieved from https://www.iea.org/reports/africa-energy-outlook-2022 (Accessed: 2023-11-28)

Izabiliza, J. (2003). The role of women in reconstruction: experience of rwanda. *Source unknown*.

Joel Mugyenyi, S. F. C. M. J. T. N. J. W., Bob Muhwezi, & Modi, V. (2022). Analysis of electricity consumption tariff changes in rwanda.

Khandker, S. R., Barnes, D. F., & Samad, H. A. (2009). Welfare impacts of rural electrification: a case study from bangladesh. *World Bank Policy Research Working Paper*(4859).

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political analysis*, *9*(2), 137–163.

Lenz, L., Munyehirwe, A., Peters, J., & Sievert, M. (2017). Does large-scale infrastructure investment alleviate poverty? impacts of rwanda's electricity access roll-out program. *World Development*, *89*, 88–110.

Matsumoto, S. (2016a). Household income structure and electrical appliance ownership: Evidence from japanese national household survey. *International Journal of Energy Economics and Policy*, *6*(1), 14–19.

Matsumoto, S. (2016b). How do household characteristics affect appliance usage? application of conditional demand analysis to japanese household data. *Energy Policy*, *94*, 214–223.

McNeil, M. A., & Letschert, V. E. (2010). Modeling diffusion of electrical appliances in the residential sector. *Energy and Buildings*, *42*(6), 783–790.

McRae, S. (2010). Reliability, appliance choice, and electricity demand. *Department of Economics, Ed., ed. University of Michigan*.

Meeks, R. C., Omuraliev, A., Isaev, R., & Wang, Z. (2023). Impacts of electricity quality improvements: Experimental evidence on infrastructure investments. *Journal of Environmental Economics and Management*, *120*, 102838. Retrieved from https://www.sciencedirect.com/science/article/pii/S0095069623000566 doi:

https://doi.org/10.1016/j.jeem.2023.102838
Meyer, A. L., Román-Palacios, C., & Wiens, J. J. (2018). Bamm gives misleading rate estimates in simulated and empirical datasets. *Evolution*, *72*(10), 2257–2266.
Muza, O., & Debnath, R. (2021). Disruptive innovation for inclusive renewable policy in sub-saharan africa: A social shaping of technology analysis of appliance uptake in rwanda. *Renewable energy*, *168*, 896–912.
Mwaura, F. M. (2012). Adopting electricity prepayment billing system to reduce non-technical energy losses in uganda: Lesson from rwanda. *Utilities Policy*, *23*, 72–79.
Nielsen, L. (1993). How to get the birds in the bush into your hand: results from a danish research project on electricity savings. *Energy policy*, *21*(11), 1133–1144.
Oceanic, N. N., & Administration), A. (2020). *5 striking facts versus myths about lightning you should know.* Retrieved from https://www.noaa.gov/stories/5-striking-facts-versus-myths-about-lightning-you-should-know (Accessed: 2023-11-29)
O'Doherty, J., Lyons, S., & Tol, R. S. (2008). Energy-using appliances and energy-saving features: Determinants of ownership in ireland. *Applied energy*, *85*(7), 650–662.
Pawar, S., & Jha, A. K. (2023). Analysis of residential location choices of different socio-economic groups and their impact on the density in a city using agent based modelling. *Applied Spatial Analysis and Policy*, *16*(1), 119–139.
Ramakrishnan, A., Kalkuhl, M., Ahmad, S., & Creutzig, F. (2020). Keeping up with the patels: Conspicuous consumption drives the adoption of cars and appliances in india. *Energy Research & Social Science*, *70*, 101742.
Rao, N. D., & Ummel, K. (2017). White goods for white people? drivers of electric appliance growth in emerging economies. *Energy Research & Social Science*, *27*, 106–116.
REG. (2020). *Reticulation standards for electricity distribution planning, construction and maintenance.*
REG. (2021). *Rwanda electricity distribution master plan.* Retrieved from https://www.reg.rw/fileadmin/user_upload/Rwanda_Electricity_Distribution_Master_Plan.PDF (Accessed: 2023-04-14)
Rezinkina, M., Babak, V., Gryb, O., Zaporozhets, A., & Rezinkin, O. (2022). Increasing the reliability of lightning protection of electric power facilities. , 281–317.
Richmond, J., & Urpelainen, J. (2019). Electrification and appliance ownership over time: Evidence from rural india. *Energy Policy*, *133*, 110862.
Sakah, M., du Can, S. d. l. R., Diawuo, F. A., Sedzro, M. D., & Kuhn, C. (2019a). A study of appliance ownership and electricity consumption determinants in urban ghanaian households. *Sustainable Cities and Society*, *44*, 559–581.
Sakah, M., du Can, S. d. l. R., Diawuo, F. A., Sedzro, M. D., & Kuhn, C. (2019b). A study of appliance ownership and electricity consumption determinants in urban ghanaian households. *Sustainable Cities and Society*, *44*, 559–581.
Sally Smith, M. A., Richard Anker, & Prates, I. (2020). *Anker living income reference value.* Retrieved from https://www.globallivingwage.org/wp-content/uploads/2020/08/Rural-Rwanda-LI-Reference-Value.pdf
Sievert, M., & Steinbuks, J. (2020). Willingness to pay for electricity access in extreme poverty: Evidence from sub-saharan africa. *World Development*, *128*, 104859.
Sinha, P., Caulkins, M. L., & Cropper, M. L. (2018). Household location decisions and the value of climate amenities. *Journal of Environmental Economics and Management*, *92*, 608–637.
Wolfram, C., Shelef, O., & Gertler, P. (2012). How will energy demand develop in the developing world? *Journal of Economic Perspectives*, *26*(1), 119–138.

Appendix 1: Data Description and Cleaning

Integrated Household Living Conditions Survey (EICV)

- Expenditure

  The expenditure variable was constructed by consolidating reported expenditure items from surveyed households. Participants were queried about their spending over the past year on durable items like clothes, accessories, furniture, and school-related expenses (uniforms, supplies, and tuition). Monthly expenditures on transport, leisure, health, beauty care, communication, and housing (including rent and electricity) were also elicited. To ensure consistency, all expenditure-related responses were transformed into a monthly cadence and aggregated into a single variable, offering an overview of each household's expenditure level

- Demographics

  Several variables from the survey were employed to characterize the demographics of households. These include gender, encompassing the sex of the household head and the overall gender makeup; the highest level of education attained by the household head and members; the age composition of household members (children under 16 and the elderly above 60); details about the dwelling, such as the number of rooms and construction material; and a variable denoting the nationality of the household head.

- Finances

  Binary variables were created to identify households involved in various financial activities, including ownership of businesses, employment status of members, presence of debt, savings in a bank account, and receipt of money transfers from friends or family. Another variable categorizes households engaged in either large-scale or small-scale agriculture. The skill level of household members is determined using reported occupation responses and classified into low, medium, or high skill levels according to the Intermation Standard Classification of Occupations ISCO-08. Additionally, the stability of a respondent's job is inferred from the number of jobs held in the recent past.

- Appliance ownership

  Households were surveyed about ownership of common household appliances, and binary variables were created to indicate ownership of each appliance within the household. This approach allows for a detailed understanding of the possession of specific household items.

Reliability Data

Reliability and grid infrastructure geo-spatial datasets were generously provided by the Rwanda Energy Group (REG). The reliability data encompasses reported outages, indicating both the time of occurrence and resolution. Additionally, it includes information on feeder outage causes, reported at the feeder level—the finest resolution on the grid. Each feeder is uniquely identified by a name and originates from a substation. The dataset spans five years, from 2016 to 2020, and is aggregated into a cohesive dataset using feeder names and origin substations to track feeders over time.

The geo-spatial dataset contains feeder names, corresponding origin substations, and a linestring illustrating the spatial extent of each feeder. A linestring is a geometric object that represents a sequence of connected line segments. Through our collaboration with the National Institute of Statistics in Rwanda (NISR), we obtained permission to integrate REG data with the EICV household survey data. The matching process required our physical presence in Rwanda due to the sensitive nature of the data, and it was conducted on NISR internal computers.

For the matching technique, we adopted a straightforward approach. Household reliability was determined based on a household's proximity to a feeder. Each household was assigned the reliability value of the nearest feeder, subject to a constraint: the household had to be within an 800-meter proximity to the feeder. This constraint aligns with REG installation requirements, as households more than 800 meters away from the closest transformer cannot be connected to the grid. This careful consideration helps minimize the probability of erroneously assigning households to distant feeder lines, especially in cases where certain grid lines may not be captured within our dataset

Rainfall Data and Spatial Interpolation

The Rwanda Meteorological Agency maintains a comprehensive historical dataset of daily rainfall collected at 18 rainfall stations distributed across Rwanda, as depicted in 5 (a). Our study has access to a 40-year daily historical record of rainfall at each station. However, the agency acknowledges occasional gaps in its daily collection of station rainfall data, which are addressed by filling them with derived outputs from satellite images collected at resolutions of 4km by 4km.

To interpolate and estimate rainfall across the entire country, we employ Kriging, a geostatistical technique known for spatial interpolation, prediction, and estimation, as depicted in 5 (a). In this process, the values of rainfall at unobserved locations are estimated using recorded rainfall data from the designated rainfall stations. Kriging involves mathematical modeling of the spatial correlation structure through a variogram, which quantifies the spatial variability of the data. The variogram is then utilized to optimize the weights assigned to observed points when predicting values at unsampled locations.

For the implementation of Kriging, we utilize pykrige, a specialized Python library for Kriging. This library offers various variogram models such as linear, Gaussian, exponential, and power. After visual inspection, we selected a "power" model for our sample as it provided the best fit to the observed data.

Lightning Data

The Lightning Imaging Sensor (LIS) is an advanced space-based lightning detection instrument situated aboard NASA's Earth Observing System (EOS) Tropical Rainfall Measuring Mission (TRMM) satellite. Designed to operate seamlessly in both day and night conditions, the LIS records the precise time of lightning occurrences, measures radiant energy, and estimates their locations with exceptional efficiency. As the TRMM satellite speeds through space at an astonishing 7 kilometers per second (nearly 16,000 miles per hour), it provides LIS with a unique vantage point, allowing for observations of a specific point on Earth or a cloud for approximately 90 seconds during each overhead pass.

Despite the relatively brief observation duration, this timeframe is sufficiently long for LIS to accurately estimate the flashing rate of most storms. The instrument's capabilities encompass recording the time of occurrence, measuring radiant energy, and determining the precise location of lightning events within its expansive field-of-view. Notably, the TRMM LIS detection efficiency demonstrates a range from 69% near noon to 88% at night.

The TRMM Lightning Imaging Sensor (LIS) dataset, collected by the LIS instrument on the TRMM satellite, serves as a valuable resource for discerning the distribution and variability of total lightning in Earth's tropical and subtropical regions. This dataset finds application in severe storm detection, comprehensive analysis, and investigations into lightning-atmosphere interactions. With its high detection efficiency during both day and night, the LIS instrument has proven instrumental in advancing our understanding of lightning phenomena.(Blakeslee, 1998)

Appendix 2: Regression Tables

- Add once we have the final results

Appendix 3: Validation of Reported Electricity Consumption Data

This section explains how the reported electricity expenditure data in EICV was validate using administrative data on electricity consumption for the Rwanda Energy Group. We first transformed the reported expenditure data into consumption data in KWh. We then matched the reported and administrative data. Finally, we conducted statistical analysis of the misreported values.

Transforming Expenditure Data into Consumption Data

We use the residential electricity tariff structure in Rwanda to transform our Expenditure Data into Consumption Data. In 2016, every household paid a tariff 182 RWF/KWh. For households surveyed in 2016, we dived the reported expenditure by the tariff level at that moment. In January 2017, a block tariff was introduced in which the household paid a tariff of 89 RWF/kWh for the first 15 kWh. That is, for the first 15 kWh, the expenditure can never be above 1,335 RWF. For the next 35 kWh, the household pays a tariff of 182 RWF/kWh (the expenditure for the following 35 units can never be above 6370). Any additional unit above the first 50 kWh pays a tariff of 189 RWF/kWh. In this context, suppose that for household i the expenditure is y_i:

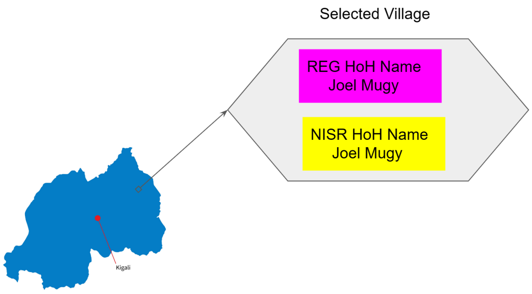
- If $y_i > 7,705$, $y_i - 7,705$ pays a tariff of 189 RWF/kWh and consumption is then 50 + $(y_i - 7,705)/189$
- If $1,335 < y_i < 7,705$, $y_i - 1,335$ pays a tariff of 182 RWF/kWh and consumption is then 15 + $(y_i - 1,335)/182$
- If $y_i < 1,335$, y_i pays a tariff of 182 RWF/kWh and consumption is then $y_i/89$

Matching for Consumption Data Validation

In both the EICV and REG datasets, the names, administrative and GPS locations are recorded for each household. Unfortunately, meters are not located at the house in several cases but on the grid pole. Hence, using GPS location to match the datasets will produce significative errors. Instead, we combine head of households names and GPS data for our matching process in a two-step algorithm.

The smallest administrative unit in Rwanda is a village. Villages common to both datasets are identified and within selected villages, respondents common to both data sets are chosen. In EICV, each household head is identified and we have access to the recorded names. In REG's data, all electricity purchases are attached the the owner of the house who is typically the household head. Therefore, given that the household head's name is listed in both datasets, it can be used as a common identifier to match households in one dataset to another. Yet, these names are not the same. Figure 14 presents an example of how the household names are display in each dataset.

Figure 14: Head of Household Names



In this context, we implement a "fuzzy" string matches to identify similar names in villages common to both data sets. A similarity score is then assigned based on the number of words which are matched. Figure 15 presents an example for the similarity score. We finally consider that names in each data set are a "match" if they have a similarity score higher than 80 implying a majority of the names are similar.

Figure 15: Fuzzy String matching

String one	String Two	Score
YANKEES	NEW YORK	14
NEW YORK METS	NEW YORK MEATS	96

We then use the GPS data to improve the matching algorithm. We calculate the distance between the houses corresponding to the matched names and the meter ID. In many cases, the meter is not installed in the house but in the electricity pole. In Rwanda, customer electricity meters especially in residential settings are typically located at the closest electricity pole which according to REG electricity connection standards should be within a 40 meter distance to the house receiving the electricity connection (REG, 2020). Hence, only matches that are within a 40 meter distance to one another are maintained, and matches outside this threshold are discarded. The reason for this decision is that people in the same village might have the same name, but the probability of two household living within 80 meters one from the other and having the same name is small.

Figure 16: Distance house to meter location

Figure 17 presents the result from our matching algorithm. Out of the 3,600 grid electrified housedholds in the EICV dataset, 693 household are matched to the REG data set. For the matched results, the respondent's consumption recorded in the month the EICV interview was conducted is compared to the respondent's reported electricity expenditure to determine the level of over or under reporting.

Figure 17: Matching Result

Dataset	**Sample Count**
Original EICV5 (2016-2017)	*14580 unique households (3600 connected to Grid)*
REG GPS	*672,561 customers*
EiCV5 - REG GPS match	1880 unique households matched
EICV5 - GPS - Consumption match	964 unique households matched
REG Consumption (2016-2017) (40 m distance threshold)	*693 unique households matched*

In this context, we are worried about sample selection bias. That it, we may have a latent variable that is only observed based on some other condition, which we will call a selection equation given by

$$s_{ij} = 1(z'_{ij}\omega + \eta_{ij} \geq 0) \tag{9}$$

where $s_{ij} = 1$ if the dependent variable s observed by the econometrician. In this case, the bias is given by the Inverse Mills ratio. We present the results for equation (9) below. As we can observe in out data, selection is not random and correlated to household characteristics.

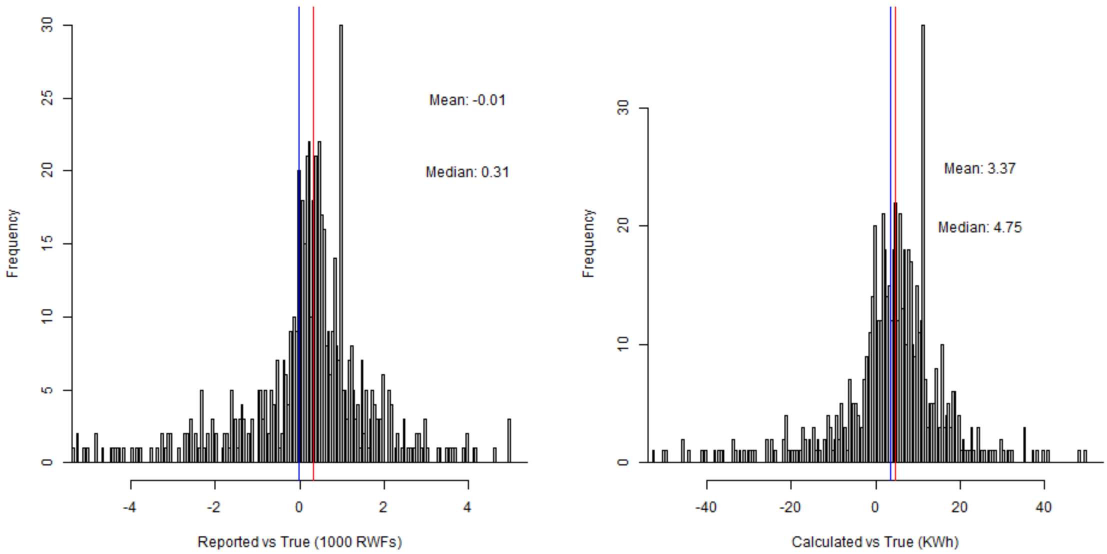
Table 9

	Dependent variable:		
	matched		
	OLS	*logistic*	*normal*
	OLS	Logit	Probit
saidi	0.026**	0.173*	0.026**
	(0.013)	(0.103)	(0.013)
saifi	−0.010	−0.167	−0.010
	(0.016)	(0.121)	(0.016)
log_exp	0.001	−0.306***	0.001
	(0.003)	(0.029)	(0.003)
savings	−0.002	−0.015	−0.002
	(0.003)	(0.021)	(0.003)
business	0.009	0.138	0.009
	(0.014)	(0.113)	(0.014)
high_skill	−0.021	−0.131	−0.021
	(0.021)	(0.175)	(0.021)
job_instability	−0.014	−0.281**	−0.014
	(0.011)	(0.109)	(0.011)
female	0.010	0.058	0.010
	(0.007)	(0.050)	(0.007)
children	−0.016**	−0.078*	−0.016**
	(0.006)	(0.046)	(0.006)
seniors	−0.058***	−0.359***	−0.058***
	(0.018)	(0.122)	(0.018)
hh_female	−0.029	−0.333**	−0.029
	(0.018)	(0.152)	(0.018)
hh_ed	−0.011	−0.083	−0.011
	(0.020)	(0.179)	(0.020)
hh_youth	−0.021	−0.267*	−0.021
	(0.016)	(0.136)	(0.016)
num_rooms	0.005	0.051	0.005
	(0.006)	(0.045)	(0.006)
multi_hhs	0.007	0.072	0.007
	(0.020)	(0.206)	(0.020)
multi_house	0.014	0.221	0.014
	(0.020)	(0.189)	(0.020)
owner	0.295***	2.862***	0.295***
	(0.020)	(0.216)	(0.020)
num_years_in_dwelling	0.002**	0.009	0.002**
	(0.001)	(0.006)	(0.001)
rural	−0.057***	−0.490***	−0.057***
	(0.017)	(0.136)	(0.017)
log_trade	0.005	0.050	0.005
	(0.008)	(0.064)	(0.008)
tv	0.064***	0.557***	0.064***
	(0.021)	(0.154)	(0.021)
phones	0.008	0.067	0.008
	(0.007)	(0.050)	(0.007)
radio	−0.004	−0.012	−0.004
	(0.011)	(0.082)	(0.011)
satellite_dish	−0.030	−0.221	−0.030
	(0.030)	(0.218)	(0.030)
decoder	−0.007	−0.003	−0.007
	(0.021)	(0.155)	(0.021)
filter	−0.064*	−0.404	−0.064*
	(0.034)	(0.264)	(0.034)
laundry	0.103	0.653	0.103
	(0.105)	(0.797)	(0.105)
computer	−0.011	−0.076	−0.011
	(0.015)	(0.113)	(0.015)
printer	0.062	0.671	0.062
	(0.068)	(0.560)	(0.068)
cooker	−0.0001	0.009	−0.0001
	(0.025)	(0.189)	(0.025)
fridge	0.054*	0.416*	0.054*
	(0.032)	(0.241)	(0.032)
hotplate	−0.014	0.248	−0.014
	(0.051)	(0.406)	(0.051)
music	−0.092**	−0.870**	−0.092**
	(0.041)	(0.407)	(0.041)
camera	0.024	0.099	0.024
	(0.038)	(0.304)	(0.038)
Observations	3,072	3,072	3,072
R^2	0.328		
Adjusted R^2	0.321		
Log Likelihood		−1,171.048	−1,163.726
Akaike Inf. Crit.		2,410.096	2,395.451
Residual Std. Error	0.355 (df = 3038)		
F Statistic	43.706*** (df = 34; 3038)		
Note:			*p<0.1; **p<0.05; ***p<0.01

Non-classical Measurement Error

Figure 18 presents a frequency histogram visualizing the measurement error in the survey report and our consumption estimates for 693 matched households.

Figure 18: Measurement error in EICV 5



(a) Monthly expenditure (b) Monthly consumption

Note: The figure shows that the measurement error depicts a normal distribution for both monthly reported electricity expenditure and monthly consumption. In the case of reported electricity expenditure, the mean difference is close to zero showing that on average the respondent's reported electricity expenditure don't suffer from under or over reporting but rather mirror a close approximation of their actual electricity consumption. In the case of the calculated monthly consumption, the mean difference is 3.37 while the median is 4.75. The reason of these larger differences might be due to taxes which are charged by the utility but we do not account in our estimates for electricity consumption.

In order to study the measurement error in our consumption data we regress the estimated measurement error on several observables, including the number of appliances. Below are the results. The results show that the measurement error is correlated to the number of some appliances owned by the household. In addition, household who own their home seem to have larger measurement error. For this reason, we need to be careful about non-classical measurment error.

Table 10: Non-classical measurment error

	Dependent variable: measurement error	
	Expenditure	Consumption
	(1)	(2)
saidi	0.133	0.861
	(0.318)	(1.789)
saifi	0.180	1.943
	(0.365)	(2.052)
log_exp	−0.112	−0.420
	(0.101)	(0.567)
savings	−0.019	−0.077
	(0.050)	(0.281)
business	−0.423	−2.333
	(0.353)	(1.986)
high_skill	0.289	1.443
	(0.581)	(3.264)
job_instability	0.220	1.483
	(0.364)	(2.044)
female	−0.180	−1.227
	(0.163)	(0.919)
children	−0.001	0.176
	(0.151)	(0.848)
seniors	−0.404	−2.204
	(0.403)	(2.268)
hh_female	−0.117	−0.765
	(0.521)	(2.927)
hh_ed	0.145	0.895
	(0.591)	(3.320)
hh_youth	−0.112	0.210
	(0.453)	(2.547)
num_rooms	−0.131	−0.504
	(0.144)	(0.809)
multi_hhs	−0.753	−4.832
	(0.622)	(3.497)
multi_house	−0.164	−1.415
	(0.564)	(3.170)
owner	1.373*	8.450**
	(0.751)	(4.221)
num_years_in_dwelling	0.022	0.140
	(0.021)	(0.120)
rural	0.218	−1.460
	(0.440)	(2.473)
log_trade	−0.016	−0.324
	(0.234)	(1.315)
tv	−0.495	−2.327
	(0.496)	(2.786)
phones	0.377**	1.968**
	(0.161)	(0.905)
radio	0.002	0.119
	(0.258)	(1.451)
satellite_dish	−1.899***	−9.759***
	(0.655)	(3.682)
decoder	0.566	2.480
	(0.501)	(2.814)
filter	−0.267	0.211
	(0.851)	(4.783)
laundry	6.479***	33.228***
	(1.776)	(9.983)
computer	−0.662**	−3.373*
	(0.325)	(1.826)
printer	4.730**	21.532**
	(1.849)	(10.395)
cooker	0.700	4.090
	(0.619)	(3.478)
fridge	−0.141	−2.460
	(0.672)	(3.776)
hotplate	−0.578	−4.189
	(1.226)	(6.890)
music	−0.730	−2.221
	(1.450)	(8.152)
camera	−1.275	−6.987
	(0.957)	(5.377)
Observations	571	571
R^2	0.096	0.110
Adjusted R^2	0.038	0.053
Residual Std. Error (df = 537)	3.847	21.627
F Statistic (df = 34; 537)	1.670**	1.942***
Note:		*p<0.1; **p<0.05; ***p<0.01