ARTICLE

<https://doi.org/10.1057/s41599-021-00848-0>

OPEN

Access to electricity and digital inclusion: evidence from mobile call detail records

Georges V. Hounbonon¹, Erwan Le Quentrec² & Stefania Rubrichi² [✉]

Despite significant progress in mobile connectivity, a large number of individuals in Sub-Saharan Africa are at risk of being left behind, especially those living in rural areas and women. In this paper, we use an original data set of mobile call detail records from Senegal and exploit variations across plausibly similar rural municipalities to assess the impact of access to electricity on mobile connectivity. We find that access to mobile connectivity for rural users, measured by mobile telephony subscriptions or smartphone ownership, increases with access to electricity, with a stronger impact for women. The impact on usage is also positive, but stems from the increased volume of incoming communications: rural mobile users with access to electricity do not place more calls than those with limited access; rather, they receive more calls and text messages from urban areas, especially those with higher living standards. These findings suggest that access to electricity can be a significant driver of digital inclusion in Sub-Saharan Africa.

¹International Finance Corporation (IFC)-World Bank Group, Washington, DC, USA. ²Orange Labs, Sociology and Economics of Networks and Services (SENSE), Châtillon, France. ✉email: stefania.rubrichi@orange.com

Introduction

Despite significant progress in mobile connectivity, a substantial number of individuals in Sub-Saharan Africa (SSA) are at risk of being left behind, especially those living in rural areas and women. The penetration rate of mobile internet in 2018 is estimated at 15% in rural areas, compared to 36% in urban areas, and 15% of women are less likely to own a mobile phone than men, while 41% are less likely to use the internet GSMA (2019). In rural areas, in particular, enabling digital inclusion, that is increasing access to mobile connectivity for the underserved population, remains constrained, among other factors, by limited access to electricity: the rate of access to electricity in 2018 is estimated at 26% for rural areas, against 74% in urban areas International Energy Agency (2018). While several studies have investigated the drivers of mobile connectivity, especially in developing countries, the link with access to electricity tends to be overlooked (Armey and Hosman, 2016), possibly because access to electricity is almost exclusively an SSA problem: in 2017, 70% of the world population without access to electricity live in this region. Another potential reason could be more technical since access to electricity and mobile connectivity may be simultaneously determined. We try to circumvent this endogeneity concern by focusing on the availability of electricity, taking advantage of a new dataset on the power sources of telecom antenna sites.

Access to electricity can affect the availability, uptake, and usage of mobile connectivity. Limited availability of a reliable source of electricity can reduce the availability of mobile connectivity because network operators seeking to roll out base stations in non-electrified areas have to incur higher costs stemming from the use of electric generators and securing fuel in remote areas. Households without access to electricity, on-grid or off-grid, need to rely on more expensive sources of energy to power their mobile handsets. Low access to electricity can limit economic opportunities, thereby constraining the demand for mobile connectivity. Inversely, mobile connectivity can enable access to electricity, for instance, thanks to the introduction of new business models (e.g., Pay-As-You-Go systems) that support affordable off-grid electricity or through increased financial or knowledge flows from urban to rural areas.

In this paper, we exploit mobile phone data collected by Orange-Sonatel, the market leader, between January and December 2013 in Senegal, to investigate the impact of access to electricity on mobile connectivity, with a focus on rural areas and women. Mobile phone data, namely Call Detail Records (CDRs), is increasingly used to investigate topics in social science with a high level of accuracy (Blumenstock et al., 2015), especially in data-poor countries where standard sources of data may not be available. Our dataset includes detailed information about the time, location, usage, and devices' characteristics of both parties involved in a call or a text message, as well as the power source of telecommunication sites. We use this data to build statistics about access to and usage of mobile phones at the level of each site, with linkages between the originating and destination sites, enabling the assessment of the impact of access to electricity on new dimensions of mobile connectivity like incoming and outgoing calls and texts according to the characteristics of the originating areas.

Our main estimation strategy identifies 'randomness' in the roll-out of the electricity grid in Senegal to compare access to and usage of mobile connectivity between rural municipalities with and without access to electricity. The low population density in rural areas means that two similar municipalities may not have equal access to electricity due to their distance to the transmission grid. We use a propensity matching score to match rural municipalities with and without access to electricity on the basis of a

number of pre-determined characteristics including access to basic infrastructures like water and fixed broadband, household assets, and individuals' level of education and socio-professional status.

Three main findings stand out from our analysis. First, access to electricity strongly drives access to mobile connectivity, measured by the number of mobile telephony subscribers and smartphone ownership. More specifically, households with access to electricity have, on average, one additional mobile subscriber, and areas with access to electricity have, on average, 3 percentage-points more smartphone owners than areas with limited access. The impact on mobile telephony access is stronger for women as 4 percentage points more own a mobile phone in households with access to electricity. Second, usage of mobile connectivity also increases with access to electricity, but this effect stems more from incoming communications. Rural areas with access to electricity receive more calls and text messages than areas with more limited access. For outgoing communications, our estimates suggest that they rely on cheaper communications options like text messages instead of calls. Third, access to electricity increases mobile connectivity with urban municipalities with higher living standards. These municipalities place more calls and send more text messages to rural municipalities with access to electricity.

The findings of this paper are related to the literature on the determinants of mobile connectivity and the economic impact of rural electrification. Kirubi et al. (2009) and Lenz et al. (2017) evaluate the impact of rural electrification, respectively, in Kenya and Rwanda with some focus on mobile phone usage but did not investigate the inclusivity aspect. Armey and Hosman (2016) assess the impact of access to electricity on mobile connectivity in developing countries but use country-level data which does not allow to assess the impact on the rural or gender connectivity gap.

The remaining of the paper is organized as follows. Section "Related literature" provides a brief summary of the literature on the determinants of mobile connectivity and the impact of access to electricity on mobile connectivity. Section "Background on mobile connectivity and access to electricity in Sub-Saharan Africa" provides some background information about mobile connectivity and access to electricity in SSA. Section "Data and descriptive statistics" describes the data and presents descriptive statistics about the correlation between access to electricity and mobile connectivity. Section "Impact of access to electricity on mobile connectivity" presents the econometric models, the estimation strategy, and the results. Section "Conclusion" concludes.

Related literature

This paper can be related to two strands of the literature. A first strand investigates the determinants of mobile connectivity but typically focuses on long-run factors such as prices, income, urbanization, education, regulation, and competition intensity. For instance, Bohlin et al. (2010) find a positive impact of per capita income, urbanization, and regulation on the diffusion of new generations of mobile telecommunications. Likewise, Lee et al. (2011) and Lin and Wu (2013) highlight the role of education and content quality on broadband adoption in OECD countries. Gruber (2001) finds a positive impact of the number of operators and market size on the speed of diffusion of mobile telephony in Central and Eastern Europe. These papers focus on advanced economies where basic infrastructures such as electricity are widely available.

Armey and Hosman (2016) provide one of the first evidence about the impact of access to electricity on mobile connectivity in low-income countries, with a focus on access to the internet. However, they rely on country-level panel data of 40 low-income

countries, of which 29 from SSA, over the period 2000–2009, and find a positive and statistically significant impact of access to electricity on the penetration rate of the internet. Mothobi and Grzybowski (2017) also provide some evidence about the impact of access to electricity on mobile connectivity but rather focus on infrastructure deficit in general, measured by the average nighttime light intensity obtained from satellite data. Interestingly, they rely on a cross-section of individual-level survey data from 11 SSA countries which enable them to measure access to electricity at the household level and find that individuals are more likely to own a mobile phone in areas with brighter nighttime light or in households with access to electricity. Mobile-money services are more likely to be used in areas or households without access to electricity.

Another strand of the literature investigates the impact of rural electrification on development outcomes, without a specific focus on mobile connectivity. Kirubi et al. (2009) conducted a detailed case study of a community-based electric micro-grid in rural Kenya and find a positive impact on internet access at the community level. Lenz et al. (2017) assess the impact of a large-scale rural electrification program in Rwanda, connecting 370,000 new users within less than four years, and find a significant increase in mobile phone ownership and usage. However, probably because they rely on survey data, these studies do not provide a detailed analysis of how access to electricity affects incoming and outgoing communications which are important in understanding innovative programs that can support increased mobile connectivity in energy-deficient areas.

Background on mobile connectivity and access to electricity in Sub-Saharan Africa

In SSA, access to mobile connectivity has significantly increased over the past few years, but large gaps remain in terms of uptake and usage, especially in rural areas and for women. More than 90% of people in the region are covered by a mobile network (ITU, 2018), 70% covered by a 3G network, and 34% covered by a 4G network (GSMA, 2018). However, the uptake of mobile connectivity is still limited, with 44% penetration of mobile telephony and 24% of mobile internet. This gap is larger in rural areas where the penetration rate of mobile internet merely reaches 15%. As highlighted by Aker and Mbiti (2010), sharing of mobile phones is common practice across SSA, mainly among rural households' members, potentially limiting the use of mobile phones for sensitive purposes like healthcare and financial services. In addition, SSA has the second-largest digital gender gap, with 15% of women less likely to own a mobile phone than men, and 41% less likely to use the internet GSMA (2019).

A number of factors have been explored to explain this connectivity gap, but most boil down to the gap in access to electricity. A survey conducted by AfterAccess (2017) across 10 SSA countries in 2016 highlights the role of both supply and demand-side factors in constraining access to mobile connectivity. On the supply side, low network coverage is generally the second most important reason for not owning a mobile phone, after device affordability. As emphasized by GSMA (2014), the vast majority of mobile towers in SSA are in areas without access to electricity or with unreliable power sources, whereas energy costs can represent up to 40% of operating expenditures, especially in energy-deficient areas. On the demand side, lack of access to electricity at home is the third most cited reason for not owning a mobile phone in 2016 AfterAccess (2017), as the cost of recharging an internet-enabled device can go up to 6% of GDP per capita Wogan (2012). Another survey conducted by the GSMA (2016) across 10 SSA countries further highlights the role of demand-side factors such as lack of awareness and locally

relevant contents, digital literacy, and skills and network coverage, all of which can be traced back to lack of access to electricity.

Meanwhile, access to electricity in SSA remains limited, especially in rural areas. According to estimates by the International Energy Agency (IEA), the rate of access to electricity in SSA was 45% in 2018, with only 26% in rural areas. The challenge is concentrated in a few countries as half of the nearly 600 million without access to electricity live in the following six countries by order of importance: Nigeria, the Democratic Republic of Congo (DRC), Ethiopia, Tanzania, Uganda, and Sudan. In relative terms, the challenge is the biggest in landlocked countries like the Central African Republic, Chad, the DRC, Burundi, South Sudan, Niger, Burkina Faso, and Malawi, where the rate of access to electricity was around 10% in 2018. At the current rate of progress, it would take another 35 years to achieve universal access to electricity in SSA IEA (2019). Projections by the IEA suggest that under the current policy scenario, there will still be 531 million people without access to electricity in SSA in 2030, corresponding to roughly 40% of the projected population, mainly in Nigeria, the DRC, and landlocked countries.

A number of interventions have been undertaken, both from the public and the private sectors, to increase the availability of electricity. Public policies typically involve the expansion of power generation capacity, mostly for urban uses. For instance, Ethiopia is building the largest African dam to alleviate the electricity access gap. However, the transmission grid remains limited to urban areas. Private sector investment is still limited due to higher risks stemming from uncertain demand in rural areas. Private sector interventions take two forms: Public-private partnerships in power generations like independent power producers, decentralized off-grid (mini-grid, solar, or wind farms) solutions with long-term power purchase agreements (PPAs) with a utility or another off-taker, sometimes at regulated prices, to fill the supply gap in remote areas, often attached to mining activity.

Innovative solutions are increasingly emerging driven by the declining cost of off-grid solutions and mobile connectivity. These include the distribution of solar home systems, often taking advantage of mobile connectivity. An example is MKoppa across East Africa. More recently, telecoms operators are getting involved in the expansion of access to electricity, typically through the distribution of solar kits in rural areas, driven by synergies with mobile-money services. Examples include the partnership between the telecom operator Safaricom and M-Koppa in Kenya to distribute solar kits, or Orange partnering with BBoxx in the DRC, Burkina Faso, and Senegal.

Data and descriptive statistics

Data. The findings of this paper are derived from three data sources that have been combined at the level of the municipality:

■ Mobile call detail records: site-level access to electricity and mobile connectivity

Mobile CDRs have recently offered a pervasive and ubiquitous opportunity to sense individual behavior in many aspects of daily life. Mobile phones have been one of the fastest-growing and pervasive technologies over the last decades, reaching urban and rural populations across all socioeconomic spectra. CDRs are data automatically generated and recorded at a large scale by carriers for billing purposes. Every time a person uses a phone makes a call, sends an SMS, or goes online, a Call Data Record is generated. They contain an enormous amount of fine-grained information on who called whom for how long and from where: among the others, they contain the identifiers of the two parties who are communicating (by SMS or calls), the location of the tower handling the communication, date/time and duration

details. Indeed, they represent an extremely rich and informative source of data for scientists, and their analysis has become a mature research field with a wide range of applications Blondel et al. (2015).

Particularly, the past few years have seen the rise of research studies that have shown how to turn these data into actionable information to support global development. Areas of application range from public health to help tackle the spread of infectious diseases (Pullano et al., 2020, Rubrichi et al., 2018), to energy, for improved estimation of electricity, needs efficient electrification planning (Salat et al., 2020a, b), to mobility and transportation planning to design more efficient transportation systems or road networks or optimize the geolocalization of services (Di Lorenzo et al., 2016), to national statistics to help mapping populations and capturing socio-economic indicators (Blumenstock et al., 2015, Deville et al., 2014), and all hold highly promising potential for service and infrastructure improvements.

For the purpose of this research, we build monthly site-level indicators measuring access to electricity and mobile connectivity using a pseudo-anonymized dataset of mobile CDRs collected by Orange-Sonatel between January and December 2013 in Senegal, and another dataset on the source of the power of Orange-Sonatel's mobile antenna sites. The CDRs dataset includes information about the mobile phone usage of nearly 9 million users. Each record contains information about the nature of communication (voice calls, SMS, applications, etc.), the time-stamp, duration, and the antenna handling the communication. The power dataset includes information about the source of power (national grid, solar panel, or fuel-based generator) of 2000 mobile antenna sites across Senegal.

There are two main challenges in associating access to electricity with mobile connectivity from these datasets. On the one hand, users are mobile, i.e., they typically move from one area to another, and vice versa, for instance for work or personal visits. As a result, a given user may use his mobile phone in several areas with different levels of access to electricity over the course of a year. To overcome this challenge, we associate the whole 'phone activity' of a given user to the cellular tower of his/her residence, i.e., the 'home site'. To estimate users' home site, we follow a standard approach that evaluates users' activity at night time, the period of the day people are most likely to be at home, and defines 'home site' as the antenna site handling most of the user's night calls or SMS. After removing holidays-periods from the dataset, we count, for each user, the number of calls and SMS made (received), by the antenna site, from 9 p.m. to 2 a.m., and select the site with the highest records associated, for each distinct day of the year. The selected home site is the one with the highest number of calls and SMS over the observed period. Indicators of mobile connectivity are, therefore, computed for each user, and aggregated at the 'home site' level using the standard mean aggregation function. These indicators include the share of smartphone owners, the number and duration of outgoing and incoming calls, and the number of SMS per user. Table S-1 in the Supplementary Information (SI) presents the list of such indicators. On the other hand, the geographical coverage of a mobile antenna is usually not aligned with the geographical boundaries of municipalities. As a result, a site may cover several municipalities with different levels of access to electricity. To overcome this challenge, we focus on the availability of electricity at the site level, measured by a dummy variable equals to 1 if the site is powered by electricity from the national grid. Sites powered by solar panels only or electric generators are considered in areas without access to electricity.

▪ **Household survey data: household-level access to electricity and mobile connectivity**

Because the CDRs dataset does not include users' characteristics, we also measure access to electricity and mobile connectivity from a household survey. The survey dataset has the advantage of being exempt from the measurement challenges from the CDRs dataset but does not include detailed information about mobile phone usage, especially the direction of communications (outgoing and incoming) and users' network of contacts.

The household data were collected from a survey conducted by the Senegalese National Institute of Statistics in 2011. This survey was stratified by the urban part of the capital city Dakar, urban areas outside Dakar, and rural areas, and covers 17,891 households, corresponding to 168,203 individuals. We focus on household members above 15 years old (included), the typical subscribers of mobile telecoms services. The survey provides information on mobile subscription and expenditure, as well as on household and individual characteristics, namely access to electricity and water, the number of rooms, the type of dwelling (hut, barrack, or modern house), the nature of the wall (mud or cement), and the annual earning of active workers, as well as basic literacy, defined as the ability to read or write in any languages, including local ones. Access to electricity is a dummy variable and equals 1 if the household's main source of lighting is from the national grid or from a solar-based power source, and 0 otherwise. Furthermore, we disentangle households with access to electricity from an informal connection to their neighbors from those which subscribe with an official energy provider. This distinction will be useful for the causal identification strategy.

▪ **Census data: socio-demographic characteristics of municipalities**

Furthermore, we build indicators about the socio-economic characteristics of municipalities in order to compare mobile phone usage between antenna sites. To do so, we rely on the Senegalese census conducted in 2013. We have access to a sample of 144,893 households randomly selected by municipalities, corresponding to 10% of the total number of households per municipality, and a total of more than 1 million individuals. We drop individuals below 15 or above 75 from the sample to focus on potential mobile phone users. The census data provides information on households' characteristics, including access to electricity, access to water, fixed telephony, and internet, as well as ownership of basic appliances such as television, fridge, and fans. The ownership of fridges and fans will proxy for the household preference for energy, whereas the ownership of TV, fixed telephony, and internet will proxy for ICT preference. Additional information includes individual characteristics, namely gender, literacy, and socio-professional category (student, formal wage worker, informal wage worker, non-wage worker, and entrepreneurs). These data have been collapsed over 551 municipalities.

Descriptive statistics. We combine information on the location of antenna sites with data on the energy transmission network, as well as population density to provide a first insight into the association between access to electricity and the availability of mobile connectivity. Figure 1 shows the high availability of electricity and mobile connectivity in high-density areas and low availability in low-density areas. Access to electricity is limited in low-density areas whereas mobile connectivity, at least through 2G and 2.5G networks, is widely available, driven by regulatory requirements that come with mobile network license acquisition. This feature will be exploited in the causal estimation strategy as it means that we will be able to find two areas with similar coverage of the mobile networks, but with very different levels of access to electricity. This difference will be critical to assess how

Fig. 1 Energy transmission network, mobile phone coverage, and population density. Senegalese municipalities are coloured by population density. Solid blue line indicates the energy transmission network. Orange dots mark mobile phone towers.

Fig. 2 Rate of access to electricity and smartphone penetration. Senegalese municipalities are coloured by electricity access rate. Green dots indicate smartphone penetration rate in the corresponding municipality.

access to electricity affects mobile phone uptake and usage in low-density, typically rural, areas.

We also combine the rate of access to electricity from the census data with smartphone ownership from the mobile CDRs to obtain a first insight into any association between access to electricity and internet access. Mobile CDRs, in fact, provide the

Type Allocation Code (TAC) of the device that has served the communication. The TAC identifies a particular model of telephone and allows us to infer the percentage of smartphone owners. Figure 2 provides a visual representation of such association, showing high smartphone ownership in municipalities with high access to electricity.

Table 1 Access to electricity and mobile connectivity-descriptive statistics.

Site with access to electricity	Urban Dakar		Other Urban areas		Rural areas		Senegal
	Yes	No	Yes	No	Yes	No	Yes
% of smartphone owners	33.7		32.8		23.3	18.0	27.0
Outgoing calls (minutes)	118.3		92.6		61.6	43.4	81.3
Outgoing SMS	92.9		89.4		28.7	11.0	51.7
Incoming calls (minutes)	122.1		113.1		77.4	62.8	93.2
Incoming SMS	80.4		77.5		24.7	12.0	44.7
Incoming calls from Dakar	72.7		28.1		21.0	16.2	38.3
Incoming SMS from Dakar	73.6		27.9		13.0	6.8	33.5
% households with access to electricity	96.4		95.0		40.0	17.9	60.3
% households with access to tap water	97.7		94.8		60.9	36.9	74.1
% households with television	91.3		93.6		37.0	16.5	56.6
% households with fridge	58.3		51.5		11.6	3.2	28.2
% households with fans	72.0		68.9		17.4	3.8	37.0
% households with fixed telephony	21.5		15.5		4.3	3.5	10.3
% households with fixed internet	21.0		10.4		1.2	0.3	8.0
% of men*	50.0		47.4		47.4	48.6	48.3
% literates*	65.0		60.8		33.2	24.5	44.5
% students*	18.1		21.5		12.9	9.1	14.9
% Formal wage workers*	19.6		10.8		3.6	1.9	9.1
% informal wage workers*	21.6		20.6		24.6	28.8	23.5
% non-wage workers*	48.1		38.1		33.6	35.8	38.5
% entrepreneurs*	2.5		2.0		1.1	0.7	1.6
Number of sites	226	0	20	0	437	230	683

Mobile CDRs from Orange Sonatel over the year 2013. Municipality characteristics data from the census conducted in 2013.

*Individual characteristics apply to household members between 15 and 75 years old.

Table 1 compares mobile connectivity, measured by smartphone ownership, as well as outgoing and incoming calls and SMS, between electrified sites and non-electrified ones. All sites in urban areas have access to electricity and, as expected, mobile connectivity is much lower in rural areas than in urban areas.

In rural areas, antenna sites with access to electricity exhibit much higher mobile connectivity than those without. The gap in smartphone ownership is 5 percentage points, and users living closer to electrified sites make longer calls (42% higher) and send more SMS (30% higher). The gap in terms of incoming communications is even larger: 161% for the duration of incoming calls and 106% for the number of SMS. However, users with access to electricity receive more calls and SMS than they place, particularly with respect to the capital city, Dakar.

Statistics provided in the second part of Table 1 suggest that these differences could be driven by municipality characteristics. On average, electrified sites are located in municipalities with higher living standards, a higher share of households with access to electricity and water, with television, fridge, fans, fixed telephony, or the fixed internet. These municipalities have a higher proportion of individuals with an advanced level of education, with a higher share of literates and formal wage workers. These differences will be controlled for in the causal estimation strategy.

The aggregate statistics are well aligned with the national estimates from public sources. The share of smartphone owners in Senegal was estimated at 16% in 2013 by the GSMA (2018), below the aggregate of 27% from the largest operator data which is used in this paper, as this latter broadly attracts more high-income customers. The rate of access to electricity in Senegal was estimated at 57% of the population in 2013, compared to the 60% household rate of access from the census data.

Table S-2 in the SI provides descriptive statistics of variables from the household survey. Consistently with statistics from the CDRs dataset, households with access to electricity have a higher

percentage of mobile phone subscribers, both in urban and rural areas. Mobile phone usage also appears higher in urban households with access to electricity. However, the outcome is less clear in rural areas. Rural households with access to electricity spend more on mobile phones, per capita or in the percentage of their income, but less per subscriber.

The impact of access to electricity on mobile connectivity

Controlling for municipality characteristics: OLS estimates.

The econometric modeling begins by investigating any association between access to electricity and mobile connectivity, controlling for potential confounding factors, including municipalities characteristics. As discussed in the literature review, mobile connectivity can be determined by demand factors, namely price, income, household size, digital literacy, but also by supply-side factors such as the cost of network deployment and the number of sites. Price is constant because mobile telephony and electricity plans are the same across Senegal municipalities. Income is proxied by socio-professional categories. Digital literacy is proxied by basic literacy, that is the ability to read or write in any language. The cost of network deployment is not observed but is proxied by the availability of other infrastructure like water.

The econometric model can be expressed as:

$$MOB_i = \alpha + \beta ELEC_i + \gamma X_i + \varepsilon_i \quad (1)$$

where MOB_i denotes smartphone penetration, the volume of calls or the number of SMS (outgoing and incoming) per user at site i . The volume of calls is measured by the total duration of calls per user.

$ELEC_i$ is a dummy variable equals 1 if site i have access to electricity. Its coefficient β corresponds to the difference in mobile phone usage between sites with and without access to electricity.

X_i is a vector of controls including the proxies for income, literacy, and network deployment cost.

ε_i denotes the residual.

Table 2 Rural electrification and mobile connectivity-OLS estimates.

	smartpen	durpu_out	smspu_out	durpu_in	smspu_in	durpu_in_dak	smspu_in_dak
Electricity	3.20*** (0.37)	11.38*** (1.73)	11.14*** (1.07)	9.34*** (1.47)	8.07*** (0.84)	1.38** (0.59)	3.52*** (0.55)
Controls	yes	yes	yes	yes	yes	yes	yes
Sites	667	667	667	667	667	667	667
R-squared (in percentage)	42.0	29.4	40.8	29.5	36.9	26.4	31.1

Robust standard errors in parentheses. Significant at 1%(*), 5%(*), and 10%(*). 'smartpen' denotes the percentage of smartphone owners. 'durpu_out' is the duration of outgoing calls per user. 'smspu_out' is the number of outgoing SMS per user, and 'durpu_in' is the duration of incoming calls per user. 'smspu_in' is the number of incoming SMS per user, and 'durpu_in_dak' is the duration of incoming calls per user from Dakar. 'smspu_in_dak' is the number of incoming SMS per user from Dakar.

Equation (1) is estimated by ordinary least squares using data at the level of antenna sites but combined at the municipality level across the various sources. The outcomes of the estimation are reported in Table 2. They lend support to a positive correlation between access to electricity and mobile connectivity. Smartphone penetration is 3.2 percentage points higher at sites with access to electricity, compared to sites without. Likewise, the volume of calls, as well as the number of SMS sent or received, are larger at sites with access to electricity. These differences are smaller than the ones observed from the descriptive statistics in Table 1, due to the confounding effects of covariates like income, literacy, and network roll-out costs.

However, these estimates may still suffer from simultaneity bias and unobserved supply-side covariates. Typically, electricity and mobile telecommunications are simultaneously provided in municipalities with higher living standards. In the next section, we employ a matching estimator in order to address these concerns.

Controlling for municipality's characteristics: propensity score matching estimates. We employ a propensity score matching estimator (Rosenbaum and Rubin, 1983), after ruling out reverse causality and finding proxies for unobserved covariates that jointly determine mobile connectivity and access to electricity. Indeed, mobile connectivity increases income (Czernich et al., 2011, Gruber and Koutroumpis, 2011, Roller and Waverman, 2001), and could drive demand for access to electricity. In this case, it would not be sufficient to control for income. Instead, we exploit the full coverage of the mobile networks, as well as the specific drivers of the supply of electricity according to population density. In rural areas, population density is low and the electric grid cannot be concentrated on a particular type of user, for instance, the richest. This feature allows us to find, in these areas, similar households with and without access to electricity. Regarding covariates, we rely on the controls from the OLS estimation in the section "Controlling for municipality characteristics: OLS estimates". We implement the propensity score matching estimator by matching each electrified site with a set of nearest sites, on the basis of their propensity scores. The propensity score is calculated as the probability of being electrified given the demand and supply factors discussed above. More formally, we assume a logistic error term of the probability of a site being electrified and estimate the following equation:

$$\frac{P(\text{Elec}_i = 1/X_i)}{1 - P(\text{Elec}_i = 1/X_i)} = \exp(\beta X_i) \quad (2)$$

where Elec_i is a dummy variable equal to 1 if site i have access to electricity and 0 otherwise. X_i represents the vector of covariates. $P(\text{Elec}_i = 1/X_i)$ is the probability of having access to electricity given the characteristics X_i . $\exp$ is the exponential function and β is a vector of coefficients to be estimated by maximum likelihood.

The nearest sites j are determined on the basis of the distance $|\hat{P}_i - \hat{P}_j|$, where $\hat{P}_i$ denotes the estimated probability that site i has access to electricity given the characteristics of its municipality. There is a trade-off between bias and precision in the choice of the number of nearest sites. The higher number of nearest sites increases the precision of the matching estimator but also increases its bias. We choose the number of nearest sites to minimize the bias while maximizing precision when possible. The performance of the matching can be checked by comparing the characteristics of the average electrified site to the ones of the average match.

Table S-3 reports the outcomes of the matching algorithm. To the extent that all sites in urban areas are electrified, the matching is restricted to rural areas. In this latter, we focus on municipalities with a single site in order to control for unobserved supply-side factors. The treatment sample is made of 91 sites with access to electricity. Each of these sites has been matched with one without access to electricity, making a total of 33 unique controls. The t -tests reported in Table S-3 cannot reject the hypothesis that the treated sites are, on average, similar to the controls.

Main results.

Access and usage of mobile phone

The first panel of Table 3 reports the propensity score matching estimates of the impact of access to electricity on access to and usage of mobile connectivity. It turns out that the share of smartphone owners increases by 3 percentage points with access to electricity, and there is no significant increase in the number or duration of outgoing calls placed. The number of SMS sent increases with access to electricity, but this effect could probably be driven by unobserved digital literacy. While we found no significant difference in terms of literacy between the treated sites and their controls, basic literacy is a conservative measure of actual digital literacy. An alternative interpretation could be that energy access does impact the usage of mobile telephony, but only in terms of outgoing SMS, since it is the least expensive way of digitally connecting (in Senegal, Orange Sonatel's customers can text their contacts free of charge for a call-back (Sonatel, 2019)).

These findings accord well with estimates of the impact of access to electricity on other well-established mobile phone-based usage behavior indicators (de Montjoye et al., 2016), namely the level of activity measured by the number of days of activity, the percentage of communication initiated, and the percentage of communications at night, and the user's communication network size measured by the number of contacts. Estimates in Table 3 suggest that access to electricity comes with a higher level of activity and larger network size.

Incoming calls and text messages

The second panel of Table 3 reports the propensity score matching estimates on incoming calls and text messages. They suggest that users residing within the vicinity of a site with access

Table 3 Main results-propensity score matching estimates.

	w/ elec.	w/o elec.	Difference
Smartphone owners (%)	20.02	16.92	3.09*** (0.65)
Monthly volume of outgoing calls per user (minutes)	44.00	41.24	2.76 (2.91)
Duration of outgoing calls per user (seconds)	42.45	41.27	1.18 (1.69)
Number of outgoing calls per user	61.19	57.98	3.21 (2.47)
Number of outgoing SMS per user	21.39	12.67	8.72*** (2.03)
Number of days of activity	17.39	16.61	0.78*** (0.24)
Duration of incoming calls per user (seconds)	87.76	86.65	1.11 (1.93)
Number of incoming calls per user	46.46	41.52	4.94*** (1.40)
Monthly volume of incoming calls per user (minutes)	67.56	59.77	7.79*** (1.83)
Number of incoming SMS per user	17.94	12.19	5.75*** (1.22)
Volume of incoming calls from Dakar	19.07	18.38	0.69 (0.81)
Number of incoming SMS from Dakar	10.48	6.63	3.85*** (0.83)
Number of unique call contacts	29.09	26.25	2.84*** (0.84)
Number of sites	91	33	

Abadie-Imbens standard errors in parentheses. Significant at 1%(**), 5%(*), and 10%(*).

to electricity receive more calls and SMS. More specifically, users receive 5 more calls, making a total of 8 min per month, and 6 more SMS. Much of the additional incoming calls come from outside the capital city Dakar, while almost all additional incoming SMS come from there, reflecting the higher level of literacy in the capital city, a strong determinant of SMS usage (Schmid et al., 2017). Moreover, the increase in the overall incoming calls may also reflect the adoption of the SMS free of charge call-back option mentioned above. Note that while there is no significant effect on outgoing calls, more calls are directed to Dakar consistently with the increase in SMS from this city.

▪ Heterogeneity with respect to the characteristics of the originating municipalities

Given that rural users with access to electricity receive more communications, we investigate how such surplus is correlated with the living standards of their contacts. A positive correlation would exemplify the potential of mobile connectivity to support development in rural areas. Rural mobile phone users could benefit from financial and knowledge transfers from contacts in urban areas, potentially enabling entrepreneurship and productivity growth in rural areas. Such correlation could also support business model innovation, for instance, a joint provision of access to electricity or healthcare services and mobile connectivity in rural areas sponsored by urban contacts.

The methodology to estimate the correlation is as follow:

We consider all mobile antenna sites run by the incumbent mobile operator in rural areas. Each site with access to electricity is matched with another site without access to electricity, on the

basis of the socioeconomic characteristics of their municipalities. The matching is limited to rural municipalities with a single site to ensure an overlapping between the geographical boundaries of the municipalities and the site coverage. The matching algorithm follows the methodology described in the section “Controlling for municipality’s characteristics: propensity score matching estimates” using the same matrix of observable socioeconomic characteristics of municipalities. Let i denote each pair of sites, with a total of 36 pairs of sites.

For each site within-pair i , we observe the incoming communications per user (y_{ij}^0, y_{ij}^1) from all urban sites j (576 in total): y_{ij}^0 is the volume of incoming communications per user from urban site j towards the non-electrified rural site in the pair i ; y_{ij}^1 is the same metric but towards the electrified rural site in the pair i . The volume of communications at the level of a site is limited to users that live in the coverage area as described in section “Data”.

The surplus of communications per user within the pair i is denoted by $\Delta y_{ij} = y_{ij}^1 - y_{ij}^0$. On average, a user living in the coverage area of a rural site with access to electricity receives approximately half a minute more calls per month than a similar user without access to electricity, and this difference is statistically significant at the 1% level.

The econometric model can be expressed as:

$$\Delta y_{ij} = \alpha + \beta X_j + \mu_i + \varepsilon_{ij} \quad (3)$$

where X_j is a matrix of socioeconomic characteristics of living standards in the urban municipality that hosts the antenna’s site j . These characteristics include literacy and employment status. μ_i is a pair-specific effect, stemming from the characteristics of each rural municipality. Equation (3) does not include the socioeconomic characteristics of the rural municipalities due to the matching. ε_{ij} is the residual. The vector of coefficients β captures the heterogeneity effects.

Equation (3) is estimated by OLS. The outcomes of the estimation are reported in Table 4. While the R-squared is relatively small, the estimation is globally significant as suggested by the p -value associated with the F-statistics. In particular, there is a positive and significant correlation between the surplus of communication in rural areas with access to electricity and the living standards in the originating urban municipalities. For instance, rural users with access to electricity receive longer calls from urban areas with more wage workers. They also receive more text messages from urban areas with a higher share of literates.

▪ Heterogeneity with respect to users’ characteristics

We further assess the heterogeneity of the impact of access to electricity on mobile connectivity using household survey data and the propensity score matching estimator.

In rural areas, households have been matched according to their size, living standards, and level of literacy. In urban areas, the electric grid has full coverage, but the connection cost falls with population density. Typically, township residents can share the cost of connection by getting access to electricity from their neighbors. Therefore, there is an exogenous source of variation in access to electricity between households living in townships and those living in the other urban areas.

Table S-4 reports the outcome of the matching between electrified and non-electrified households. In rural areas, the treatment sample is made of 657 households with access to electricity, but without access to water. Fixing the status with respect to access to water is a way to control the unobserved cost of network deployment. In addition, access to water typically comes with access to electricity, reducing the probability of finding a match. Given a large number of households without access to electricity and to improve the precision

Table 4 Heterogeneity with respect to the characteristics of the incoming municipality-OLS estimates.

	$\Delta durpu$	$\Delta ncallpu$	$\Delta durpm$	$\Delta smspu$
% literates in urban areas	−2.530 (2.335)	0.128 (0.632)	−16.328 (15.370)	5.520** (2.200)
% formal wage workers in urban areas	19.716* (10.069)	0.582 (2.553)	174.860*** (65.133)	−9.953 (8.563)
% informal wage workers in urban areas	20.740* (10.746)	0.538 (2.752)	169.411** (72.451)	−9.302 (9.164)
% non-wage workers in urban areas	−16.465* (9.731)	0.043 (2.495)	−163.704*** (62.994)	4.970 (8.226)
% entrepreneurs in urban areas	9.172 (12.215)	−0.791 (3.186)	86.496 (87.007)	−9.652 (11.031)
% unemployed in urban areas	9.077 (8.309)	2.673 (1.665)	−2.560 (37.619)	−7.776 (4.988)
% students in urban areas	9.062 (5.864)	0.947 (1.404)	28.182 (34.961)	−10.394** (4.754)
Pairs fixed effects	Yes	Yes	Yes	Yes
Constant	−1.080 (2.001)	−0.332 (0.473)	15.376 (11.326)	1.171 (1.456)
Observations	16398	16398	16398	12980
Number of pairs	36	36	36	36
R-squared (in percentage)	1.56	3.30	2.54	1.70
F-Statistics	8.42	15.67	10.93	14.42

Robust standard errors in parentheses. Significant at 1%(*), 5%(*), and 10%(*).

Dependent variables are surplus in the volume of calls per user (*durpu*), the number of calls per user (*ncallpu*), the duration of calls per user (*durpm*), and the number of SMS per user (*smspu*). The p-value of the Fisher statistics is below 1% in all regressions.

Table 5 Mobile subscription and expenditures-propensity score matching estimates.

	Urban areas			Rural areas		
	w/ elec.	w/o elec.	Difference	w/ elec.	w/o elec.	Difference
Number of mobile subscribers	2.06	2.03	0.03 (0.21)	2.84	2.12	0.72*** (0.09)
Monthly household expenditures	8677.6	9223.8	−546.2 (3391.1)	11688.3	9289.4	2398.9*** (798.7)
Number of households	167	80		657	972	

Abadie-Imbens standard errors in parentheses. Significant at 1%(*), 5%(*), and 10%(*). The matching procedure equalizes household size across groups: 3.5 in urban areas and 6 in rural areas (See Table S-4). Expenditures in local currency (FCFA).

of the estimate, each electrified household is matched with three non-electrified households, making a sample of 972 unique controls. The *t*-tests reported lend support to a similarity between these controls and the treated households. The treatment sample in an urban area is made of 167 households living in the capital city Dakar that get access to electricity from their neighbors. Each of these households has been matched with one non-electrified household living in an urban area outside the capital city Dakar, making a sample of 80 unique controls. According to the *t*-tests, these controls are indeed similar to the treated households. The reported bias in terms of household size, living standards, or literacy is not statistically significant.

Table 5 reports the propensity score matching estimates of the impact of access to electricity on mobile subscription and expenditures. In urban areas, we find no statistically significant difference in mobile subscription or expenditures between electrified and non-electrified households. These findings contrast with the positive and statistically significant correlation highlighted by the comparisons in Table S-2. This is probably due to the strong simultaneity bias between access to electricity and mobile connectivity in urban areas. Indeed, in these high population density areas, users tend to be spatially concentrated

and access to electricity is typically purchased by those households with strong preferences for modern services such as mobile communications.

Using the household data and the matching estimator presented in section “Controlling for municipality’s characteristics: propensity score matching estimates”, we compare the gender, age, and literacy status of mobile subscribers between households with access to electricity and the comparable ones without access to electricity. The outcomes of the propensity score matching estimates are reported in Table 6. It turns out that the share of women subscribing to mobile telephony is 4 percent higher in households with access to electricity. This difference is statistically significant at the 5% level, suggesting that more women subscribe to mobile phones as their households get access to electricity. However, we find no statistically significant difference with respect to age and literacy, although the average subscriber in an electrified household tends to be younger and literate.

Conclusion

The findings of this paper lend support to the hypothesis that access to electricity supports digital inclusion. More specifically, we

Table 6 Profile of rural mobile subscribers-propensity score matching estimates.

	w/ elec	w/o elec	Difference
Percentage of women	37.43%	33.20%	4.23** (1.92)
Average age	36.16	36.34	−0.18 (0.61)
Percentage of literates	64.66%	61.82%	2.84 (2.01)
Number of households	604	952	

Abadie-Imbens standard errors in parentheses.
Significant at 1%(**), 5%(*), and 10%(*).

find that access to electricity increases mobile connectivity in rural areas, both in terms of access to mobile telephony and smartphone ownership. This effect is stronger for women, thereby contributing to closing the digital gender gap. In addition, we also find that access to electricity enables increased usage of mobile connectivity, but in an unexpected manner: rural mobile phone users with access to electricity rely on the cheapest communication means, typically short messages services, to communicate with their contacts, and receive more calls and text messages, especially from areas with higher living standards, than users with limited access to electricity. These findings accord well with the outcomes of another analysis of the relationship between access to electricity and mobile connectivity by Armev and Hosman (2016) in developing countries.

Our findings suggest that mobile connectivity, when combined with access to electricity, can support development in rural areas through financial and knowledge transfers as rural users with access to electricity receive more calls and text messages from municipalities with higher living standards. They imply that public policies aiming at increasing digital inclusion should account for access to electricity as an additional determinant on top of long-run factors like price, income, and digital literacy. However, it remains unclear how public policies would support access to electricity, especially with the goal of increasing digital inclusion, as the objectives of energy policies typically go beyond mobile connectivity. One approach could be a gradual provision of access to electricity in rural areas using decentralized solutions like mini-grids and solar home systems, given that uptake often significantly lags behind availability (Bernard, 2012). A study of the impact of these decentralized solutions on the demand for electricity in rural areas would guide public policies in rural electrification.

Data availability

Mobile phone data are proprietary and confidential. For legal reasons, the full dataset is not publicly available. However, access to the aggregated data can be requested from Orange authors on a contractual basis. The 2013 Senegalese census data are available under license to the registered users through the official website of the Senegalese National Statistical Institute ANSD. Household survey data are also available upon request from the ANSD.

Received: 4 September 2020; Accepted: 17 June 2021;

Published online: 14 July 2021

References

AfterAccess (2017) The inside internet story of Africa, Asia and Latin America. Report. AfterAccess

- Aker JC, Mbiti IM (2010) Mobile phones and economic development in Africa. *J Econ Perspect* 24:207–232
- Armev LE, Hosman L (2016) The centrality of electricity to ICT use in low-income countries. *Telecommun Policy* 40:617–627
- Bernard T (2012) Impact analysis of rural electrification projects in sub-Saharan Africa. *World Bank Res Observer* 27:33–51
- Blondel VD, Decuyper A, Krings G (2015) A survey of results on mobile phone datasets analysis. *EPJ Data Sci* 4:10
- Blumenstock J, Cadamuro G, On R (2015) Predicting poverty and wealth from mobile phone metadata. *Science* 350:1073–1076
- Bohlin A, Gruber H, Koutroumpis P (2010) Diffusion of new technology generations in mobile communications. *Inform Econ Policy* 22:51–60
- Czernich N, Falck O, Kretschmer T, Woessmann L (2011) Broadband infrastructure and economic growth. *Econ J* 121:505–532
- de Montjoye YA, Rocher L, Pentland AS (2016) bandicoot: a python toolbox for mobile phone metadata. *J Machine Learn Res* 17:6100–6104
- Déville P, Linard C, Martin S, Gilbert M, Stevens FR, Gaughan AE, Blondel VD, Tatem AJ (2014) Dynamic population mapping using mobile phone data. *Proc Natl Acad Sci USA* 111:15888–15893
- Di Lorenzo G, Sbodio ML, Calabrese F, Berlingiero M, Pinelli F, Nair R (2016) Allaboard: visual exploration of cellphone mobility data to optimise public transport. *IEEE Trans Visual Comput Graphics* 22
- Gruber H (2001) Competition and innovation: the diffusion of mobile telecommunications in central and eastern Europe. *Inform Econ Policy* 13:19–34
- Gruber H, Koutroumpis P (2011) Mobile telecommunications and the impact on economic development. *Econ Policy* 26:387–426
- GSMA (2014) Tower power Africa: Energy challenges and opportunities for the mobile industry in Africa. Report. GSMA and IFC
- GSMA (2016) Consumer barriers to mobile internet adoption in Africa. Report. GSM Association
- GSMA (2018) Mobile statistics. Database GSMA Intelligence
- GSMA (2019) The mobile gender gap report 2019. Report. GSMA
- IEA (2019) World energy outlook-electricity access database. Dataset. International Energy Agency
- International Energy Agency (2018) World energy outlook 2018. <https://www.iea.org/reports/world-energy-outlook-2018> [Accessed: May 2020]
- ITU. (2018) ICT statistics. Database. International Telecommunications Union
- Kirubi C, Jacobson A, Kammen DM, Mills A (2009) Community-based electric micro-grids can contribute to rural development: evidence from Kenya *World Dev* 37:1208–1221
- Lee S, Marcu M, Lee S (2011) An empirical analysis of fixed and mobile broadband diffusion. *Inform Econ Policy* 23:227–233
- 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 Dev* 89:88–110
- Lin MS, Wu FS (2013) Identifying the determinants of broadband adoption by diffusion stage in OECD countries. *Telecommun Policy* 37:241–251
- Mothobi O, Grzybowski L (2017) Infrastructure deficiencies and adoption of mobile money in sub-Saharan Africa. *Inform Econ Policy* 40:71–79
- Pullano G, Valdano E, Scarpa N, Rubrichi S, Colizza V (2020) Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the covid-19 epidemic in France under lockdown: a population-based study. *Lancet* 2:e638–e649
- Roller L-H, Waverman L (2001) Telecommunications infrastructure and economic development: a simultaneous approach. *Am Econ Rev* 91:909–923
- Rosenbaum PR, Rubin DB (1983) The central role of propensity score in observational studies for causal effects. *Biometrika* 70:41–55
- Rubrichi S, Smoreda Z, Musolesi M (2018) A comparison of spatial-based targeted disease mitigation strategies using mobile phone data. *EPJ Data Sci* 7:1–15
- Salat H, Schlapfer M, Smoreda Z, Rubrichi S (2020a) The impact of electrification on rural attractiveness in Senegal. Preprint at arXiv 2005.12464
- Salat H, Smoreda Z, Schlapfer M (2020b) A method to estimate population densities and electricity consumption from mobile phone data in developing countries. *PLoS ONE* 15(6):e0235224
- Schmid T, Bruckschen F, Salvati N, Zbiranski T (2017) Constructing socio-demographic indicators for national statistical institutes by using mobile phone data: estimating literacy rates in Senegal. *J Royal Statist Soc A* 180:1163–1190
- Sonatel O (2019) Call back service. <https://www.orangebusiness.sn/2/professionnels/1/26/rappellemoi-267678.html> [Accessed: May 2019]
- Wogan D (2012) Charging a mobile phone in rural Africa is insanely expensive. <https://blogs.scientificamerican.com/plugged-in/charging-a-mobile-phone-in-rural-africa-is-insanely-expensive/> [Accessed: May 2020]

Acknowledgements

GVH acknowledges funding, in 2018, from the Orange research program 'Digital Emerging Countries' directed by Valerie Pla.

Author contributions

GVH, ELQ, and SR conceived and designed the study. GVH and SR acquired and analyzed the data. GVH drafted the manuscript. All authors interpreted the results, contributed to writing and to the critical revision of the article. All authors read and approved the final manuscript to be published.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-021-00848-0>.

Correspondence and requests for materials should be addressed to S.R.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021