CSAE Working Paper WPS/2022/10

Community Effects of Electrification: Evidence from Burkina Faso's grid extension

Maika Schmidt* Alexander Moradi† ‡

September 21, 2022

Abstract

Using true and pseudo panel data of localities and households, we study the effects of Burkina Faso's large scale electricity grid expansion 2008-2017. We show that the timing of electrification was driven by engineering constraints and thus largely exogenous. We investigate the effects of electrification using a staggered difference-in-differences (DiD) approach, where not-yet treated communities serve as the control group. Despite low uptake of electricity at the household level, we find strong positive effects on luminosity at the community level. In terms of public goods provision, we find an increase in infant vaccination rates, electrified schools and drinking water provision. At the household level, we find increases in the ownership of electric appliances as well as an increase in bank patronage. Importantly, effects spill over to households that do not have an electricity connection.

***JEL* Classification: N77, O13, O18, O20; Q40**
Keywords: Electrification; Africa; panel event study; spill overs

*Department of Economics, University of Sussex, UK. E-mail: M.Schmidt@sussex.ac.uk.
†Department of Economics, University of Bozen-Bolzano, Italy. Also affiliated with the Center for the Study of African Economies, University of Oxford. E-mail: Alexander.Moradi@unibz.it.
‡We thank conference and seminar participants at Bolzano and the CSAE for valuable comments. We also thank Steve Stillman, Michael Grimm, Lucie Gadenne, Tilman Altenburg, and Mattia Di Ubaldo. We thank Sidiki Soubeiga for sharing with us the place names of LSMS clusters.

1 Introduction

One of the Sustainable Development Goals (SDGs) is to ensure universal access to affordable, reliable, and modern energy services by 2030. Despite significant progress in improving electricity access, sub-Saharan Africa is lagging behind the rest of the world. In 2018, nearly half the population lacked access to electricity. The International Energy Agency (IEA) estimated that around $60 billion a year would be necessary to reach the entire population by 2040 (IEA, 2019). Given the high costs of expanding electricity access, it is important to understand its contribution to development. There is a large literature documenting positive effects of electricity on household income and expenditure in developing countries (Akpandjar and Kitchens, 2017; Barron and Torero, 2014; Chakravorty et al., 2014; Chaplin et al., 2017; Grogan and Sadanand, 2013; Van de Walle et al., 2017).[1] However, findings for sub-Saharan Africa are more mixed - several studies find no positive income effects, or mixed effects (Bensch et al., 2011; Bernard and Torero, 2015; Dinkelman, 2011; Lenz et al., 2017). Several studies have questioned the cost-benefit case for electrification, particularly in rural settings of sub-Saharan Africa (Lee et al., 2020; Lenz et al., 2017; Bos et al., 2018).

In this paper we estimate the impact of community electrification in Burkina Faso on a wide range of development outcomes, taking into account both private effects at the household level, and public benefits at the community level – including both, public goods provision and spill over effects from electrified households. We exploit the roll-out of electrification across 558 Burkinabe localities between 2008 and 2017.[2] We show that the government connected more developed localities to the national grid. However, among localities that do receive electricity, the timing is much less dependent on local development and much more on engineering concerns. Hence, our identification relies on the quasi-exogenous timing of the roll-out, and our sample is composed

[1] Beyond income, positive effects have been shown to exist with educational outcomes, women's employment out of the house (Akpandjar and Kitchens, 2017; Dinkelman, 2011), and reduced indoor air pollution and incidences of respiratory disease (Barron and Torero, 2017).

[2] Each locality refers to a settlement (either a town or village) that we consider a community (see Web Appendix A for the exact definition). We use the terms locality and community interchangeably.

only of communities that are all eventually electrified. We use a DiD approach based on conditional parallel pre-trends to study the effects of community electrification, comparing electrified and non-yet electrified communities. Additionally, we employ an instrumental variable strategy based on a theoretical "least cost" grid expansion strategy.

Our findings point to significant positive community effects of electrification, despite relatively low uptake of household electricity. In terms of public goods, we find an increase in the number of electrified schools, drinking water systems, and infant vaccination rates. Additionally, luminosity increased significantly, suggesting the presence of street lights and/or increased economic activity more broadly. We also find suggestive evidence of an increase in population. In terms of private goods, we find an increase in the ownership of electrical appliances and financial inclusion that extends to households without private access to electricity. Despite of this, evidence of effectiveness remains mixed. Unlike Dinkelman (2011) for South Africa and Akpandjar and Kitchens (2017) for Ghana, we find little effect of electrification on female labor force participation and employment. We also do not find significant effects on school enrollment and children's nutrition and health outcomes.

This paper contributes to the literature in several ways. Firstly, many benefits of electricity access accrue at the community level, also benefiting unconnected households, e.g. due to improved provision of public goods and services and spill overs from connected households. In fact, many alleged benefits work through this channel (World Bank, 2008; Chaplin et al., 2017). Few studies have estimated community effects, and most of them focus only on labor market outcomes and household incomes (Akpandjar and Kitchens, 2017; Chaplin et al., 2017; Dinkelman, 2011; Lipscomb et al., 2013; Van de Walle et al., 2017).[3] More importantly, almost all studies explore community benefits at the *intensive margin*, meaning they study the effect of an increase in connection rates within already electrified communities. Those that study the *extensive margin* of

[3]Lee et al. (2020) point to potential externalities in rural Kenyan communities, where connected households share the benefits of electricity with neighbors, by letting them watch television or charge their mobile phones. Lighting from households with electricity has been shown to increase the perceived safety of the neighborhood (Bensch et al., 2011; Chaplin et al., 2017).

community electrification point to positive spill over effects (Khandker et al. (2013) for Vietnam and Van de Walle et al. (2017) for India). If community effects of electrification are positive, comparing households with and without electricity within the same community can potentially underestimate the total effect of electrification. While household's access to electricity can be relatively easily observed, spill over effects to the surrounding community are not well understood. Exceptions are Khandker et al. (2013) and Van de Walle et al. (2017), but these studies focus on the Asian context.

Secondly, many works study effects up to a few years after electrification. However, the full effect of electrification may have not been reached by then, because households do not have the liquidity to pay for connection fees, and broader knock-on effects unfold over time. We estimate the dynamics of a wide array of development outcomes for up to six years after electrification. Our estimates suggest that in our context a plateau in household connections and asset ownership is reached after about three years. Luminosity levels off after approximately four years. Overall, the dynamics point to a level but not growth effect of electrification.

Thirdly, we present methodological innovations for policy evaluation in the context of infrastructure expansion. There are a number of recent studies that provide experimental (Barron and Torero, 2017; Lee et al., 2020) and quasi-experimental evidence (Burlig and Preonas, 2022).[4] Our identification strategy is different. We apply a staggered DiD set-up based on the methodology of Callaway and Sant'Anna (2021). We compare localities that were connected with those that have not yet been connected. We argue that this reduces endogeneity concerns, because the exact timing of grid expansion was largely motivated by engineering considerations. In addition, we present a new instrument that may also be applicable in other contexts. We determine a hypothetical least cost grid expansion strategy connecting the existing grid with *development poles*. The development poles were localities that the Burkinabe government explicitly prioritized for electri-

[4] Barron and Torero (2017), like Lee et al. (2020) employed a randomized encouragement design, which consisted in offering electrical connection subsidies to a randomly selected sub-sample of households. Burlig and Preonas (2022) use population discontinuities in the eligibility to India's rural electrification program.

fication. The distance of the communities to this least cost electricity network then serves as an instrument for the timing of electrification, simultaneously controlling for concentric distances to the development poles that formed this network. In effect, the instrument captures localities that were connected earlier because they were located "on the way", mirroring a strategy that has been used for railroad lines (Jedwab and Moradi, 2016). Finally, we develop a new methodology for geographically matching DHS household clusters to localities, based on satellite data of settlements and population heat maps.

The remainder of the paper is organized as follows. Section 2 introduces the data sources and explains how we created our data set. Burkina Faso's electrification strategy is detailed in section 3, and section 4 explains the estimation strategy. Results are presented in section 5 and discussed in section 6.

2 Data

We compiled a geocoded data set on electrification roll-out in Burkina Faso from 1954 to 2017. To this we added data at the locality and household level. Details are explained in Web Appendix A.

Electrification. We have rich spatial information on the electrification strategy implemented in 2008, which effect we wish to study (Ministère des Mines, des Carrières et de l'Energie, 2008, thereafter MEPRED). We created a list of all localities connected to the national grid and their years of electrification.[5] The *Annuaire Statistique 2017* was our main source (INSD, 2018). With the help of a map showing the electricity grid in 2017, we geocoded the electrified localities as well as the power lines used to connect them. Overall, there were 660 localities which were connected to the grid between 1954 and 2017. For 20 localities, we were unable to match the locality name to a location. Thus, our data set consists of 640 georeferenced localities. Out of these, 558 localities were electrified between 2008 and 2017.

[5]We do not have information on isolated mini-grids or photovoltaic systems. They are not considered part of the national grid.

Luminosity. We use cloud free satellite imagery of stable night-light from i) the Defense Meteorological Satellite Program (DMSP) and ii) the Visible Infrared Imaging Radiometer Suite (VIIRS). DMSP measures luminosity in digital numbers (DN), for the years 1992-2013, ranging between 0 and 63 (capped), available at 30 arc-second grid resolution (ca. 1x1 km).[6] VIIRS measures luminosity in radiance in nano Watts per square centimeter per steradian, for the years 2012-2018, at 15 arc-second grid resolution. VIIRS offers the higher quality recordings detecting light at lower spectra with higher spatial accuracy (Gibson et al., 2021).

We create two measures of luminosity based on the highest pixel value lit within the locality's polygon.[7] First, we create a measure of extensive margin, i.e. a dummy variable indicating whether the locality was visible from outer space at all. Second, we create a measure of intensive margin, which quantifies the degree of luminosity for the electrified localities. We only use ever lit localities and we combine the two different series of satellite data. The two data series overlap for two years in 2012 and 2013, and we combine them based on their correlation during these two years.[8]

Other Data at the Locality Level. Furthermore, we collected detailed information on population, health facilities, schools, roads, drinking water provision, and irrigation. The purpose of collecting this data is two-fold. Firstly, these community characteristics allow us to determine the locality's development status prior to electrification, and thus let us study selection into electrification. Secondly, some of the measures are available over time and thus allow us to estimate the impact of electrification on community characteristics.

We have population counts from the 1996, 2006, and 2019 censuses (Institut National de la

[6]Top-coding is not a concern in our context. For localities electrified after 2008, the highest DN is 42.

[7]Alternatively, one could calculate mean pixel values within the locality's polygon, or create buffers of different sizes around the centroid or the highest lit pixel of the locality. Our results are robust to the choice of statistic (mean) and base (2 and 5 km buffer) in line with Min et al. (2013), Dugoua et al. (2018) and Burlig and Preonas (2022).

[8]Li et al. (2020) introduce a different approach to combining data from these two series. They "harmonized" and merged the two series into one. However, in our context, this data does not appear to really harmonize the two series. Figure B.4 shows luminosity for three years of electrification using the "harmonized" data. We find a large discontinuity in 2014 at the change from one series to the other. While this should be orthogonal, there is also implausibly large heaping (or censoring) at DN=6 in the harmonized data, strongly reducing annual variation. Thus, instead of using the "harmonized" data, we predicted radiance from DN running a locality fixed effects regression over ever lit localities.

Statistique et de la Démographie, 2000, 2011; Ministère de l'Eau et de l'Assainissement, 2020b).[9] We added population data for 2015 using the Burkina Faso - High Resolution Settlement Layer from Facebook and CIESIN (2019).[10] We retrieved data on health facilities in 2007 from Ministère de la Santé (2008). We have information on five types of facilities: i) CHR: Centre hospitalier régional (N=12); ii) CM: Centre médical (N=34); iii) CMA: Centre médical avec antenne chirurgicale (N=41); and vi) CSPS: Centre de santé et de promotion sociale (N=1,355) and v) dispensaries (N=86).[11] We geocoded annual census data of primary schools for the years 2003-2013 (Ministère de de l'Éducation Nationale et de l'Alphabétisation, 2019). The source informs about electricity at school and the number of students entering primary school. In addition, we reconstructed the location of private secondary schools in 2007 from Ministère des l'Enseignements Secondaire et Supérieur (2014). We retrieved the location of asphalt roads from Michelin road maps (Jewab and Storeygard, 2020). We also have information on the timing of AEP drinking water installations (Ministère de l'Eau et de l'Assainissement, 2020a). Finally, we digitized the location of irrigation systems in 2008 (N=97) (Figure 5, MEPRED, 2008) and major, permanent cattle markets based on the 2019 Survey of Livestock Markets (Table 1, Minot and Elahi, 2020).

Household survey data. We took household and individual level data from the 1993, 2003, 2010, and 2014 rounds of the Burkina Faso DHS (INSD and ICF). For confidentiality reasons, DHS displaced the true GPS coordinates of urban and rural clusters by up to 2 and 5 km respectively (Burgert et al., 2013). In our setting, this built-in measurement error – DHS clusters are erroneously matched to electrified/not yet electrified localities and thus – may lead to potentially large

[9] In repeated censuses, localities may be omitted, merged, renamed, misspelled, etc. We could not fully track 52 of the electrified localities across the censuses.

[10] The settlement layer provides population estimates at a 1 arc-second resolution (approximately 30m), based on satellite images and building recognition, to which then population derived from sub-national census data was proportionally allocated. We assigned population raster grids to localities following a procedure similar to Baruah et al. (2020) and AFRICAPOLIS (2022). First, we removed single hamlets from the layer. Second, we defined the boundary of a locality as the continuous built-up area within 100 meters. Thus, we constructed a buffer of 100 meters and joined all overlapping polygons. Finally, we calculated the total population within each polygon of a georeferenced electrified locality.

[11] The CSPS are the first level of care corresponding to primary healthcare centers. Staffed with about three health workers and two unqualified volunteers, they are responsible for 5 to 23 villages, providing basic outpatient services, including maternity care and administering vaccination programs in the villages. CM and CMA are the first referral level and are equipped with beds and surgical facilities (Marschall and Flessa, 2011). CHR are regional hospitals.

attenuation bias. We therefore apply a novel technique to assign DHS household clusters to electrified localities. For each DHS cluster, we calculated a *probability score* that describes the probability of the cluster lying in a given locality. We then only included DHS clusters in our analysis with a probability score of at least 25%, excluding multiple matches with lower scores for the same DHS cluster. We show that this strategy leads to a sample representative of treatment times and administrative regions. Details of the matching strategy are described in Web Appx. A.3.

The 1993, 2003, 2010 rounds are standard DHS surveys, whereas the 2014 round is a Malaria Indicator Survey (MIS) offering less information on health and well-being. This means that we could only observe some development outcomes until 2010. To overcome this limitation, we added a survey from the Living Standards Measurement Study (LSMS) carried out in 2014.[12] We identified LSMS cluster locations by place name (plus region, province, and commune name). Thus, we can perfectly match LSMS clusters to the correct electrified community and assign all of them a probability score of 1.[13] Overall, we have 219 DHS and 67 LSMS clusters, with roughly 22 households per cluster.

Summary statistics are reported in Web. Appx. Table B.2.

3 Background: Electrification in Burkina Faso

Burkina Faso is a small, poor, overwhelmingly rural country in West Africa.[14] It is also one of the least electrified countries in the world.[15] In 2004, the government announced a new national electrification strategy aimed at reducing poverty and providing electricity equitably across provinces, at the least cost possible. Planning activities became more concrete over the next 4 years. MEPRED

[12]This is the only LSMS survey available for our time frame of interest.

[13]Web Appx. Table A.1 report the number of DHS and LSMS clusters assigned to localities electrified from 2008 onward. We excluded the 1998 DHS survey round from our analysis. Only a few, selected places were sampled during this survey. As a consequence, the survey does not represent the electrified localities well.

[14]In 2000, Burkina Faso counted a population of 11.6 million, with a GDP/c (PPP) of 1,300 (constant 2017 international $) and 80% of the people employed in agriculture; in 2004, an estimated 57% of the population lived below the $1.90 a day poverty line (World Bank, 2020b). Between 2000 and 2018, per capita income grew steadily by an average of 3.5% per year (sd: 0.6).

[15]Burkinabe household's access to electricity ranked among the lower 20th to 30th percentile of sub-Sahara African countries in the 2000s and roughly maintaining this rank later on (Web Appx. Figure B.1).

developed the methodology and reviewed options.[16] At the end, Burkina Faso opted for building a unified national grid and linking this grid to the neighboring countries (African Development Fund, 2009, thereafter ADF).[17]

Figure 1 summarizes the progress in electrification in Burkina Faso. About 6% of households had electricity in 1993, increasing to about 20% in 2017. During the same time, the number of localities connected to the national grid grew from 21 to 660 in 2017, or by 16% per year. Access to electricity exhibits a strong urban-rural divide; in 2014, 58.4% of urban households, whereas only 4.7% in rural areas (INSD and ICF).

Figure 2 maps the electrified localities and the distribution network at two points in time. One can see that cities electrified by 2007 are larger in population. Indeed, the capital Ouagadougou and the second largest town got access to electricity as early as 1954. All 30 towns with more than 10,000 inhabitants in 1990 (except Bena and Yalgo) were connected by the year 2000. After 2007, access to electricity expanded to less populated localities. In our study, we focus on electrification 2008-2017.

Synergies of Electrification. The government's electrification plan specifically proposed improving the provision of public goods providing electricity for i) public lighting, ii) lighting systems in schools, iii) lighting and refrigeration to health centers, iv) water pump systems, and v) irrigation pump systems for off-season crops (Ministère de l'Économie et du Dèveloppement, 2004; Ministère des Mines, des Carrières et de l'Energie, 2008).[18] Effects were assumed to be broad. ADF (2009) speculated that semi-urban electrification would help the cottage industry[19] and trading activities, engaged in mainly by women. In the health sector, electric refrigerators could replace unreliable gas refrigerators thereby improving the storage of drugs and vaccines. Furthermore, small medical electronic devices and sterilization equipment could be used, and new

[16] MEPRED stands for Mainstreaming Energy for Poverty Reduction and Economic Development and was co-funded by the European Commission and DANIDA.

[17] Other supply options considered more expensive were: i) the rehabilitation and new construction of diesel thermal plants, ii) the development of photovoltaic solar energy and iii) the use of biomass (ADF, 2009).

[18] Prioritizing the connection of major public facilities is common, e.g. market centers, secondary schools, and health clinics in the case of Kenya (Lee et al., 2020).

[19] The cottage industry refers to business an manufacturing carried out in the home.

health equipment installed. In education, evening classes and new internet-based information technologies could be introduced. Furthermore, electrification would improve working conditions for teachers and health personnel positively affecting attendance and retention (ADF; MEPRED).

Nevertheless, the connection to the national grid may not have been the primary method for powering health facilities and schools. According to MEPRED (2008b), for example, Photovoltaic remained the technology for primary health care facilities (CSPS), even if connected to the distribution networks once available. Indeed we later show that the timing of the connection of a locality to the national grid is not associated with a substantial take-up of electrification at schools at that locality.

Selection into Electrification. MEPRED (2008) proposed to extend the 33kV network to localities with a significant demand for electricity linked to a strong development potential. For each locality, MEPRED (2008) calculated a composite index of developmental potential (DPI). The DPI covers three key dimensions: health, education and economic (each with a weight of 1/3). Sub-indices include the presence of health facilities, access to water, presence of formal and non-formal educational facilities, population size, agricultural facilities, markets, road infrastructure and financial institutions (see Web Appx. Table B.1 for a detailed list of all measures and their weights). Seeking impact for the benefit of the greatest number of people, the population in the catchment areas of the localities was then used to identify *development poles* that would be prioritized for electrification. In the analysis we control for whether a locality is considered a development pole, its DPI[20], as well as individual individual sub-indices.

Ministère des Mines, des Carrières et de l'Energie (2007) deemed the cheapest strategy of electrification to expand the network of 33kV lines in a zone of 50-60km around the existing 33kV distribution lines. The 33kV lines would typically serve the regional centers. From 33kV lines, smaller localities could be connected through single-wire earth return lines (SWER). SWER lines

[20]The government published DPI scores only for a subset of localities. We thus calculated the DPI for all localities based on available sub-indices. Web Appx. Figure B.2 shows that our DPI predicts reported DPIs extremely well. Moreover, regression coefficients reproduce the weights of the sub-indices.

are cheaper, because they use earth as the return path for the current avoiding the need for a second wire; lines have a lower power transfer capacity and are thus only a cost-efficient solution for localities with low electricity demand. Then, for localities further away from the national grid, for which an extension would be too costly, isolated off-grid diesel power plants were proposed. Finally, it was decided to extend the distribution network to localities located close enough to the new lines to justify in economic terms.

Actual electrification certainly deviated from planning.[21] So what type of localities got electrified? Table 1 tests for the covariates of electrification 2008-2017. Col. (1) confirms a significant positive correlation between the status of being a development pole and the probability of getting connected to the grid between 2008 and 2017. In addition, the locality's own population and DPI exhibit a positive correlation. The three variables are jointly highly significant $\chi^2 = 806.7$). Col. (2) adds the individual DPI sub-indicators as well as dummies for administrative capitals. The Pseudo R^2 increases indicating that the implicit weighting in the DPI may not reflect actual importance attached to those indicators in the electrification process. In other words, whether a community was selected for electrification depended on its development status beyond what is captured in the DPI. In col. 3, we include measures of distance to the 33kV line network as of 2007. We also include the distance to the hypothetical 33kV network that minimizes the length of lines when connecting all development poles (MST). We find localities further away from the pre-existing and optimal grid to be less likely to become connected.

Selection in the Timing of Electrification. In col. (4)-(6) of Table 1 we use the same set of explanatory variables to test for correlations with the year of electrification. Higher DPI localities got connected earlier (col. (4)). However, the correlation with locality development characteristics is rather weak. In fact, in col. (6) it turns out that the distance variables to the existing grid and the optimal grid (MST) are highly significant, whereas the development indicators that enter the DPI are jointly not significant (p-val: 0.18). Thus, the timing of electrification did not depend on

[21] MEPRED (2008b), for example, identified 619 development poles; 845 / 199 / 947 localities to be connected by 33kV lines / SWER / diesel by 2014. Instead, only 527 localities got access to electricity by 2017.

a locality's development status beyond what is captured in the DPI. It appears that it is governed by engineering concerns, i.e. that a network has to be expanded from the existing network.

Overall, localities were selected for electrification based on their size, and development and growth potential. The correlations highlight endogeneity concerns when comparing electrified to non-electrified localities. However, we follow a different strategy comparing localities that were actually connected, but in different years. Those localities should be more similar, as they all qualified for electrification. Indeed, we found that the timing is far less correlated with development. The state of 33kV distribution network determined the timing when localities were electrified. Localities closer to the existing network will be connected earlier. Moreover, it appears that there is more noise in the timing - in fact, we do not find any correlation between planned and actual date of electrification (p-val: 0.55).[22] Nevertheless, endogeneity issues may remain a concern, which we try to mitigate in three ways. First, we use observable, pre-existing indicators as controls. In particular, we control for the DPI score of each location. Second, we empirically test the assumption of conditional parallel pre-trends. Third, as an extension we exploit an instrumental variable strategy based on the engineering aspect of the electricity grid.

4 Estimation Strategy

We study the impact of electrification in Burkina Faso at two levels. First, we examine effects at the locality level using a panel data set of localities electrified from 2008 onward. This enables us to study community-level outcomes such as night-lights and population growth. It also allows us to test whether electrification came as a package, with improvements in other public goods such as water systems and schools. Second, we match household data from the DHS and LSMS surveys to the electrified localities and use these repeated cross-sections to analyze the effects of electrification at the household and individual level. Here, we can disentangle private and community benefits - whether effects also exist among non-electrified households in electrified

[22]The proposed date of electrification comes from MEPRED (2008).

localities.

Baseline estimation for the panel of localities. The effects of electrification may unfold over time. With a panel event study approach we take account of the dynamics. We have a panel of localities i over calendars years t. Electrification represents the *event*. The year of electrification g_i varies across localities. We define electrification as the *treatment*, electrified villages as the *treatment group* and not yet electrified villages as the *control group*. Localities that are treated stay treated. $T_{it} = t - g_i$ denotes the number of years that have passed since electrification for localities in group G_g. Group G_g includes all localities that were electrified in calendar year g. In the absence of staggered treatment timing (or under homogeneous treatment effects), we could simply stay in the canonical difference-in-difference (DiD) set up and use a dynamic two-way fixed effects estimator (TWFE) based on the following equation:

$$Y_{it} = \alpha + \sum_{k=1}^{K} \gamma_k 1(T_{it} = k) + \mu_i + \delta_t + X_{it}'\Gamma + \epsilon_{it} \tag{1}$$

where Y_{it} is the outcome variable, μ_i and δ_t are locality and year fixed effects, and X_{it} are a set of controls. With T_{it} being time relative to the electrification event, $1(T_{it} = k)$ is a dummy variable that takes the value one when t is k years after electrification.

However, in the presence of staggered treatment across groups (i.e. treatment happening at difference points in time for different localities) and heterogeneous treatment effects, this canonical difference-in-difference event study model based on the TWFE yields biased results (Borusyak et al., 2021; De Chaisemartin and D'Haultfoeuille, 2022; Goodman-Bacon, 2021; Sun and Abraham, 2021). To circumvent this issue, we follow Callaway and Sant'Anna (2021) in estimating "group-time average treatment effects" - causal estimates of the treatment effect for a specific treatment group at a specific point in time. We then aggregate these estimates across groups into time average treatment effects.

Specifically, we use the outcome regression (OR) estimator, based on not-yet treated localities without anticipation effects, from Callaway and Sant'Anna (2021) to identify the group-time average treatment effects:

$$ATT^{ny}_{OR}(g,t) = \mathbb{E}\left[\frac{G_g}{\mathbb{E}[G_g]}\left(Y_t - Y_{g-1} - m^{ny}_{g,t}(X)\right)\right] \tag{2}$$

where g indicates the treatment time of group G, and Y is the outcome of interest. D_{tg} is a dummy variable that denotes treatment and takes the value 1 if the locality has been electrified (i.e. if $t \geq g$) and $m^{ny}_{g,t}(X) = \mathbb{E}[Y_t - Y_{g-1}|X, D_t = 0, G_g = 0]$ is the population outcome regression for the group of localities that are not-yet-treated by time t, based on observable pre-treatment (or time invariant) covariates X. We then aggregate the ATTs across treatment groups g. Additionally, we aggregate them across time into a short-term event window $T = [0, 1]$ and a medium-term event window $T = [2, 6]$.

Baseline estimation for the repeated cross-section of survey data. In the context of the repeated cross-sections of surveys, we use a DiD set-up. Observation i describes an individual (or household) observed at time t who lives in a location from group G_g electrified at time g. D_{tg} is a dummy variable that denotes treatment and takes the value 1 if the locality has been electrified (i.e. if $t \geq g$). Again, in the absence of staggered treatment timing, we could use a TWFE estimator based on the following equation:

$$Y_{itg} = \alpha + \beta\ D_{tg} + \mu_g + \delta_t + X'_{it}\Gamma + \epsilon_{itg} \tag{3}$$

where D_{tg} is a dummy variable whether group g was electrified by time t. Y_{itg} is the outcome variable and X_{it} are a set of controls. Time fixed effects are denoted δ_t, and μ_g are fixed effects for the year of electrification (similar to locality fixed effects).

However, due to the presence of staggered treatment across groups, we once again follow Callaway and Sant'Anna (2021) in estimating "group-time average treatment effects" - as shown in equation

2. We then aggregate the ATTs across treatment groups g. In the survey data setting, where less time periods are available, we aggregate these into a single summary estimator of the overall treatment effect $T = (0, 6)$.

Community versus Private Effects. The survey data includes information on whether a household has private access to electricity. This allows us to disentangle the private effects of electrification from community or spill over effects. We re-estimate equation 2 for the sample of households *without* private access to electricity. This yields a lower-bound estimate of the community- or spill over effects. It is a lower-bound estimate, because we are comparing never-taker households (i.e. households that remain unconnected despite the community's connection to the national grid) in electrified localities with both never-taker and complier households in the not-yet-electrified localities. Complier households are those that would be unconnected in an unelectrified community but connected once the community is electrified. It is reasonable to assume that complier households differ substantially from never-taker households in their socio-economic status. In order to limit this problem, we control for pre-determined socio-economic characteristics of households. Specifically, we control for the gender and age of the household head, as well as for primary and secondary education of adult women in the household.[23]

Identification. An obvious threat to identification of the causal effect is selection into electrification. Treated localities may differ from non-treated ones in level and trend. We showed that selection is of limited concern when we limit the control group to those localities that will be electrified (up to 2017) but were not yet at time t. Indeed, one would expect localities that were electrified to be more similar (compared to those that were not chosen for electrification) thereby reducing the selection bias. We argue that eligibility criteria were common to all the electrified localities and that the state of 33kV distribution network largely determined the timing of when localities were electrified. To ease selection concerns, we control for the distance to this existing

[23]We do not have information on the education of men across all surveys, thus we do not include this.

network in 2007. We also control for the Development Potential Index which was used to measure electricity demand and eligibility during the planning process, as well as other determinants of electrification. However, this may not remove the endogeneity issue entirely. Towns that are electrified earlier may differ from those that were electrified later and a common trend assumption is still key for identification. The estimator from Callaway and Sant'Anna (2021) is particularly useful in this setting, because it requires only a weak form of the conditional trends assumption, as discussed below.

Identifying assumptions. The ATT OR estimator we employ based on "not-yet-treated" units without anticipation, imposes the following two assumptions (Callaway and Sant'Anna, 2021).

1. No treatment anticipation. The units of our analysis (localities) are not the ones choosing the date of electrification. Moreover, while localities may anticipate treatment, they are unlikely to be able to predict the date. In fact, if anyone relied on the proposed connection date by MEPRED (2008), they must have been disappointed given the zero correlation between planned and actual date of electrification (p-val: 0.55). In such context of highly uncertain public investments, there is little incentive to anticipate.

2. Conditional parallel trends based on not-yet-treated groups. Conditional on covariates, the average outcomes for the group first treated in period g and for the "not-yet-treated" groups would have followed parallel paths in the absence of treatment. This assumption holds after conditioning on covariates X, and allows for covariate-specific time-trends.

Callaway and Sant'Anna (2021) propose three different estimators for the ATT: ordinary regression (OR), inverse probability weighting (IPW), and doubly robust (DR). The advantage of the OR approach over the alternative estimators using propensity scores (IPW and DR) is that it does not require the overlap assumption. The overlap assumption requires "that a positive fraction of the population starts treatment in period g, and that, for all g and t, the generalized propensity score is uniformly bounded away from one"(Callaway and Sant'Anna, 2021). If propensity scores are sufficiently close to one, using either the IPW or the DR estimator can lead to a problem of irregular inference procedures. Since we have a limited sample (especially for the survey data) and a large number of covariates, we choose to employ the OR estimator, which does not use propensity scores. However, we do show results from the DR estimator for the community level analysis (where the sample is larger than in the household survey analysis).

In all specifications we control for the covariates that were significant predictors of the year of electrification, as in col(6) in Table 1 (development pole, distance to development pole, within 60km distance to existing network 2007, provincial/district capital, number of primary students 2007) as well as DPI, and region fixed effects. For the analysis using household survey data, we also control for population before the treatment period (in natural logarithms), as well as survey fixed effects, and the probability score which captures the probability that the household cluster was correctly matched to the electrified locality.[24] Based on these controls, we find that the conditional parallel trends assumption is not violated (see Figures 4, 5, 6).

5 Results

5.1 Community-level Effects (Locality Data)

Figure 4 shows the dynamics in outcome indicators at the locality level on an annual basis, whereas Table 2 reports aggregated ATTs for the short- and medium-term event window (T=[0,1] and [2,6] respectively). We begin by examining the conditional pre-trends assumption. For none of the outcomes do we observe meaningful differences in pre-treatment trends (see Figure 4). Only a small and insignificant increase in luminosity can be observed at $T = -1$ (Subfigures 4a and 4b).[25] The proportion of localities for which night-lights are visible from outer space was just 3.6% in $T = -1$, starts to increase from $T = 0$, and levels off at around 4 years after electrification.

[24] Lower probability scores are associated with greater measurement error and larger attenuation bias.

[25] It may be an artifact due to overglowing from near-by localities connected earlier. See Web Appx. Figure B.3 for an example of overglowing in our sample. Because it is so small in our context, considering $T = -2$ as baseline does not alter conclusions.

The coefficient is very large indicating an increase in localities with visible night-lights by ca. 45 percentage points (pp). An increase in luminosity can also be observed at the intensive margin. Again, we find that the effect reaches a plateau after about 3 years of electrification. We cautiously interpret this as a temporary growth effect. If we follow the literature and do not distinguish between intensive and extensive margin, we obtain a radiance of 0.32 at $T = 4$. This appears slightly larger than the results in the literature. For Senegal and Mali, for example, Min et al. (2013) reported an average increase of about 0.5 DN. Burlig and Preonas (2022) estimated an effect of 0.4 DN for India. Our estimate corresponds to about 0.7 DN.[26]

The evidence on other public goods is mixed. We find that electricity provision at the locality level leads to a modest increase in the supply of AEPS drinking water systems (Subfigure 4c). We do not find any effect on the number of public primary schools (Subfigure 4d) and student enrollment (Table 2). This can be rationalized by the observation that AEPS could be powered by electrical pumps, whereas schooling quantity is rather independent from electricity. Subfigure 4d indicates that a number of schools was electrified, simultaneously with connection to the grid and without delay, at $T = 0$. The effect of ca. 10 pp increase in the percentage of schools that are electrified is relatively large compared to the proportion of electrified schools in $T = -1$, which was 20% (Table 2).

Finally, it appears that electrified localities may outgrow not-yet electrified localities in terms of population. The timing of the effect is not measured precisely, because it is identified by only two measurements in time (2006 and 2015). Moreover, one would expect such an effect to be more gradual. Nevertheless, the estimate from Table 2 suggest a small yet insignificant effect. We reexamine this effect in more detail later.

Table 2 also reports estimates of the doubly robust (DR) DiD estimator (Callaway and Sant'Anna, 2021). Estimates tend to be larger using the DR estimator than when using the outcome regression (OR) estimator. Hence, OR DiD estimands appear to rather on the conservative side. Due to

[26]Based on the correlation between DN and radiance in 2012/13.

sample size constraints, we use only the OR estimator in our analysis of the survey data.

5.2 Community versus Private Effects (Survey Data)

Figure 5 shows the effects of community electrification on household electricity uptake and asset ownership over time. Reassuringly, we do not find any significant pre-trends. Confidence intervals tend to be large after $T = 3$. This is due to the fact that we only have observations until $t = 2014$, meaning we can only observe $T = 4$ for localities that were electrified between 2008 and 2010. Similarly, we only observe $T = 5$ and $T = 6$ for those electrified in 2008 or 2009, and 2008 respectively.

Table 3 reports average treatment effects 0 to 6 years after electrification. Uptake of electricity is significant but, with 7-8% of households, relatively low. Prior to community electrification, the share of households with an electric connection is virtually zero, with no significant pre-trends (see Subfigure 5a). For the other household assets, we estimate both the *total effects* of community electrification (pooling households with and without an electric connection), as well as the *spill over effects* on non-electrified households. We do not find an effect on radio ownership, probably because radios are most often battery powered. In fact, an average of 51% of households owned a radio in the year before the locality got connected to the grid. However, we find a strong increase in electric appliances that are truly complementary to electricity access. Television ownership increases after community electrification. Interestingly, this is the case even among households without household access to electricity. We also see an increase in the ownership of refrigerators, but this seems to be driven mainly by households with an electric connection. Electricity does not lead to more households having private access to piped water inside their houses. However, Subfigure 5d suggests that the time it takes to get to the main source of drinking water reduces a few years after electrification. The average treatment effect on "Time to Water Source" is not statistically significant at conventional levels for the aggregate effect (T=[0,6]), but the medium-term treatment effect (T=[2,6]) is highly significant (see Table B.3 in the Appendix). This is consistent with the community-level finding of an increase in AEPS drinking water systems.

Table 4 shows the effects of community electrification on the labor market for women. We find no significant effect of electrification on female labor force participation, paid work (either in kind or in cash) and non-agricultural work. Figure 6 shows the dynamics. There are no pre-or post-treatment differences. Confidence intervals are very large, especially for $T > 2$. Unfortunately, as we move up in event time, sample size decreases as those labor market variables were not collected in the MIS 2014.

Table 5 shows ATT estimates for health outcomes and inputs. We do not find any significant effect of community electrification on child health outcomes like infant and child mortality rates and child malnutrition (stunting and wasting). While we find generally positive effects on health inputs, coefficients are imprecisely estimated. The only significant effect is an increase in the share of children receiving the oral polio vaccine at birth. Burkina Faso has achieved very high vaccination rates in the 2010s. Hence, any effects of electrification must necessarily be small. This is not the case for prenatal care, birth attendance and medical treatment for fever. Nevertheless, coefficients are positive but remain statistically insignificant.

Taking a closer look at the immunization rates for infants reveals that there do seem to be medium-term effects of community electrification, particularly for the polio vaccines. Table 6 shows the effects of community electrification on infant vaccination rates separately for the short-term (0 to 1 year after electrification) and the medium-term (2 to 6 years after electrification). Polio vaccination rates increase in the medium run after electrification, both for the oral polio vaccine given at birth, and for the first dose of the inactivated polio vaccination given in the first year of life. There also seems to be an increase in the share of infants receiving the BCG vaccine against tuberculosis in the medium-term.

5.3 Extensions

The dynamic staggered regression estimator is our preferred specification, because it nicely illustrates the dynamics and makes efficient use of sample size. In this section we apply alternative identification strategies.

Canonical DiD setup. We follow a standard DiD set-up and estimate a static two-way fixed effects estimator (TWFE), as in equation 3, but with only two periods and two groups:

$$Y_{itg} = \alpha + \beta\ D_{tg} + \mu 1(g = 2008, 2009) + \delta 1(t = 2010 - 2014) + X_{it}'\Gamma + \epsilon_{itg} \tag{4}$$

where g is now divided into a treatment group (localities electrified in 2008-2009) and a control group (localities electrified in 2015-2017). We observe both groups in a pre-treatment period (2003-2007) and in a post-treatment period (2010-2014). D_{tg} is a dummy variable indicating that treatment has occurred. δ captures differences between the pre-and post periods. μ captures time-invariant differences between treatment and control group. β is the DiD estimate. Standard errors are clustered at the locality level. For the locality level analysis, we have yearly data for the pre- and post-treatment periods, whereas for the household level analysis the pre-period consists of the 2003 DHS survey and the post-period consists of the DHS and LSMS surveys from 2010 and 2014.

Instrumental Variable Approach. We argued that the timing of grid connection is less endogenous than selection into the grid and we showed that there are no differences in conditional pre-trends between localities electrified in different years. However, as a robustness check, we also introduce an instrumental variable approach. The timing of grid expansion is heavily constrained by engineering considerations. First, localities tend to be connected later if they are further away from the existing grid. Second, localities that happen to be located along the expanding network may be connected “on the go”. On this basis, we construct an instrumental variable.

We take note that MEPRED (2008) identified localities as development poles that promised high returns to electrification and should, therefore, have been prioritized. Table 1 indicates that the government followed MEPRED's advise relatively closely. Taking the existing grid network of 2007 and the development poles as vertices, we construct a least-cost grid network. Because network length is the major cost driver, we compute the minimum spanning tree (MST) that connects all vertices, with the minimum total network length possible. The derived MST is shown in Figure 3.

Our IV set-up is then as follows. At the locality level, outcomes are measured between 2010 and 2014. Localities electrified in 2008-2009 are the treatment group. Localities electrified in 2015-2017 represent the comparison group (as they were not yet treated in 2010-2014). Our instrument is log distance to the MST. We simultaneously control for whether the locality is a pole and distance to the poles to take into account that economic activities may be spaced concentrically around poles. Compliers are then localities that were connected earlier because they were located "along the optimal grid network". We estimate the reduced form (RF) and the 2SLS. The reduced form can be written as

$$Y_{irt} = \alpha + \beta ln(Distance\ to\ MST)_{ir} + \mathrm{X}'_{irt}\gamma + \delta_r + \epsilon_{irt} \tag{5}$$

for locality i in region r and years $t = [2010, 2014]$. Control variables X include indicator variables for the locality being a development pole, 60km to the nearest 33kV line in 2007, province and commune capital, and number of primary school students in 2007, logged distance to pole and the DPI. Standard errors are clustered at the locality level.

Results for the community level variables are reported in Table 7. We first report the ATT using the outcome regression DiD estimator based on the treatment (electrified 2008/2009) and control group (electrified 2015-2017) under study. Results from Table 2 $[T = 2, 6]$, using the full set of electrified localities, are largely confirmed. We find a significant increase in the proportion of lit localities, AEPS drinking water systems and electrified schools. The estimated effects differ in magnitude but without any striking pattern. For example, the effect on luminosity is smaller whereas the one on AEPS is larger. Next, the canonical 2x2 DiD estimator yields effects of indeed very similar magnitude as compared to the outcome regression DiD estimator in the full sample. Finally, the instrumental variable regression produces point estimates similar to the baseline OLS. We interpret this as a lack of evidence for strong endogeneity biases. As indicated by the Kleibergen-Paap and Montiel-Pflueger F-statistic the instrument is sufficiently strong, except for luminosity at the intensive margin. The instrument will necessarily be weak in this specification, owed to the fact that a large number of non-lit localities drop from the control group. Generally, standard errors lack precision, which is owed to the inefficiency in the IV estimation. Overall, the results from these different identification strategies line up with our interpretation from the staggered DiD model.

Results at the household level are reported in Table 8. The outcome regression DiD lacks power in this reduced sample. Point estimates, however, are very similar to the full sample in Table 3. The canonical 2x2 DiD estimator then yields consistent results under smaller standard errors. The only difference is that the percentage of households with finished floor and piped water appears to have increased after treatment, suggesting an increase in the quality of housing. We find similar effects in the IV approach. In line with the main specification, we find a strong increase in household electricity connections, ownership of televisions and refrigerators. In addition, the prevalence of a finished floor increased. The Kleibergen-Paap and Montiel-Pflueger F-statistics are reasonably large, suggesting that the instrument is not weak.

An important advantage of the IV strategy is that it only requires cross-sectional data, allowing us to expand the analysis to data that were not collected in the earlier survey rounds. Table 9 shows results for additional household level variables available only for 2010 and/or 2014. Just as before, localities electrified in 2008-2009 are the treatment group and localities electrified in 2015-2017 represent the control group. OLS results are provided as a benchmark. The instrument is again the log distance to the minimum distance spanning tree (MST). Panel A of Table 8 shows the treatment effects for the full sample, while panel B shows the spill over effects on non-electrified households. We find a large increase in the ownership of mobile phones, regardless of whether the household has an electric connection. This is what one expects if community electrification improves the opportunities for mobile phone charging for everyone. Importantly, we also find a large increase in the share of households that have a bank account. We cautiously interpret this as improvement in economic opportunities.

6 Discussion and Conclusion

In this paper we studied Burkina Faso's grid extension. The rich and detailed information on planning helped us to shed light on likely sources of endogeneity. The government explicitly targeted more developed localities in terms of their own population, their catchment area and existing private and public infrastructures. Grid supply followed assumed demand. This is probably a typical setting in developing countries. It would let us expect positively biased estimates of the impact when comparing electrified against non-electrified localities. We argued that the timing of electrification is more exogenous, because grid expansion is bound by engineering considerations. Expanding the grid means that localities closer to the existing network will be connected earlier. SWER lines are extended from 33kV lines. Costs are reduced by minimizing total network length. We showed that those engineering variables can predict the timing of electrification well not only trumping but rendering economic variables insignificant. This motivated us to choose not-yet electrified localities as control group and a staggered DiD approach as preferred identification strategy.

Evidence on the effectiveness of electrification is mixed. We found positive effects on public goods that are complementary to electricity. With the arrival of the grid, the number of electrified schools and infant vaccination rates increased. In contrast, there was no simultaneous increase in the number of schools suggesting that the electrification strategy was not a package where development expenditures were poured into electrified localities indiscriminately. Moreover, school enrollment has been unresponsive in our context, thus contradicting claims of broad schooling benefits. Similarly, we found evidence of drinking water systems installed in electrified communities, which are likely powered by grid electricity.

In line with much of the literature (Burlig and Preonas, 2022; Min et al., 2013; Min and Gaba, 2014; Dugoua et al., 2018), we found a significant increase in luminosity after a community got connected to the grid. If, as argued by Min et al. (2013), luminosity reflects mostly street lights, our results suggest significant public investment into street lighting. Of course, luminosity may also reflect private economic activity, household wealth and GDP more broadly (Henderson et al., 2012; Hodler and Raschky, 2014). In fact, our finding of increases in the household ownership of certain assets underlines this possibility.

However, we also found a number of non-results contradicting claims of the transformative power of electrification. We did not find significant effects of electrification on children's nutrition and health status, possibly because health inputs did not improve substantially either. Similarly, we found little to no evidence that community electrification increased female labor force participation. This contrasts the findings for South Africa (Dinkelman, 2011) and Ghana (Akpandjar and Kitchens, 2017). Of course, it may be specific to our context. The electrified localities are relatively small, with an average population of ca. 3,000. They are also more rural: a larger share of working women in Burkina Faso are employed in agriculture (World Bank, 2020a). Nevertheless, these findings are in line with a more pessimistic view recently expressed particularly with respect to African countries (Lee et al., 2020). Grid electricity may not be the bottleneck holding back income generating activities.

Our findings highlight the importance of considering the "arrival of the grid" (the extensive margin) rather than attempts to increase the electricity uptake of households under an existing connection (the intensive margin). Many of the benefits materialized within 2-4 years. This is even more important when positive spill overs exist at the community level. In our context, this was the case. With the arrival of the grid, more households own electric appliances that require electricity such as TVs, refrigerators and mobile phones. The effect also extends to households that do not have their own electricity connection. One explanation is that electricity is sourced by other means such as a power generator or solar panel. However, it appears implausible that these sources become en-vogue just at the time when the locality was connected to the national grid. Rather, there may by theft or on-selling of electricity from the households with a connection. We find particularly strong increases in the ownership of mobile phones and televisions, neither of which require a constant electric connection. There is much less spill overs in the ownership of refrigerators, which require a constant source of power. This is suggestive of the possibility that electricity is sold on in the community, with connected households giving temporary access to their neighbors. Indeed, Peters and Sievert (2016) find that across rural settings of sub-Saharan Africa, the most common income-generating use of household electricity consists of providing services to the local community, such as mobile phone charging or "cinemas". Similarly, Lenz et al. (2017) find that in Rwanda households connected to the grid rent out mobile phones to make calls or show films and football matches for a fee. Non-connected households spend less money on batteries, kerosene, and mobile phone charging if they live in an electrified area, suggesting informal sharing among friends and neighbors.

We also demonstrated that electrification increases the share of households who have a bank account, including for unconnected households. Financial exclusion remains a large problem in Burkina Faso. However, mobile money services have been expanding rapidly (N'dri and Kakinaka, 2020). ICTs more broadly are an important determinant for whether households hold a bank account (Karakara and Osabuohien, 2019). The fact that mobile phone ownership and bank accounts simultaneously expand in electrified localities can be seen in this light.

We found suggestive evidence that electrified localities may have outgrown not-yet electrified localities in terms of population. Most likely, this result stems from migration from never-electrified localities (outside our sample). People probably will not move to essentially similar localities in order to benefit from a slightly earlier date of electrification. We would rule out natural growth. First, we did not find large measurable effects on child mortality. Second, we would expect, if at all, a negative effect of electrification on fertility (Akpandjar and Kitchens, 2017; Grimm et al., 2015; Grogan, 2016).

Overall, this paper has demonstrated that there is an important community dimension to the benefits of electrification, comprising both the provision of public goods, as well as spill over effects from electrified households to the community. The community dimension is well recognized for other infrastructure investments, such as roads and water provision, but the literature studying the welfare effects of electrification has so far focused mainly on the private benefits of household electric connections at the intensive margin. An accurate assessment of the welfare effects of grid extension needs to take into account public goods at the community level and spill over effects.

REFERENCES

African Development Fund (2009). Electricity infrastructure strengthening and rural electrification project. project appraisal report. Burkina Faso. 8, 9

AFRICAPOLIS (2022). Defining urban. https://africapolis.org/en/about/defining-urban. 6, 46

Akpandjar, G. and Kitchens, C. (2017). From darkness to light: The effect of electrification in Ghana, 2000–2010. *Economic Development and Cultural Change*, 66(1):31–54. 1, 2, 24, 25

Barron, M. and Torero, M. (2014). Electrification and time allocation: experimental evidence from northern El Salvador. *MPRA Paper 63782*. 1

Barron, M. and Torero, M. (2017). Household electrification and indoor air pollution. *Journal of Environmental Economics and Management*, 86:81–92. 1, 3

Baruah, N. G., Henderson, J. V., and Peng, C. (2020). Colonial legacies: Shaping African cities. *Journal of Economic Geography*, 21(1):29–65. 6, 46

Bensch, G., Kluve, J., and Peters, J. (2011). Impacts of rural electrification in Rwanda. *Journal of Development Effectiveness*, 3(4):567–588. 1, 2

Bernard, T. and Torero, M. (2015). Social interaction effects and connection to electricity: experimental evidence from rural Ethiopia. *Economic Development and Cultural Change*, 63(3):459–484. 1

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*. 12

Bos, K., Chaplin, D., and 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. 1

Burgert, C. R., Colston, J., Roy, T., and Zachary, B. (2013). *Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys*. DHS Spatial Analysis Reports No. 7. ICF International, Calverton, Maryland. 6, 49

Burkina Faso (2021). Nombre d'abonnés en électricité par région 2006-2017. https://www.data.gov.bf/search/field_topic/energie-46/. 30

Burlig, F. and Preonas, L. (2022). Out of the darkness and into the light? Development effects of rural electrification. *Journal of Political Economy, forthcoming*. 3, 5, 17, 24

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230. 3, 12, 13, 15, 17, 37, 38, 39, 40, 58

Chakravorty, U., Pelli, M., and Marchand, B. U. (2014). Does the quality of electricity matter? evidence from rural India. *Journal of Economic Behavior & Organization*, 107:228–247. 1

Chaplin, D., Mamun, A., Protik, A., Schurrer, J., Vohra, D., Bos, K., Burak, H., Meyer, L., Dumitrescu, A., Ksoll, C., et al. (2017). Grid electricity expansion in Tanzania by MCC: Findings from a rigorous impact evaluation. *Report submitted to the Millennium Challenge Corporation, Washington, DC: Mathematica Policy Research*. 1, 2

de Bellefo, M.-P., Combes, P.-P., Duranton, G., Gobillon, L., and Gorin, C. (2021). Delineating urban areas using building density. *Journal of Urban Economics*, 125. 49

De Chaisemartin, C. and D'Haultfoeuille, X. (2022). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. Technical report, National Bureau of Economic Research. 12

Dinkelman, T. (2011). The effects of rural electrification on employment: New evidence from South Africa. *American Economic Review*, 101(7):3078–3108. 1, 2, 24

Dugoua, E., Kennedy, R., and Urpelainen, J. (2018). Satellite data for the social sciences: measuring rural electrification with night-time lights. *International Journal of Remote Sensing*, 39(9):2690–2701. 5, 24

Facebook and CIESIN (2019). High resolution settlement layer (hrsl). source imagery for hrsl. https://energydata.info/dataset/burkina-faso-high-resolution-settlement-layer-2015. 6, 44, 46, 49

Gibson, J., Olivia, S., Boe-Gibson, G., and Li, C. (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics*, 149:102602. 5

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277. 12

Grimm, M., Sparrow, R., and Tasciotti, L. (2015). Does electrification spur the fertility transition? Evidence from indonesia. *Demography*, 52(5):1773–1796. 25

Grogan, L. (2016). Household electrification, fertility, and employment: Evidence from hydroelectric dam construction in Colombia. *Journal of Human Capital*, 10(1):109–158. 26

Grogan, L. and Sadanand, A. (2013). Rural electrification and employment in poor countries: Evidence from Nicaragua. *World Development*, 43:252–265. 1

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *The American Economic Review*, 102(2):994–1028. 24

Hodler, R. and Raschky, P. A. (2014). Regional favoritism. *The Quarterly Journal of Economics*, 129(2):995–1033. 24

IEA (2019). *Africa Energy Outlook*. International Energy Agency, Paris. 1

INSD (2018). *Tableau 15.07: Electrification des localités du Burkina Faso*, pages 251–252. INSD. 4, 30, 44

INSD and ICF (various). Burkina Faso Demographic and Health Surveys. `https://dhsprogram.com/`. 6, 8, 30

Institut National de la Statistique et de la Démographie (2000). *Recensement Général de la Population et de l'Habitation de 1996 - Fichier des villages du Burkina Faso*. INSD, Ouagadougou. 5

Institut National de la Statistique et de la Démographie (2011). *Recensement Général de la Population et de l'Habitation de 2006 - Fichier des villages du Burkina Faso*. INSD, Ouagadougou. 6, 30, 60

Jedwab, R. and Moradi, A. (2016). The permanent effects of transportation revolutions in poor countries: Evidence from Africa. *Review of Economics and Statistics*, 98(2):268–284. 4

Jewab, R. and Storeygard, A. (2020). The average and heterogeneous effects of transportation investments: Evidence from sub-Saharan Africa 1960-2010. *Journal of the European Economic Association*, 20:1–38. 6

Karakara, A. A. and Osabuohien, E. S. (2019). Households' ICT access and bank patronage in West Africa: Empirical insights from Burkina Faso and Ghana. *Technology in Society*, 56:116–125. 25

Khandker, S. R., Barnes, D. F., and Samad, H. A. (2013). Welfare impacts of rural electrification: A panel data analysis from Vietnam. *Economic Development and Cultural Change*, 61(3):659–692. 3

Lee, K., Miguel, E., and Wolfram, C. (2020). Experimental evidence on the economics of rural electrification. *Journal of Political Economy*, 128:1523–1565. 1, 2, 3, 8, 24

Lenz, L., Munyehirwe, A., Peters, J., and Sievert, M. (2017). Does large-scale infrastructure investment alleviate poverty? impacts of Rwanda's electricity access roll-out program. *World Development*, 89:88–110. 1, 25

Li, X., Zhou, Y., Zhao, M., and Zhao, X. (2020). A harmonized global nighttime light dataset 1992–2018. *Scientific Data*, 7(168):1–9. 5, 60

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

Marschall, P. and Flessa, S. (2011). Efficiency of primary care in rural Burkina Faso. a two-stage DEA analysis. *Health Economics Review*, 1(1):5. 6, 48

Min, B. and Gaba, K. M. (2014). Tracking electrification in vietnam using nighttime lights. *Remote Sensing*, 6(10):9511–9529. 24

Min, B., Gaba, K. M., Sarr, O. F., and Agalassou, A. (2013). Detection of rural electrification in Africa using DMSP-OLS night lights imagery. *International Journal of Remote Sensing*, 34(22):8118–8141. 5, 17, 24

Ministére de de l'Éducation Nationale et de l'Alphabétisation (2019). Evolution des inscrits en alphabétisation initiale. 47

Ministère de de l'Éducation Nationale et de l'Alphabétisation (2019). Liste des écoles avec électricité du burkina faso de 2003-2013. `https://data.gov.bf/`. 6, 47

Ministère de la Santé (2008). *Carte sanitaire 2007*. INSD, Ouagadougou. 6, 48

Ministère de l'Eau et de l'Assainissement (2020a). Résultats de l'inventaire national des ouvrages d'approvisionnement en eau potable de 2019: 11. liste des aeps-pea. `https://www.eauburkina.org/index.php/resultats-ino`. 6, 47

Ministère de l'Eau et de l'Assainissement (2020b). Résultats de l'inventaire national des ouvrages d'approvisionnement en eau potable de 2019: 18. taux d'accés par quartier. `https://www.eauburkina.org/index.php/resultats-ino`. 6, 46

Ministère de l'Économie et du Dèveloppement (2004). *Cadre Stratégique de Lutte contre la Pauvreté*. Burkina Faso. 8

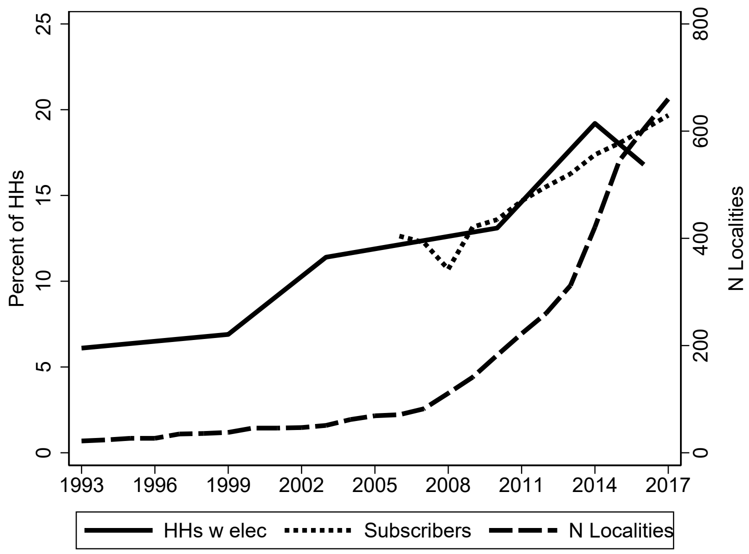
Ministére de l'Energie (2017). *Tableau 16: Localités électrifiées par région*, pages 19–22. Ministére de l'Energie. 44

Ministère des l'Enseignements Secondaire et Supérieur (2014). *Liste des Établissement Privés d'Enseignement secondaire Général 2014-2015*. Ministère des l'Enseignements Secondaire et Supérieur. 6, 47

Ministère des Mines, des Carrières et de l'Energie (2007). Stratégie de développement de l'électrification

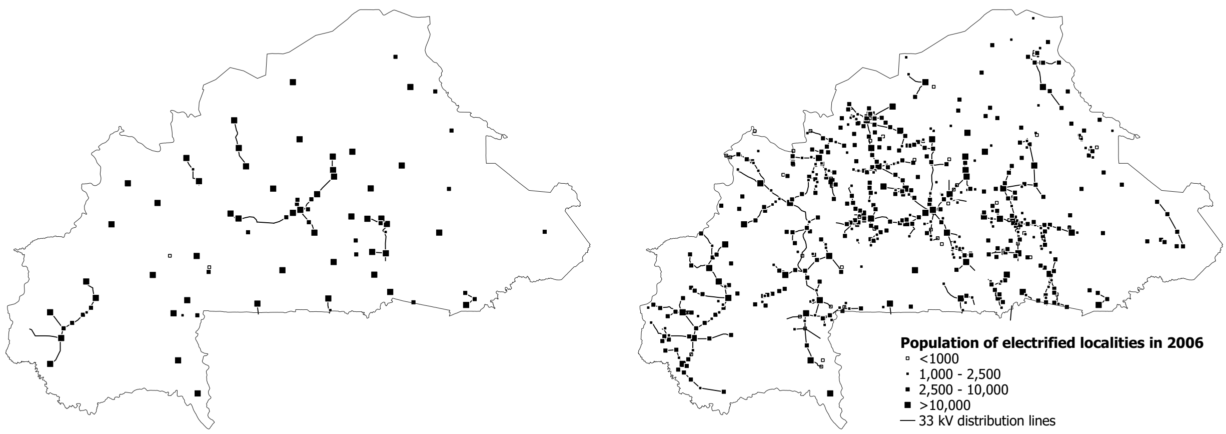
rurale au Burkina Faso. 9
Ministère des Mines, des Carrières et de l'Energie (2008). *Programmation régionale MEPRED – Rapport final*. Burkina Faso. 4, 6, 9, 11, 20, 21, 31, 35, 45, 46, 47, 56, 59
Ministère des Mines, des Carrières et de l'Energie (2008). *Vision 2020. De l'accés aux services énergétiques modernes*. Burkina Faso. 8, 9, 10, 15, 45
Ministère des Mines, des Carrières et de l'Energie (2020). Système d'information géographique sur l'état et la projection de l'électrification rurale au Burkina Faso. https://www.improves-re.com/sig/SIG/th_Reseau/default.asp. 44
Minot, N. and Elahi, S. (2020). Livestock markets in Burkina Faso: Characteristics and impact on the local economy. *International Food Policy Research Institute*. 6, 47
Moner-Girona, M., Bódis, K., Huld, T., Kougias, I., and Szabó, S. (2016). Universal access to electricity in Burkina Faso: Scaling-up renewable energy technologies. *Environmental Research Letters*, 11(8):084010. 45
Montiel Olea, J. L. and Pflueger, C. E. (2013). A robust test for weak instruments. *Journal of Business and Economic Statistics*, 31(8):358–369. 41, 42, 43
Moriconi-Ebrard, F., Harre, D., and Heinrigs, P. (2016). *Urbanisation Dynamics in West Africa 1950–2010*. OECD. 49
N'dri, L. M. and Kakinaka, M. (2020). Financial inclusion, mobile money, and individual welfare: The case of Burkina Faso. *Telecommunications Policy*, 44(3):101926. 25
NGA (2017). GEOnet Names Server. https://geonames.nga.mil/gns/html/. 44, 46
OCHA (2020). Burkina Faso - subnational administrative boundaries. 45
Peters, J. and Sievert, M. (2016). Impacts of rural electrification revisited–the African context. *Journal of Development Effectiveness*, 8(3):327–345. 25
Rios-Avila, F., Sant'Anna, P. H. C., and Callaway, B. (2021). CSDID: Difference-in-differences with multiple periods. *Journal of Econometrics*. 36, 37, 38, 39, 40, 58
Sant'Anna, P. H. C. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122. 36
SONABEL (2018). *Liste des projets d'infrastructures énergétiques et d'études électriques envisagés ou en cours à la SONABEL*. SONABEL, Ouagadougou. 44
Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199. 12
UCF (2018). *Carte A1: Carte du réseau de transport à fin mars 2018*, page 74. UCF. 44, 45
Van de Walle, D., Ravallion, M., Mendiratta, V., and Koolwal, G. (2017). Long-term gains from electrification in rural India. *The World Bank Economic Review*, 31(2):385–411. 1, 2, 3
World Bank (2008). The welfare impact of rural electrification: A reassessment of the costs and benefits. 2
World Bank (2020a). Enterprise surveys. http://www.enterprisesurveys.org. 24
World Bank (2020b). World development indicators. 7, 59
Wu, J., He, S., Peng, J., Li, W., and Zhong, X. (2013). Intercalibration of DMSP-OLS night-time light data by the invariant region method. *International Journal of Remote Sensing*, 34(20):7356–7368. 45

Figure 1: Electrification in Burkina Faso 1993-2017



Notes: This graph shows various measures of access to electricity. The percentage of households with electricity is from the 1993, 1999, 2003, 2010, 2014 and 2016 DHS (INSD and ICF, ious). We also calculated the percentage of households that subscribed to the services of the electricity providers SONABEL and FDE. We divided the number of subscribers reported in Burkina Faso (2021) by the average household size derived from INSD and ICF (ious). Finally, the number of electrified localities (secondary axis) is sourced from INSD (2018).

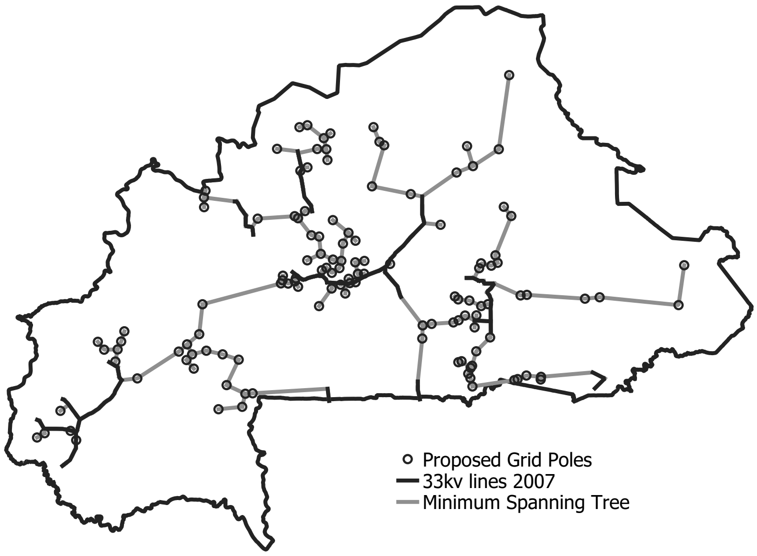
Figure 2: Electrified Localities 2007 and 2017



(a) 2007 (b) 2017

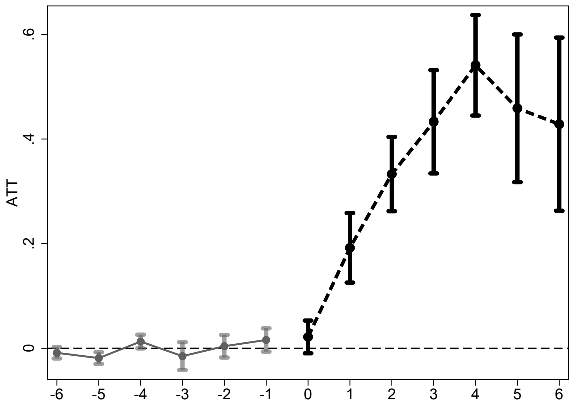
Notes: The maps show the electrified localities and their population in 2006 (Institut National de la Statistique et de la Démographie, 2011). Subfigure 2a shows the situation in 2007. N(localities <1,000; 1,000-2,500; 2,500-10,000; >10,000)=2; 6; 26; 47. Subfigure 2b shows the situation in 2017. N(localities)=61; 184; 277; 56. The 33 kV power lines are the main distribution network. See Web Data Appendix for data sources.

Figure 3: Proposed Grid Poles and Minimum Spanning Tree

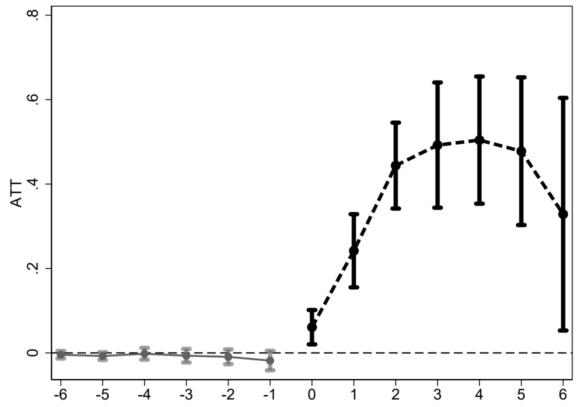


Notes: The map shows the proposed development poles by MEPRED (2008). Taking into account the existing 33kv grid of 2007 the network that would connect the development poles minimizing network length is the Minimum Spanning Tree (MST).

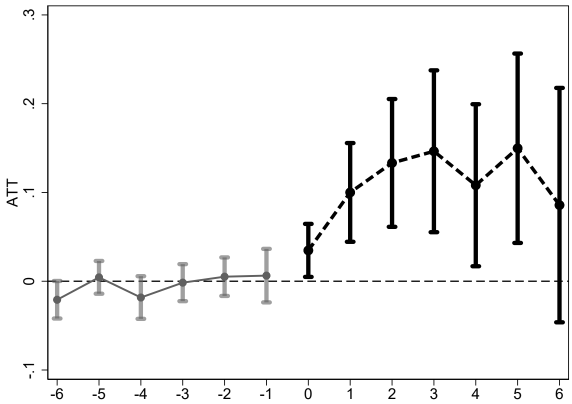
Figure 4: Luminosity, Drinking Water, Schooling and Population Before/After Electrification



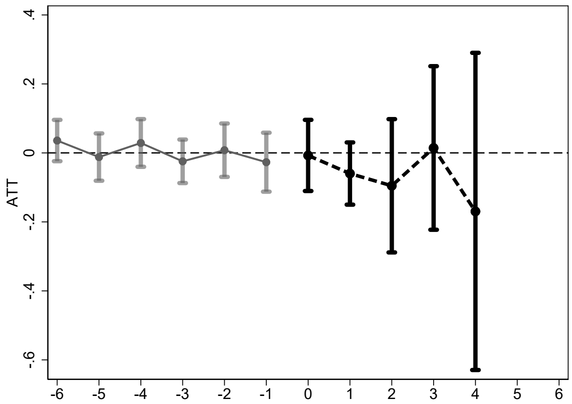
(a) Luminosity, Extensive Margin



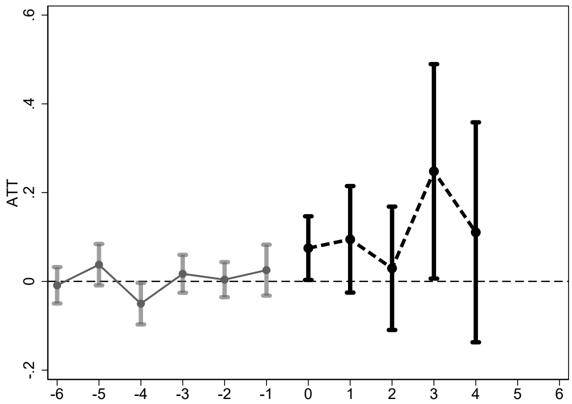
(b) Luminosity, Intensive Margin



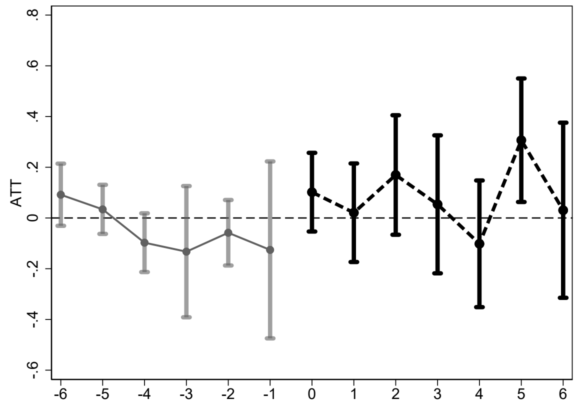
(c) Drinking Water (AEPS), Extensive Margin



(d) N Public Primary Schools



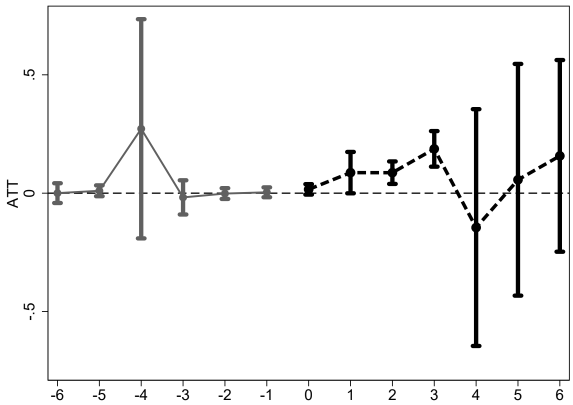
(e) N Electrified Schools



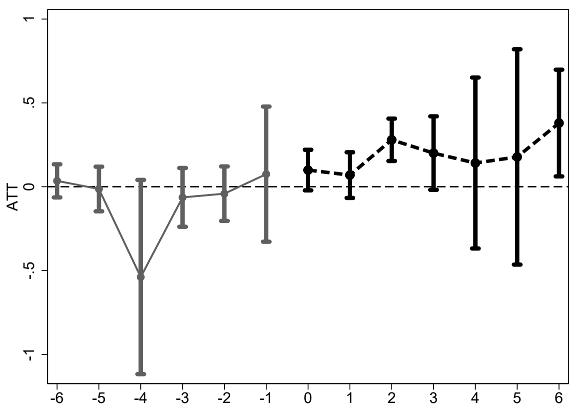
(f) LN(Population)

Notes: This Figure shows average treatment effects by length of exposure to electrification with 95% confidence intervals. The 0 line refers to the full pre-treatment estimates. The horizontal axis indexes the length of exposure to the treatment T=[-6,6]. Estimator is outcome regression DiD estimator with controls; standard errors are clustered at the locality level. Pre-treatment parameters are reported as "short differences" $(Y_t - Y_{t-1})$. Subfigure 4a shows the proportion of localities with non-zero night-lights (extensive margin). Subfigure 4b displays maximum radiance within ever lit localities (intensive margin). See Web Data Appendix for data sources.

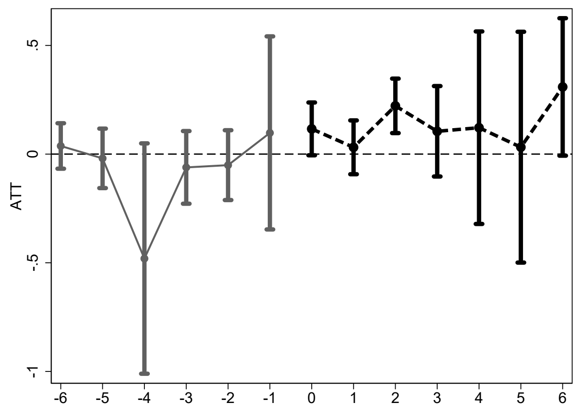
Figure 5: Household Access to Electricity and Electric Appliances Before/After Electrification



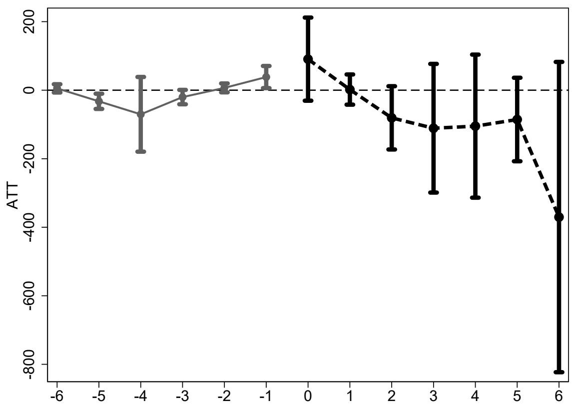
(a) Household has Electricity



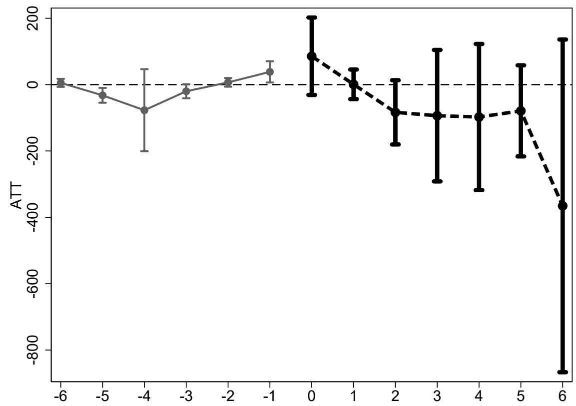
(b) Household Owns Television, Total Effect



(c) Household Owns Television, Spill-over Effect on Non-electrified Households



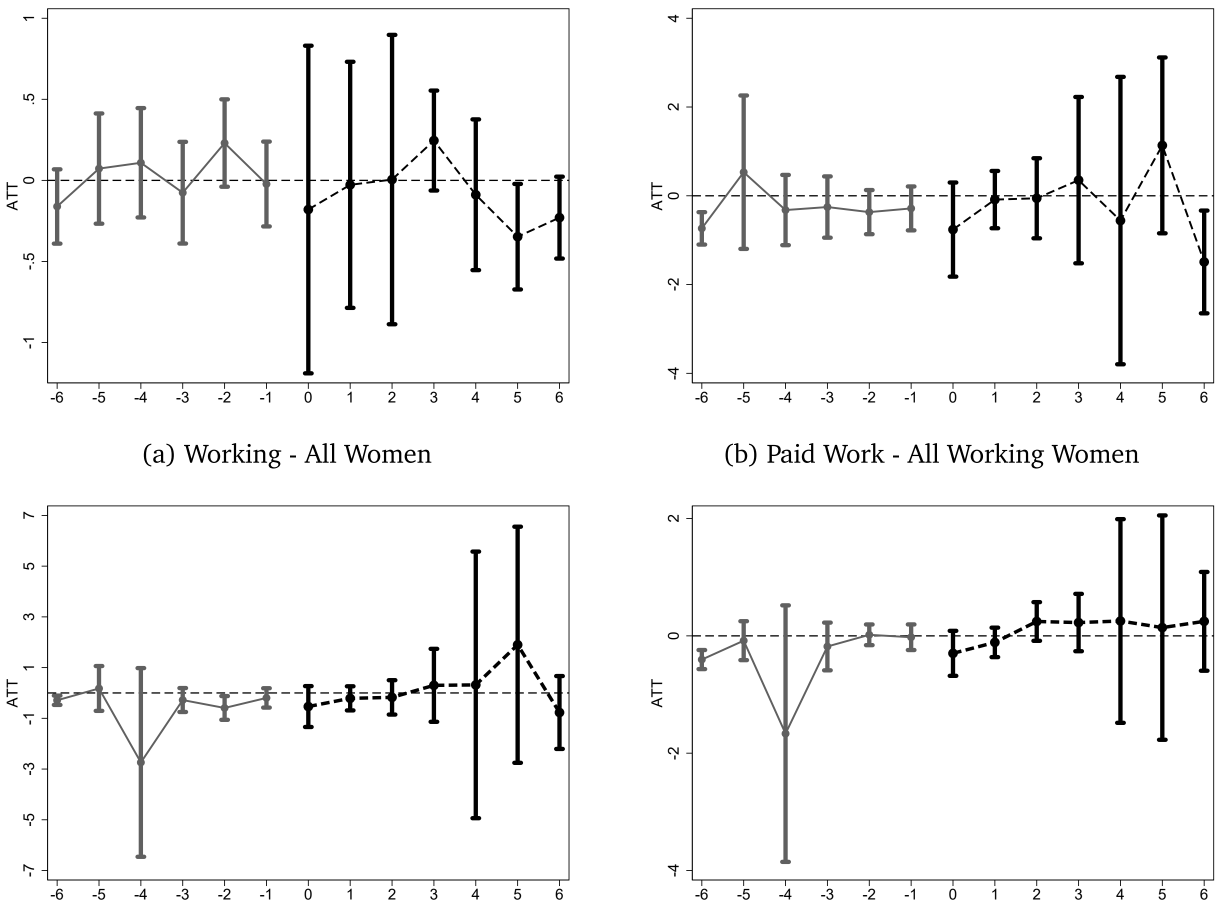
(d) Household's Time to Drinking Water (min), Total Effect



(e) Household's Time to Drinking Water (min), Spill-over Effect on Non-electrified Households

Notes: This Figure shows average treatment effects by length of exposure to electrification with 95% confidence intervals. The 0 line refers to the full pre-treatment estimates. The horizontal axis indexes the length of exposure to the treatment T=[-6,6]. Estimator is outcome regression DiD estimator with controls; standard errors are clustered at the locality level. Pre-treatment parameters are reported as "short differences". See Web Data Appendix for data sources.

Figure 6: Female Labor Force Participation Before/After Electrification



(a) Working - All Women

(b) Paid Work - All Working Women

(c) Paid Cash for Work - All Working Women

(d) Non-Agricultural Work - All Working Women

Notes: This Figure shows average treatment effects by length of exposure to electrification with 95% confidence intervals. The 0 line refers to the full pre-treatment estimates. The horizontal axis indexes the length of exposure to the treatment T=[-6,6]. Estimator is outcome regression DiD estimator with controls; standard errors are clustered at the locality level. Pre-treatment parameters are reported as "short differences". See Web Data Appendix for data sources.

Table 1: Covariates of Electrification 2008-2017

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var. / Estimator	Electrified 2008-2017 / Probit			Year of Electr. / Ordered Probit		
Sample	All Localities Not Yet Electrified in 2007			Only Localities Electrified 2008-2017		
Development Pole	0.81***	0.38***	0.42***	0.08	0.26**	0.28**
	(0.07)	(0.09)	(0.09)	(0.11)	(0.12)	(0.12)
LN(Population)	0.30***	0.30***	0.33***	0.07	0.17**	0.09
	(0.04)	(0.04)	(0.05)	(0.08)	(0.08)	(0.09)
DPI	4.44***			-2.56***		
	(0.39)			(0.56)		
Within 60km to grid 2007			0.13			-0.47***
			(0.10)			(0.16)
LN(Dist. to grid 2007 (km))			-0.08**			0.08
			(0.04)			(0.06)
LN(Dist. to dev. pole (km))			0.07**			-0.01
			(0.05)			(0.07)
LN(Dist. to MST (km))			-0.13***			0.21***
			(0.04)			(0.07)
Observations	7,672	7,342	7,332	502	502	502
Region FE	YES	YES	YES	YES	YES	YES
DPI indicators	NO	YES	YES	NO	YES	YES
Dev. Indicators	NO	YES	YES	NO	YES	YES
Pseudo R2	0.290	0.341	0.354	0.044	0.078	0.091

Notes: This table tests covariates of electrification. Unit of observation are localities. Localities electrified before 2008 were excluded. In col. (1)-(3), the dependent variable is a dummy indicating whether the locality was electrified between 2008 and 2017. In col. (4)-(6), the dependent variable is the year of electrification. Development pole is a dummy indicating whether the locality was identified as such by (Ministère des Mines, des Carrières et de l'Energie, 2008) due to a high DPI value and population in the catchment area. The Development Potential Index (DPI) is a summary measure of development indicators as shown in Web Appx. Table B.1 and predicted for all localities based on available data. The Minimum Spanning Tree (MST) was calculated using the proposed 33kV poles. "DPI indicators" refers to the joint inclusion of development indicators that enter the calculation of the DPI. "Dev. Indicators" specifies additional indicators including the square root of the number of primary school children and dummy variables for province and commune capitals. All regressions include region fixed effects. Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

Table 2: Average Treatment Effects by Length of Exposure to Electrification

Dependent Variable	Mean Dep.Var	Outcome regression DiD		Doubly robust DiD	
	T=[-1]	T=[0, 1]	T=[2, 6]	T=[0, 1]	T=[2, 6]
1. Luminosity, Ext. Margin	0.036	0.107***	0.421***	0.179***	0.485***
		(0.020)	(0.039)	(0.023)	(0.047)
2. Luminosity, Int. Margin	0.112	0.149***	0.464***	0.212***	0.630***
		(0.028)	(0.057)	(0.040)	(0.077)
3. AEPS Drinking Water, Ext. Margin	0.287	0.063***	0.129***	0.022	0.078*
		(0.018)	(0.035)	(0.017)	0.046)
4. N Public Primary Schools	1.446	-0.039	-0.062	-0.041	-0.202
		(0.039)	(0.093)	(0.055)	(0.159)
5. LN(N Primary School Students)	6.143	-0.010	-0.123	-0.003	-0.046
		(0.012)	(0.075)	(0.013)	(0.044)
6. N Electrified Schools	0.200	0.078*	0.100	0.097*	0.270*
		(0.042)	(0.081)	(0.058)	(0.140)
7. LN(Population)	7.772	0.073	0.124*	0.071	0.227
		(0.072)	(0.075)	(0.106)	(0.188)

Notes: This Table shows the average treatment effects over two time horizons i) short-term t=[0, 1] and ii) mid-term [2-6]. Unit of observation is the locality. Each cell represents a separate estimation model and shows the average treatment effect of electrification on the dependent variable. Using Stata's csdid command by Rios-Avila et al. (2021) we followed two estimation approaches. The first estimator is an outcome regression DiD estimator based on ordinary least squares. The second estimator is a doubly robust DiD estimator based on stabilized inverse probability weighting and ordinary least squares (Sant'Anna and Zhao, 2020). Luminosity combines two different measures: i) DN values and ii) radiance based on DMSP and VIIRS, respectively. The extensive margin of luminosity indicates any positive value in the two measures. The intensive margin takes the observed locality-specific difference between these two measures in 2012 and 2013 to merge them into one measure. AEPS Drinking Water indicates the existence of such water system in the locality. Population is only available for the years 2006 and 2015. Coefficients are average treatment effects relative to the period first treated, across all cohorts. Standard errors, clustered at the locality level, in parentheses. See Web Data Appendix for data sources.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: Household Asset Ownership - Average Treatment Effects (0-6 Years after Electrification)

Dependent Variable	*Mean Dep.Var.*	**Total Effect**		**Spill-over Effect**	
	T=-1	No HH Cont.	HH Cont.	No HH Cont.	HH Cont.
1. Electricity	*0.00*	0.08 ***	0.07 *		
		(0.03)	(0.04)		
2. Radio	*0.53*	-0.20	0.03	-0.22	-0.01
		(0.14)	(0.07)	(0.16)	(0.08)
3. Television	*0.12*	0.11 +	0.16 **	0.06	0.12 **
		(0.07)	(0.07)	(0.07)	(0.06)
4. Refrigerator	*0.00*	0.02 ***	0.02 **	0.01 *	0.00
		(0.01)	(0.01)	(0.00)	(0.01)
5. Piped Water	*0.00*	-0.02	-0.03	-0.02	-0.04
		(0.02)	(0.04)	(0.02)	(0.04)
6. Time to Water Source	*20.13*	-40.11 +	-48.82 +	-35.19	-43.01
		(25.41)	(33.34)	(25.72)	(33.63)
7. Finished Floor	*0.43*	-0.09	-0.10	-0.12	-0.13
		(0.16)	(0.13)	(0.16)	(0.13)

Notes: This table shows the average treatment effects based on Callaway and Sant'Anna (2021) for the time frame T=[0,6]. Unit of observation is the household. Each cell represents a separate estimation model and shows the average treatment effect of electrification on the dependent variable. The columns entitled "Total" calculate the average treatment effects for all households, while the columns entitled "Spill-over" calculate the average treatment effects only for households who do not have access to electricity. Using Stata's csdid command by Rios-Avila et al. (2021) we estimated outcome regression DiD estimator based on ordinary least squares. "Time to Water Source" is measured in minutes and gives the time it takes to go to the primary source of drinking water. The other dependent variables are all binary, taking the value 1 if the household owns (at least one of) the item. Household controls comprise the sex and age of the household head, as well as the share of adult women in the household who have any primary and secondary education, respectively. Standard errors, clustered at the locality level, in parentheses. See Web Data Appendix for data sources.
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15.

Table 4: Labor Market for Women - Average Treatment Effects (0-6 Years after Electrification)

	Mean Dep.Var.	**Total Effect**		**Spill-over Effect**	
	T=-1	No Educ. Cont. +No HH Cont.	Educ. Cont. +HH Cont.	No Educ. Cont. +No HH Cont.	Educ. Cont. +HH Cont.
1. Working	*0.77*	-0.06	-0.07	-0.05	-0.06
		(0.35)	(0.31)	(0.38)	(0.07)
2. Paid Work	*0.37*	-0.24	-0.20	-0.21	-0.17
		(0.72)	(0.28)	(1.01)	(0.28)
3. Paid Cash	*0.37*	-0.05	-0.04	-0.02	0.00
		(0.73)	(0.34)	(1.00)	(0.35)
4. Non Agricultural Work	*0.56*	-0.03	-0.03	-0.03	-0.03
		(0.17)	(0.17)	(0.17)	(0.17)

Notes: This Table shows average treatment effects based on Callaway and Sant'Anna (2021) for the time frame t=[0,2]. Unit of observation is the woman. Each cell represents a separate estimation model and shows the average treatment effect of electrification on the dependent variable. The columns entitled "Total" calculate the average treatment effects for all women, while the columns entitled "Spill-over" calculate the average treatment effects only for women in households which do not have access to electricity. Using Stata's csdid command by Rios-Avila et al. (2021) we estimated outcome regression DiD estimator based on ordinary least squares. The dependent variables are all binary. "Working" takes the value one if the woman reports being currently in work. "Paid Work" takes the value one if the woman receives payment for work - either in cash or in kind (the base is all working women.) "Paid Cash" takes the value one if the woman receives payment for work in cash. "Non-Agric. Work" takes the value one if the woman works in a sector other than agriculture (the base is again all working women.) Education controls include whether the woman has any primary education and whether she has any secondary education. Household controls comprise the sex and age of the household head, as well as the share of adult women in the household who have any primary and secondary education, respectively. Standard errors, clustered at the locality level, in parentheses. See Web Data Appendix for data sources.
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15.

Table 5: Child Health - Average Treatment Effects (0-6 Years After Electrification)

Dependent Variable	*Mean Dep.Var.*	**Total Effect**		**Spill-over Effect**	
	T=-1	No HH Cont.	HH Cont.	No HH Cont.	HH Cont.
Health Outcomes:					
1. Infant Death	*0.05*	0.04	0.03	0.03	0.03
		(0.07)	(0.07)	(0.08)	(0.07)
2. Child Death	*0.08*	0.00	-0.03	0.02	-0.01
		(0.06)	(0.06)	(0.06)	(0.06)
3. Height for Age Z-Score	*-1.19*	0.14	0.12	-0.07	-0.07
		(2.16)	(0.93)	(2.19)	(0.86)
4. Weight for Height Z-Score	*-1.10*	0.53	0.49	0.36	0.34
		(0.97)	(1.72)	(0.98)	(1.77)
Health Inputs:					
5. Prenatal Care	*0.47*	0.12	0.16	0.17	0.25
		(0.25)	(0.33)	(0.25)	(0.36)
6. Skilled Birth Attendance	*0.80*	0.16	0.25	0.19	0.30
		(0.20)	(0.23)	(0.16)	(0.26)
7. Medical Treatment for Fever	*0.66*	0.34	0.40	0.32	0.37
		(0.49)	(0.36)	(0.48)	(0.38)
8. BCG Vaccine	*0.97*	0.08	0.10	0.07	0.10
		(0.08)	(0.07)	(0.09)	(0.07)
9. Oral Polio Vaccine	*0.90*	0.13 +	0.15 **	0.12	0.15 *
		(0.08)	(0.08)	(0.10)	(0.08)
10. Inactivated Polio Vaccine (1st Dose)	*0.99*	0.07	0.06 +	0.07	0.06
		(0.05)	(0.04)	(0.06)	(0.05)
11. DPT Vaccine (1st dose)	*0.92*	0.05	0.06	0.05	0.07
		(0.11)	(0.10)	(0.12)	(0.10)

Notes: This Table shows the average treatment effects based on Callaway and Sant'Anna (2021). Each cell represents a separate estimation model and shows the average treatment effect of electrification on the dependent variable. Unit of observation is the live birth for rows 1,2,5,6,8,and 9; here treatment is defined as being born after electrification. For rows 3,4, and 7 unit of observation is the living child under the age of 5 and treatment is defined as living in an electrified community. For rows 10 and 11 the unit of observation is all children who survived (at least) to their first birthday and the treatment is being born after electrification. The columns entitled "Total" calculate the average treatment effects for all households, while the columns entitled "Spill-over" calculate the average treatment effects only for households who do not have access to electricity. Using Stata's csdid command by Rios-Avila et al. (2021) we estimated outcome regression DiD estimator based on ordinary least squares. Infant death takes the value 1 if the child died before their first birthday, child death is 1 if it died before its 5th birthday. Prenatal care takes the value 1 if the mother attended at least 4 prenatal appointments during pregnancy. Skilled birth attendance takes the value 1 if a healthcare professional was present at birth. Medical treatment for fever takes the value 1 if medical treatment was sought to treat a child's fever (base are all children under the age of 5 who had a fever in the past two weeks). Rows 8 to 11 refer to binary variables that take value one if the child has received the vaccination in question. The BCG and Oral Polio vaccine are administered at birth, thus the base for these variables are all live births. The first doses of the inactivated polio vaccine and the DPT vaccine are administered within the first year of live, thus the base for these are all children that survived to their first birthday (both living and dead at the time of survey). HH-Level controls include whether the woman has any primary education and whether she has any secondary education, as well as the sex and age of the household head. All specifications control for child's age in months (theoretical age at interview for deceased children), sex, and whether it was a multiple birth. Additionally, all specifications control for mother's age and religion, as well as community controls. Standard errors, clustered at the locality level, in parentheses. See Web Data Appendix for data sources.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$, + $p<0.15$.

Table 6: Infant Immunizations - Short and Long Term Treatment Effects

Panel A: Short Term Effects, T=[0,1]					
Dependent Variable	*Mean Dep.Var.*	**Total Effect**		**Spill-over Effect**	
	T=-1	No HH Cont.	HH Cont.	No HH Cont.	HH Cont.
1. BCG Vaccine	*0.97*	0.15 +	0.01	0.15	0.01
		(0.09)	(0.07)	(0.11)	(0.07)
2. Oral Polio Vaccine	*0.90*	0.19 **	0.06	0.20 *	0.06
		(0.09)	(0.07)	(0.11)	(0.07)
3. Inactivated Polio Vaccine (1st Dose)	*0.99*	0.03	0.02	0.03	0.01
		(0.07)	(0.05)	(0.07)	(0.06)
4. DPT Vaccine (1st dose)	*0.92*	-0.04	-0.01	-0.03	0.00
		(0.09)	(0.11)	(0.10)	(0.10)
Panel B: Long Term Effects, T=[2,6]					
Dependent Variable	*Mean Dep.Var.*	**Total Effect**		**Spill-over Effect**	
	T=-1	No HH Cont.	HH Cont.	No HH Cont.	HH Cont.
1. BCG Vaccine	*0.97*	0.00	0.20 *	-0.04	0.21 *
		(0.12)	(0.10)	(0.14)	(0.11)
2. Oral Polio Vaccine	*0.90*	0.05	0.25 **	0.02	0.26 **
		(0.14)	(0.11)	(0.15)	(0.12)
3. Inactivated Polio Vaccine (1st Dose)	*0.99*	0.11 **	0.11 **	0.12 *	0.12 **
		(0.05)	(0.05)	(0.07)	(0.06)
4. DPT Vaccine (1st dose)	*0.92*	0.18	0.17	0.18	0.18
		(0.16)	(0.13)	(0.18)	(0.14)

Notes: This Table shows the average treatment effects based on Callaway and Sant'Anna (2021). Unit of observation is the live birth for rows 1,2 and for rows 3,4 children who survived (at least) to their first birthday. The treatment is being born after electrification. The columns entitled "Total" calculate the average treatment effects for children from all households, while the columns entitled "Spill-over" calculate the average treatment effects only for children in households who do not have access to electricity. Using Stata's csdid command by Rios-Avila et al. (2021) we estimated outcome regression DiD estimator based on ordinary least squares. The BCG and Oral Polio vaccine are administered at birth, thus the base for these variables are all live births. The first doses of the inactivated polio vaccine and the DPT vaccine are administered within the first year of live, thus the base for these are all children that survived to their first birthday (both living and dead at the time of survey). HH-Level controls include whether the woman has any primary education and whether she has any secondary education, as well as the sex and age of the household head. All specifications control for child's age in months (theoretical age at interview for deceased children), sex, and whether it was a multiple birth. Additionally, all specifications control for mother's age and religion, as well as community controls. Standard errors, clustered at the locality level, in parentheses. See Web Data Appendix for data sources.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$, + $p<0.15$.

Table 7: Community Effects, Alternative Identification Strategies

	Treatment Group: Community Electrified 2008-2009						
	Control Group: Community Electrified 2015-2017						
Dependent Variable	Outcome Regr. DiD T=[1,5]	Simple OLS DiD 2x2 T=[1,5]	OLS t=[2010, 2014]	Reduced Form t=[2010, 2014]	2SLS t=[2010, 2014]	Kleibergen-Paap F	Montiel-Pflueger F
1. Luminosity, Ext. Margin	0.193***	0.38***	0.33***	-0.06***	0.46***	18.94	18.93
	(0.04)	(0.05)	(0.06)	(0.02)	(0.18)		
2. Luminosity, Int. Margin	0.13	0.342***	0.16	-0.05	0.39	4.12	4.23
	(0.08)	(0.07)	(0.10)	(0.03)	(0.31)		
3. AEP Water, Ext. Margin	0.14***	0.08	0.06	0.01	-0.08	19.02	19.01
	(0.04)	(0.06)	(0.07)	(0.02)	(0.15)		
4. N Public Primary Schools	-0.02	0.08	0.06	-0.06	0.44	13.38	12.98
	(0.09)	(0.08)	(0.12)	(0.06)	(0.50)		
5. N Electrified Schools	0.229**	0.10	0.29**	-0.06	0.47	18.81	18.81
	(0.10)	(0.11)	(0.12)	(0.04)	(0.31)		
6. Ln(Population)	0.05	0.138***					
	(0.03)	(0.05)					
7. Population growth			0.176**	-0.03	0.22	18.01	18.22
			(0.09)	(0.03)	(0.19)		

Notes: N (Control)=238 , N(Treatment)=59. Standard errors, clustered at the locality level, in parentheses. The last two columns report F-statistics testing for weak instruments, according to Kleibergen-Paap and Montiel Olea and Pflueger (2013). See Web Data Appendix for data sources.
*** p<0.01, ** p<0.05, * p<0.1.

Table 8: Effects of Community Electrification on Household Asset Ownership, Alternative Estimation Strategies

	Treatment Group: Community Electrified 2008-2009						
	Control Group: Community Electrified 2015-2017						
Dependent Variable	Outcome Regr. DiD	Simple OLS DiD 2x2	OLS	Reduced Form	2SLS	Kleibergen-Paap F	Montiel-Pflueger F
	T=[1,5]	T=[1,5]	t=[2010, 2014]	t=[2010, 2014]	t=[2010, 2014]		
1. Electricity	0.09	0.21***	0.23***	-0.04*	0.25*	16.43	17.33
	(0.17)	(0.05)	(0.06)	(0.02)	(0.13)		
2. Radio	-0.09	0.10	0.08	-0.08***	0.52***	16.45	17.37
	(0.12)	(0.07)	(0.07)	(0.01)	(0.14)		
3. Television	0.13	0.09**	0.10**	-0.04***	0.23**	16.43	17.33
	(0.23)	(0.04)	(0.04)	(0.01)	(0.09)		
4. Refrigerator	0.02	0.04***	0.02**	-0.01+	0.04*	16.43	17.33
	(0.02)	(0.01)	(0.01)	(0.00)	(0.02)		
5. Piped Water	-0.08	0.02**	0.01**	0.00*	-0.02+	16.45	17.37
	(0.18)	(0.01)	(0.01)	(0.00)	(0.01)		
6. Time to Water	-103.76+	6.74	-5.69	0.38	-2.79	11.24	13.16
	(63.72)	(5.92)	(6.38)	(1.41)	(9.85)		
7. Finished Floor	-0.14	0.16**	0.20***	-0.07***	0.43***	16.45	17.37
	(0.54)	(0.08)	(0.06)	(0.02)	(0.14)		

Notes: Unit of analysis is the household, unit of treatment is the locality. The difference in difference estimation is based on 94 localities, 20 of which are in the treatment group. OLS, reduced form, and 2SLSL estimations are based on 64 localities, 14 of which are in the treatment group. All estimations control for socio-economic background of the household, as well as community-level controls and region fixed effects. The last two columns report F-statistics testing for weak instruments, according to Kleibergen-Paap and Montiel Olea and Pflueger (2013). Standard errors, clustered at the locality level, in parentheses . See Web Data Appendix for data sources.
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15.

Table 9: Household Asset Ownership - Cross-Sectional Estimation of Treatment Effects

Treatment Group: Community Electrified 2008-2009 Control Group: Community Electrified 2015-2017 Cross-Sectional Analysis using Data from 2010, 2014						
Panel A: Total Effects						
Dependent Variable	Mean Dep. Var in Control	OLS	Reduced Form	2SLS	Kleibergen-Paap F	Montiel-Pflueger F
1. Mobile Phone	0.71	0.00	-0.03***	0.19**	16.45	17.37
		(0.04)	(0.01)	(0.09)		
2. Bank Account	0.07	0.09***	-0.01+	0.09*	11.24	13.17
		(0.02)	(0.01)	(0.06)		
Panel B: Spill-Over Effects (Non-Electrified Households Only)						
Estimation Method	Mean Dep. Var in Control	OLS	Reduced Form	2SLS	Kleibergen-Paap F	Montiel-Pflueger F
1. Mobile Phone	0.71	-0.03	-0.03***	0.20*	11.56	12.58
		(0.04)	(0.01)	(0.12)		
2. Bank Account	0.07	0.06***	-0.03***	0.19***	6.77	7.79
		(0.02)	(0.00)	(0.07)		

Notes: Unit of analysis is the household, unit of treatment is the locality. OLS, reduced form, and 2SLS estimations are based on 64 localities, 14 of which are in the treatment group. The dependent variables are all binary, taking the value 1 if the household owns at least one of the item. All estimations control for socio-economic background of the household, as well as community-level controls and region fixed effects. The last two columns report F-statistics testing for weak instruments, according to Kleibergen-Paap and Montiel Olea and Pflueger (2013). Standard errors, clustered at the locality level, in parentheses. See Web Data Appendix for data sources.
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15.

A Web Data Appendix

A.1 Electrification

Electrified Localities 1954-2017. We built a database of electrified localities and their respective years of electrification from various government reports. Our main source was the *Annuaire Statistique 2017* (INSD, 2018). We found a few errors and discrepancies that we rectified. For example, for the year 2013 the reported total number and itemized localities differed (N=51 vs N=31). For the year 2015, INSD (2018) reported only 16 electrified localities, whereas the 2016 statistical yearbook of the Ministére de l'Energie reported 125 ones. For those years, we used Ministére de l'Energie (2017) as ultimate source. This yielded a list of 660 localities that were connected to the electricity grid between 1954 and 2017.

We then retrieved the geographic coordinates of those localities using i) GEONet data of populated places in Burkina Faso (NGA, 2017), ii) an online map of places and (planned) electrification points as of 2006 (Ministère des Mines, des Carrières et de l'Energie, 2020) and iii) the 2019 georeferenced locality census. We matched observations by place name. Identifying African places is challenging. Alternative spellings are common; several localities had the same name. We therefore added information on the region, province, and commune on a subset of electrified localities that we obtained from Ministére de l'Energie (2017) and SONABEL (2018). Remaining ambiguities were resolved by mapping localities onto a map of the Burkina Faso electricity grid as of March 2018 prepared by the state energy provider SONABEL and published in UCF (2018). 20 localities would not match any place in our place name databases and electrification maps. In a final step, we improved precision of the coordinates by assigning the point location of electrified locality to the pixel with maximum population as indicated by 2015 Population Settlement Layer with 120m x 120m resolution (Facebook and CIESIN, 2019). Overall, we work with a dataset of 640 georeferenced localities.

Distribution Network. The 33kV power lines are the main distribution network. We digitized the 33kV transmission lines in 2007 from Ministère des Mines, des Carrières et de l'Energie (2008), the one in 2014 from Moner-Girona et al. (2016, Figure 5) and the one in 2016 from UCF (2018).

Electrification Strategy. MEPRED (2008) developed the methodology for the national electrification strategy. The same document listed the localities that were considered *development poles*. These are localities with good access to public goods and high growth potential as indicated by the development potential index (Web Appx. Table B.1) and a large population in the catchment area. In addition, for the complete set of Burkina Faso's localities the document lists the proposed connection methods: A - 33kV line, B - isolated and C - clustered diesel generators, D - Single-wire earth return (SWER) lines, and E - pre-electrification (i.e. not be connected to the grid). For A and D, the year of connection was proposed as well. Using the Geo-MST plugin in QGIS, we computed the *minimum spanning tree* (MST) that would connect all 312 development poles at minimum total network length. We then calculated for each electrified locality the distance to the MST (in km).

A.2 Other Data at the Locality Level

Administrative units. We extracted region, province and commune of the electrified locality using a commune boundary shapefile for Burkina Faso as of January 2007 from OCHA (2020).

Night-lights. We use cloud free satellite imagery of stable night-light from i) the Defense Meteorological Satellite Program (DMSP) and ii) the Visible Infrared Imaging Radiometer Suite (VIIRS). Wu et al. (2013) explained that differences in atmospheric absorption, solar altitude angle, terrain illumination, and sensor calibration cause large differences between data from the same year obtained by different DMSP satellites. We want to minimize measurement error as much as possible. We therefore refrained from averaging night-lights across satellites of and rely on imagery of one satellite per year. In particular, we use satellites F10, F12, F14, F15, F16, and F18 for the years 1992-93, 1994-96, 1997-99, 2000-03, 2004-09 and 2010-13 respectively. DMSP measures luminosity in digital numbers (DN), for the years 1992-2013, ranging between 0 and 63 (capped), available at 30 arc-second grid resolution (ca. 1x1 km). VIIRS measures luminosity in radiance in nano Watts per square centimeter per steradian, for the years 2012-2018, at 15 arc-second grid resolution.

We create two measures of luminosity based on the highest pixel value lit within the locality's polygon derived from Facebook and CIESIN (2019). First, we create a measure of extensive margin, a dummy variable indicating whether the locality was visible from outer space at all. Second, we create a measure of intensive margin. We only use ever lit localities. We merge the DMSP and VIIRS series based on their correlation in 2012 and 2013 predicting radiance from DN after running a locality fixed effects regression over ever lit localities.

In addition, we calculated mean pixel values, i) within the locality's polygon and ii) within buffers of 2 and 5 km around the highest lit pixel of the locality. Buffers may be less precise, because they potentially overlap the space of neighboring localities, thus wrongly assigning light that originated from somewhere else, similar to overglowing (Web Appx Figure B.3).

Population. We have population counts for the years 1996, 2006, 2015 and 2019. The 2006 data comes from MEPRED (2008) and is based on the census. It is available for all localities. We accessed the 2019 census data through the Ministère de l'Eau et de l'Assainissement (2020b). The source listed geographic coordinates at the quarter (Admin 5) level only for "rural" areas; we added latitude and longitude for the "urban" localities from NGA (2017). For 2015, we used the Burkina Faso - High Resolution Settlement Layer from Facebook and CIESIN (2019). The layer provides population estimates at a 1 arc-second resolution (approximately 30m). It is based on satellite images and building recognition, to which then population derived from sub-national census data was proportionally allocated. We followed a procedure similar to Baruah et al. (2020) and AFRICAPOLIS (2022). First, we removed single hamlets from the layer. Second, we defined the boundary of a locality as the continuous built-up area within 100 meters. Hence, we constructed a buffer of 100 meters and joined all overlapping polygons. Finally, we calculated the total population within those polygons for each electrified locality.

Roads. From the 2007 Michelin Map *741 Africa North & West* we digitized the network of asphalt roads. We then computed for each locality the distance to the nearest asphalt road (in km).

Schools. We obtained panel census data of primary schools for the years 2003-2013. The source informs about the type of school (public/private/denominational), the number of children entering school (Ministére de de l'Éducation Nationale et de l'Alphabétisation, 2019) and whether the school has electricity (Ministère de de l'Éducation Nationale et de l'Alphabétisation, 2019). The source does not report geographic coordinates, so we matched schools to localities by region, province, commune and place name. For each locality and year we calculated the number of electrified schools, public schools and primary school children. In addition, we used information in Ministère des l'Enseignements Secondaire et Supérieur (2014) to build a panel of private secondary schools. The source states the year of school opening. We therefore were able to reconstruct a panel based on the schools that existed in 2014. Using the school's region, province and village name, we georeferenced their locations.

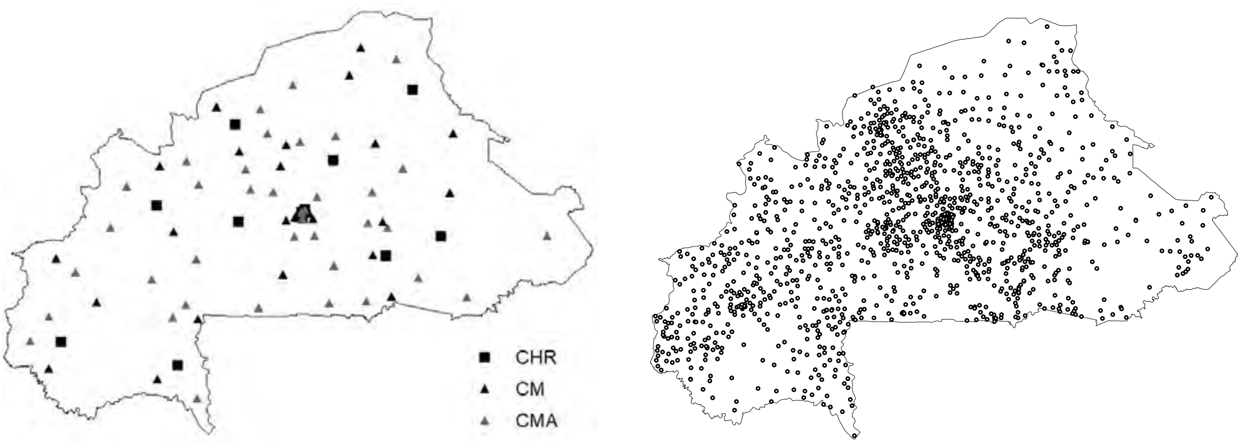
Irrigation. We digitized the location of 97 irrigation systems in 2008 from MEPRED (2008, Figure 5). In Burkina Faso, rain-fed agriculture dominates.

Drinking Water. The Ministère de l'Eau et de l'Assainissement (2020a) provides data on drinking water supply systems in the localities and the year when they were installed. The source informs about the number of water points (”borne fontaine”) and distinguishes between types of supply systems such as "Approvisionnement en eau potable" (AEP) and and "Poste d'eau autonome" (PEA). Both treat water to make it fit for human consumption and transport to consumption points. PEAs are autonomous water stations without a wider distribution network. For our purpose, we only use AEPs to be consistent with MEPRED's definition of drinking water access. One drawback of the data is that we do not know whether the systems were in operation every year and whether water pumps were powered by grid electricity.

Cattle Markets. The data is based on the 2019 Survey of Livestock Markets, which contains a sample of 45 major, permanent livestock markets in Burkina Faso (Table 1, Minot and Elahi, 2020).

Health Facilities. From Ministère de la Santé (2008) we digitized the names of localities with health facilities in 2007. The accompanying maps helped us georeferencing 1,402 out of the 1,445 facilities (or 97.0%). We have information on four types of facilities: i) CHR: Centre hospitalier régional (N=12); ii) CMA: Centre médical avec antenne chirurgicale (N=41); iii) CM: Centre médical (N=34); and vi) CSPS: Centre de santé et de promotion sociale (N=1,355) and v) dispensaries (N=86).[27] The CSPS are the first level of care and correspond to primary healthcare centers. Staffed with about three health workers and two unqualified volunteers, they are responsible for 5 to 23 villages, providing basic outpatient services, including maternity care and administering vaccination programs in the villages. The CM and CMA are the first referral level and are equipped with beds and surgical facilities (Marschall and Flessa, 2011). CHR are regional hospitals. We calculated the distance between each locality and health facility. We then assigned health facilities to the closest locality. The mean distance to CSPS and higher level health facilities is 2.5 km and 19.8 km (sd=3.6 and sd=13.4) respectively.

Figure A.1: Health Facilities in Burkina Faso in 2007



(a) CHR, CM, CMA

(b) CSPS

Notes: Data is from Ministère de la Santé (2008).

[27] We detected minor inconsistencies with the summary statistics reported in Ministère de la Santé (2008): N(CSPS)=1,268; N(CM)=33; N(CMA)=42; N(CHR)=12.

A.3 Matching DHS Clusters to Electrification Points

To ensure confidentiality, the DHS surveys masked the true location of households (Burgert et al., 2013). Geographic coordinates of urban and rural clusters were displaced up to a distance of 2 and 5 km respectively.[28] The displacement followed a "random direction, random distance" method, while ensuring that the newly assigned coordinates are located in the true administrative unit (=region). Overall, the method "produces a near uniform distribution, with an average displacement of 1.0 kilometers for urban areas and 2.5 kilometers for rural areas" (Burgert et al., 2013, p.11). This displacement of household coordinates introduces measurement error: Households that appear within/outside an electrified community may actually be located outside/within.

In an effort to reduce attenuation bias and improve the precision of the estimates we calculated *probability scores* that approximate the probability that a DHS cluster is located within a locality.

Step 1: We determined the geographic extent of the localities. The 2015 High Resolution Settlement Layer from the Facebook Connectivity Lab and Center for International Earth Science Information Network (Facebook and CIESIN, 2019) provided us with population estimates at a 1 arc-second resolution (ca. 30m) based on satellite imagery and computerized building recognition. We defined localities as contiguous populated settlements. Operationally, for urban places we took the boundary shapefile from Africapolis (Moriconi-Ebrard et al., 2016). For rural places, we merged all buildings within a 100 meters distance from each other and that formed a contiguous shape into one settlement.[29] We defined the electrified locality as the one which contains the electrification point.

Step 2: We calculated the population within 2 and 5 km of a DHS cluster.

Step 3: We calculated probability scores as the proportion of population within radius of the DHS

[28] Furthermore, a randomly-selected 1% of rural clusters is displaced by up to 10 km. This small proportion is unlikely to affect estimates. For the sake of simplicity, we therefore ignore this.

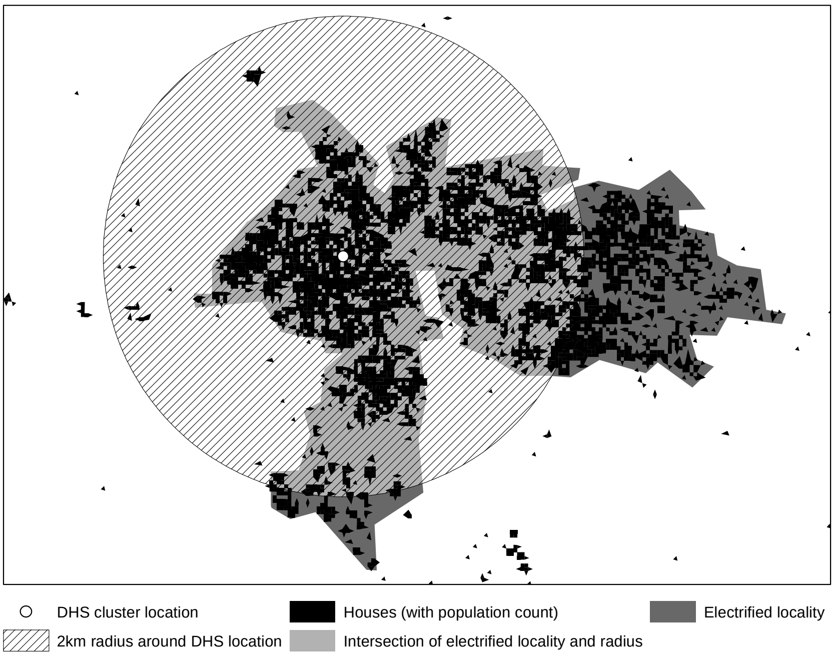
[29] This follows Moriconi-Ebrard et al. (2016) who defined an urban agglomeration as a continuously built-up and developed area, with less than 200 meters between two buildings. The criterion was then followed by de Bellefo et al. (2021). For rural places, the choice of 200 meter is too conservative, as it would include large low-populated spaces that encroach into the area of other villages (as indicated by the Housing Census and the list of electrification points).

cluster that falls in the electrified community. We apply the 2 km and 5 km radius to DHS clusters classified as urban and rural respectively.

$$probscore = \frac{pop(radius \cap electrified)}{pop(radius)}$$

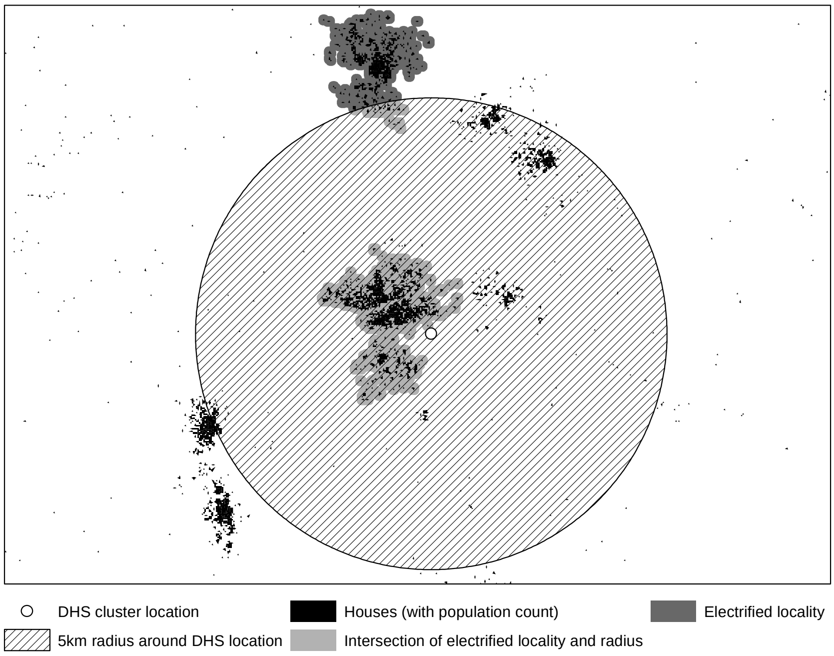
Figure A.2 illustrates the computation of the probability score for an urban DHS cluster. In this case the probability score is very high, because almost all of the population in the 2km radius around the cluster is located within the electrified locality. Figure A.3 illustrates the case of a rural DHS cluster for which the 5km radius overlaps with two localities. One locality lies completely in the radius and contains most of the population within the radius. The probability score for this electrified community is very high, meaning that the DHS survey was most likely carried out in this community. The second electrified community (at the top of the figure) contains only a small fraction of the population within the radius, yielding a much lower probability score.

Figure A.2: Calculating Probability Scores for DHS Clusters - Urban Setting



Notes: This figure illustrates the construction of the probability score for an urban DHS cluster. The gray area shows the electrified community. The white dot denotes the reported location of the DHS cluster (subject to random displacement). The shaded area shows a 2km radius around the reported DHS cluster location.

Figure A.3: Calculating Probability Scores for DHS Clusters - Rural Setting



Notes: This figure illustrates the construction of the probability score for a rural DHS cluster. The gray areas show two different electrified communities. The white dot denotes the reported location of the DHS cluster (subject to random displacement). The shaded area shows a 5km radius around the reported DHS cluster location.

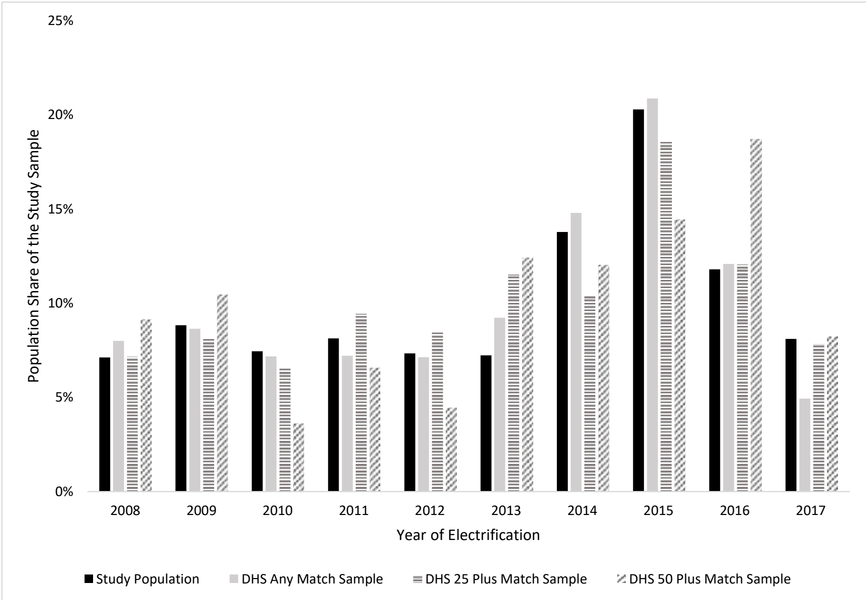
We use this methodology to match our electrified communities to data from DHS surveys from 1993, 1998, 2003, 2010, and 2014. As illustrated in Figure A.3, a DHS cluster can be matched to more than one electrified locality. For some matches the probability score is very low, meaning it is unlikely the DHS cluster actually lies in the electrified locality. We want to only consider DHS clusters with a high probability of representing our treatment areas.

What is an acceptable probability score? There is a trade off. Setting the probability score too high will reduce sample size and increase standard errors. Setting it too low will increase measurement error and attenuate the estimated coefficient. Complicating a theoretical derivation is the fact that probability scores may depend on settlement patterns. Localities that are finely delineated with few localities in close vicinity will have higher probability scores. In order to better understand the consequences of choosing certain assignment rules, we construct two samples of DHS clusters following two different definitions of "high probability". The first sample considers all matches with a probability score of at least 50% (the "50% plus sample"). In the second sample we consider all matches with a probability score of at least 25%, provided it is the most likely match for the DHS cluster (the "25% plus sample"). This ensures that within both samples, DHS clusters are assigned to one electrified locality. For example, imagine a DHS cluster with two matches - one with a probability score of 45% and one with 30%. For the second sample we would include the former match, but exclude the latter. For comparison, we constructed an "any match sample" that includes all DHS clusters that can be assigned to an electrified locality (provided it is the most likely match).

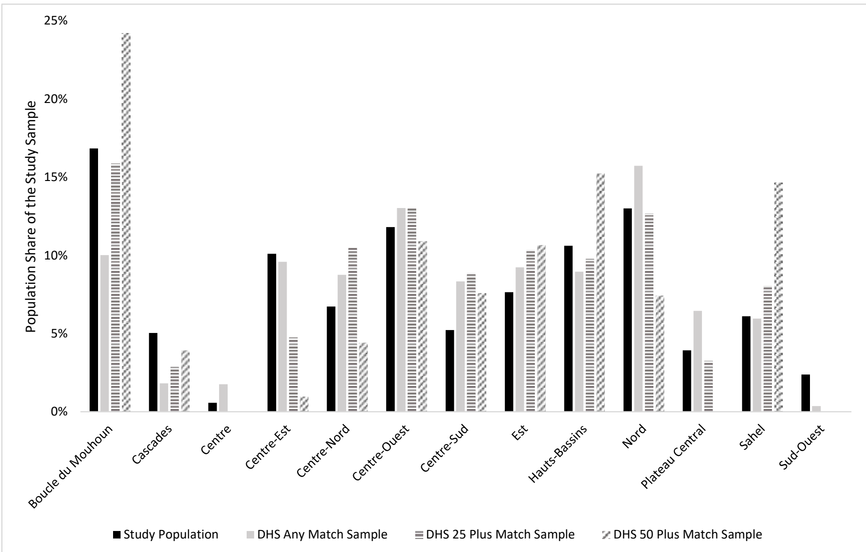
From census data, we know the population of each electrified locality and how it evolves over time. We use this information to check whether our two DHS samples (the "25% plus sample", and the "50% plus sample") are representative of the study population. Figure A.4 illustrates this exercise for the year 2010. The black bars show the true distribution of the study population across treatment times (years of electrification) and across administrative regions. The grey and dashed bars show the distribution of three possible DHS samples (adjusted for DHS population weights). The distributions of the "25% plus sample" are similar to those of the true population, while the distribution of the "50% plus sample" is significantly different. This could be due over-sampling of larger towns/ exclusion of smaller towns in the "50% plus sample". The "25% plus sample" outperforms the "Any Match sample" with respect to the spatial distribution, Figure A.4b. Thus, we chose the "25% plus sample" for our analysis.

For the analysis, we are interested in localities that were electrified from 2008 onward. Table A.1 illustrates the number of DHS clusters matched to electrified localities for each treatment year. This includes only DHS clusters where the probability score is at least 25% (with each cluster assigned only to the most likely electrification point). We exclude the 1998 survey round from our analysis, because coverage is very low, owing to a change in the DHS sampling strategy that year as shown in Figure A.5. In 1993, 230 clusters were sampled across the country. In 1998 a similar number of clusters were sampled (2010), but this time they were more strongly focused on the larger urban centers (especially Ouagadougou and Bobo-Dioulasso). This leads to a much lower

Figure A.4: Population Distributions in the full Study Population and various DHS Study Samples (2010)



(a) Population Distributions across Treatment Times



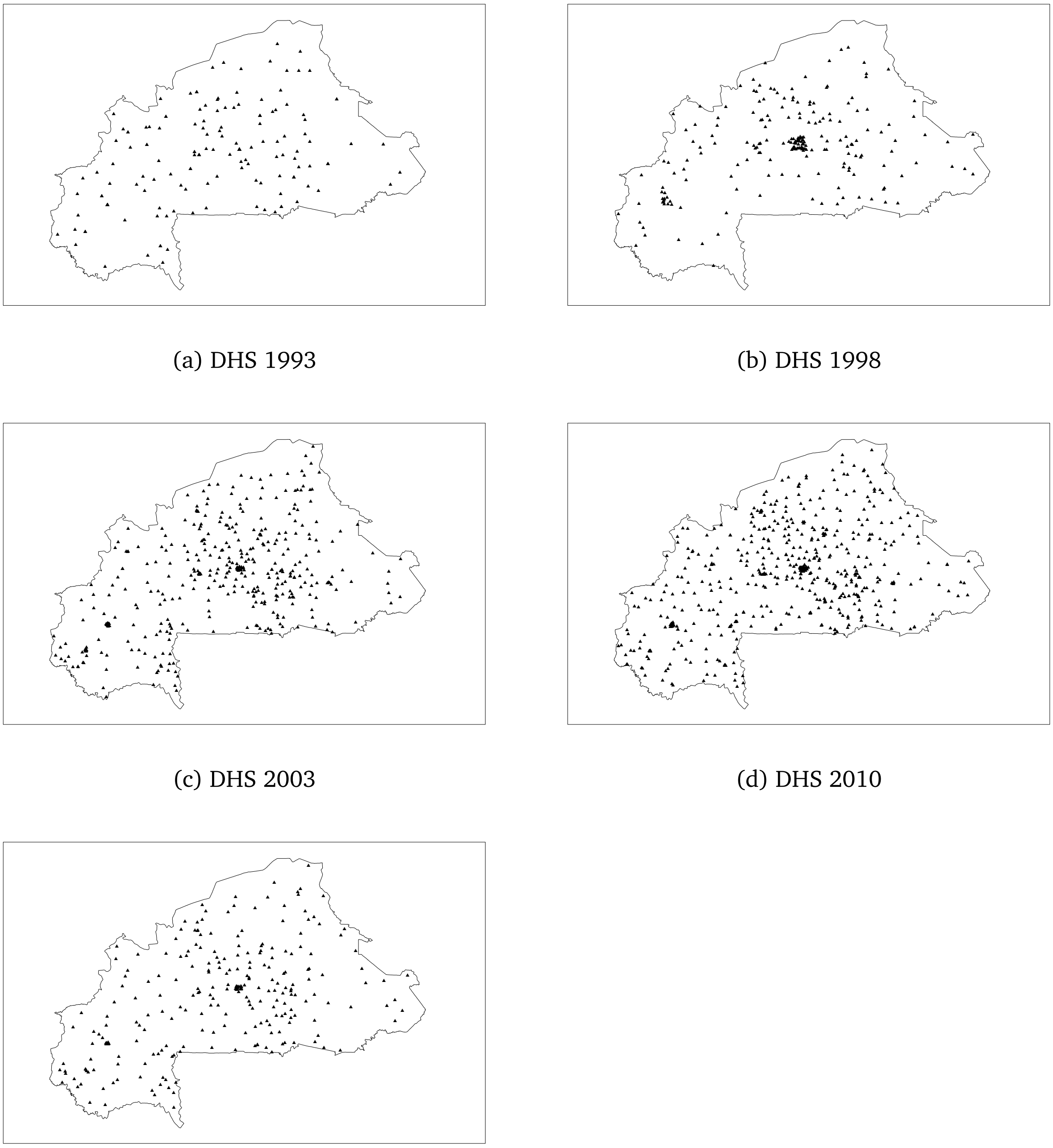
(b) Population Distributions across Administrative Regions coverage of the smaller towns that make up our sample for analysis. In 2003, the DHS sampling continued to focus on large urban areas, but the overall number of clusters was increased to 400 - once again providing good coverage in our areas of interest. The survey carried out in 2014 was a Malaria Indicator Surveys (MIS) with a smaller sample.

Table A.1: Number of matched DHS & LSMS Clusters by Year of Electrification

Year of electrification	Survey & Round						
	DHS1993	DHS1998	DHS2003	DHS2010	MIS2014	LSMS2014	Total
2008	6	3	7	6	5	3	30
2009	5	1	12	7	5	3	33
2010	4	1	5	4	3	2	19
2011	3	0	3	6	5	7	24
2012	7	0	2	6	1	8	24
2013	0	1	8	7	3	8	27
2014	2	2	9	8	7	14	42
2015	8	2	12	15	7	10	44
2016	5	0	1	9	3	7	25
2017	5	3	4	4	0	5	21
Total	45	13	63	72	39	67	232

Notes: This includes all DHS clusters with a probability score of at least 25%, assigned to their most likely electrified locality.

Figure A.5: Geographic Distribution in DHS clusters



(a) DHS 1993

(b) DHS 1998

(c) DHS 2003

(d) DHS 2010

(e) DHS 2014

B Web Appendix - Tables & Figures

Table B.1: Calculation of the Development Potential Index (DPI)

Component	Weight	Criteria	Weight	Indicator	Value
Health	1/3	Health Facilities	1/2	Medical Center equipped with surgical facilities (CMA)	1
				Medical Center (CM)	0.6
				Health and Social Promotion Center (CSPS)	0.4
				Dispensary or Maternity	0.2
		Access to Water	1/2	Simplified drinking water supply system (AEPS)	1
				Borehole	0.5
				Well	0.2
Education	1/3	Non-formal education	1/3	Existence of a training structure (CEBNF)	1
				Literacy facility (CPAF)	0.6
		Formal Education	2/3	Higher institution	1
				Vocational training establishment	0.5
				Secondary School	0.5
				Primary School	0.3
Local Economy	1/3	Population	1/4	More than 5,000 inhabitants	1
				2,501-5,000 inhabitants	0.5
				1,001-2,500 inhabitants	0.2
		Agriculture	1/4	Cattle market	1
				Irrigation	0.7
				Cereal Bank	0.5
		Markets	1/6	Daily Market	1
				Regular Market	0.5
				Occasional Market	0.2
		Distance to nearest Asphalt Road	1/6	0 km	1
				0-10 km	0.5
		Financial institutions	1/12	Bank	1
				Microfinance	0.75
		Telephone line	1/12	Available	1

Notes: This table shows the elements and weights of measures that enter into calculating the Development Potential Index, which is used to identify localities to be electrified.
Source: Ministère des Mines, des Carrières et de l'Energie (2008).

Table B.2: Summary Statistics, Table 1

	Table 1, col. 1-3			Table 1, col. 4-6		
VARIABLES	N	mean	sd	N	mean	sd
Electrified 2008-2017	7,784	0.072	0.258	558	1	0
LN(Population)	7,708	6.992	0.950	538	7.934	0.740
DPI	7,708	0.078	0.077	538	0.185	0.111
Development Pole	7,730	0.070	0.255	506	0.441	0.497
CMA	7,782	0.002	0.042	558	0.016	0.126
CM	7,782	0.002	0.041	558	0.013	0.111
CSPS	7,782	0.143	0.350	558	0.509	0.500
Dispensary	7,782	0.003	0.057	558	0.009	0.094
Missing value in AEPS 2007	7,784	0.118	0.323	558	0.116	0.321
AEPS Drinking Water 2007	7,782	0.029	0.167	558	0.168	0.375
Private Secondary School 2007	7,782	0.003	0.053	558	0.020	0.139
Primary School	7,784	0.572	0.495	558	0.841	0.366
Cattle Market	7,782	0.004	0.066	558	0.029	0.167
Irrigation 2007	7,782	0.008	0.088	558	0.013	0.111
Asphalt 0 km 2007	7,784	0.051	0.220	558	0.154	0.361
Asphalt 0-10 km 2007	7,784	0.162	0.369	558	0.185	0.388
SQRT (N Primary students)	7,782	7.628	7.586	558	16.62	9.022
Regional capital (Admin2)	7,782	0.001	0.023	558	0.005	0.073
Provincial capital (Admin3)	7,782	0.033	0.180	558	0.292	0.455
Within 60km to grid 2007	7,784	0.541	0.498	558	0.699	0.459
LN(Dist. to grid 2007 (km))	7,418	3.738	1.105	558	3.272	1.414
LN(Dist. to dev. pole (km))	7,428	3.137	0.944	558	2.694	1.188
LN(Dist. to MST (km))	7,428	2.727	1.272	558	2.096	1.494

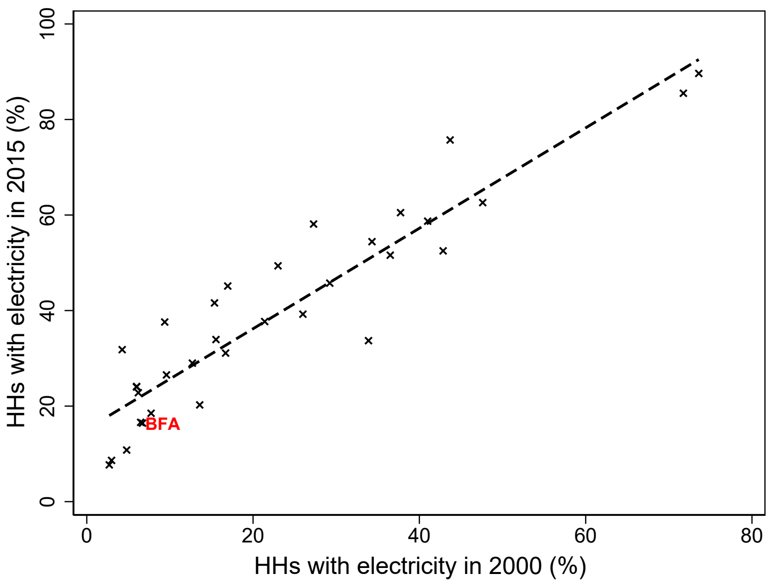
Notes: This Table reports summary statistics for Table 1. Variables are expressed in terms of their contribution to the DPI as shown in Web. Appx. Table B.1. For example, if a locality had a CMA, CM, CSPS as well as a dispensary, only the CMA indicator variable would take on the value 1.

Table B.3: Household Asset Ownership - Short- and Medium-Term Treatment Effects

Panel A: Short-Term Effects, T=[0,1]					
	Mean Dep.Var.	**Total Effect**		**Spill over Effect**	
	T=-1	No HH Cont.	HH Cont.	No HH Cont.	HH Cont.
1. Electricity	*0.00*	0.05 **	0.05 **		
		(0.02)	(0.02)		
2. Radio	*0.53*	-0.03	0.05	-0.05	0.04
		(0.07)	(0.08)	(0.07)	(0.08)
3. Television	*0.12*	0.04	0.08 +	0.03	0.08 +
		(0.05)	(0.05)	(0.04)	(0.05)
4. Refrigerator	*0.00*	0.01 **	0.02 **	0.01	0.01 *
		(0.01)	(0.01)	(0.00)	(0.01)
5. Piped Water	*0.00*	0.00	0.01	0.00	0.01
		(0.00)	(0.00)	(0.01)	(0.00)
6. Time to Water Source	*20.13*	33.23	45.91	31.90	43.10
		(26.63)	(34.38)	(26.20)	(33.27)
7. Finished Floor	*0.43*	0.05	0.00	0.05	0.00
		(0.08)	(0.09)	(0.08)	(0.09)
Panel B: Medium-Term Effects, T=[2,6]					
	Mean Dep.Var.	**Total Effect**		**Spill over Effect**	
	T=-1	No HH Cont.	HH Cont.	No HH Cont.	HH Cont.
1. Electricity	*0.00*	0.10 **	0.08		
		(0.05)	(0.07)		
2. Radio	*0.53*	-0.34	0.01	-0.39	-0.05
		(0.26)	(0.09)	(0.29)	(0.12)
3. Television	*0.12*	0.17 +	0.24 **	0.09	0.16 *
		(0.11)	(0.10)	(0.11)	(0.09)
4. Refrigerator	*0.00*	0.02 **	0.02	0.01	0.00
		(0.01)	(0.01)	(0.00)	(0.01)
5. Piped Water	*0.00*	-0.03	-0.07	-0.05	-0.09
		(0.04)	(0.07)	(0.04)	(0.07)
6. Time to Water Source	*20.13*	-104.90 ***	-136.08 ***	-97.39 **	-126.64 **
		(36.89)	(48.72)	(38.95)	(51.65)
7. Finished Floor	*0.43*	-0.22	-0.20	-0.28	-0.25
		(0.25)	(0.20)	(0.27)	(0.21)

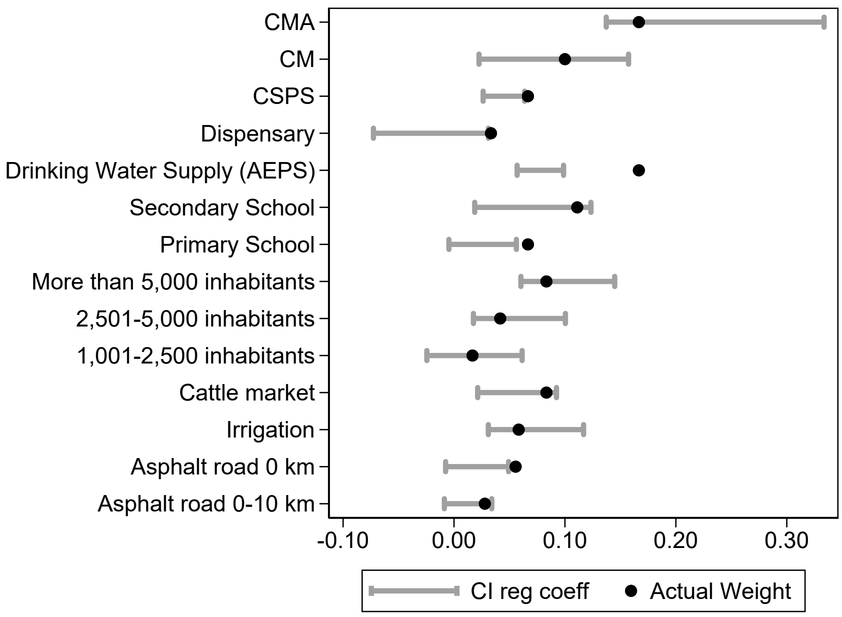
Notes: This Table shows the average treatment effects based on Callaway and Sant'Anna (2021). Unit of observation is the household. The columns entitled "Total" calculate the average treatment effects for all households, while the columns entitled "spill over" calculate the average treatment effects only for households who do not have access to electricity. Using Stata's csdid command by Rios-Avila et al. (2021) we estimated outcome regression DiD estimator based on ordinary least squares. "Time to Water Source" is measured in minutes and gives the time it takes to go to the primary source of drinking water. The other dependent variables are all binary, taking the value 1 if the household owns (at least one of) the item. Household controls comprise the sex and age of the household head, as well as the share of adult women in the household who have any primary and secondary education, respectively. Standard errors in parentheses. See Web Data Appendix for data sources.
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15.

Figure B.1: Electrification in Comparative Perspective, Sub-Saharan Africa, 2000-2015



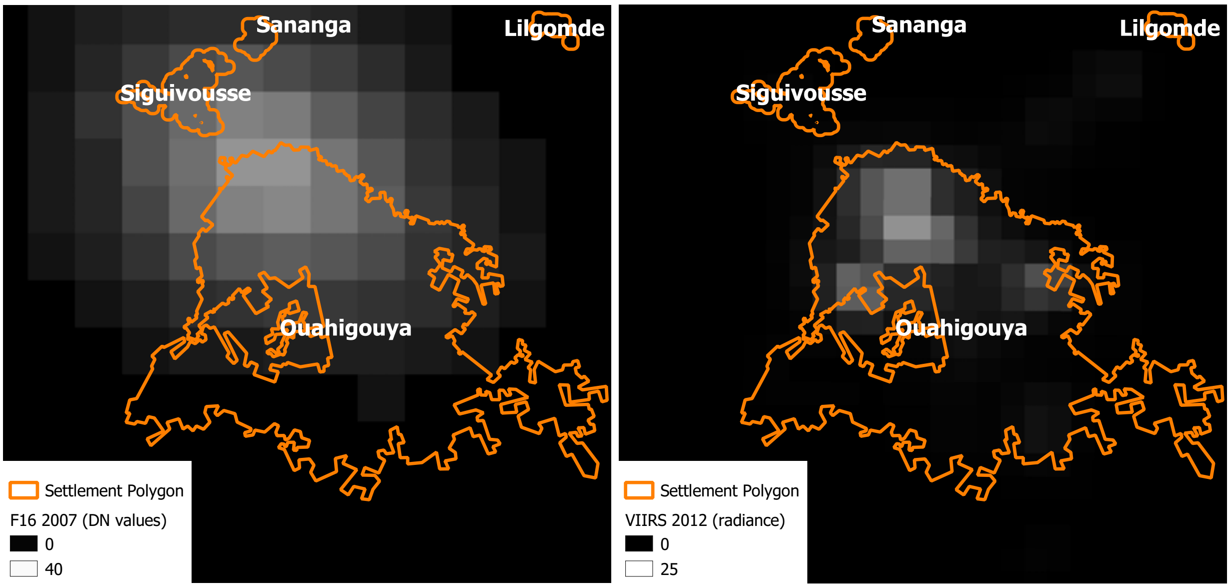
Notes: This Figure shows the percentages of households with electricity in 2000 and 2015 across sub-Sahara African countries (World Bank, 2020b). The dashed line represents the linear fit. Burkina Faso was among the lower tercile of countries and it roughly stayed there in 2015 despite of an increase from 7% to 19%. On average, the percentage of households with electricity increased in sub-Saharan African countries by 1% per year.

Figure B.2: Correlation between DPI Weights and Regression Coefficients



Notes: This Figure shows the weight given to the indicators in the calculation of the Development Potential Index (black circle) and the 95% confidence interval from a multivariate regression of the indicators on non-missing DPI Index values as reported in Ministère des Mines, des Carrières et de l'Energie (2008). Sample excludes localities that were connected to the grid before 2008. An overlap between the weight and the confidence interval indicates that our data matches the applied weights when predicting the DPI values for those localities for which Ministère des Mines, des Carrières et de l'Energie (2008) does not report DPI Index values. N=535 localities. R^2=0.37.

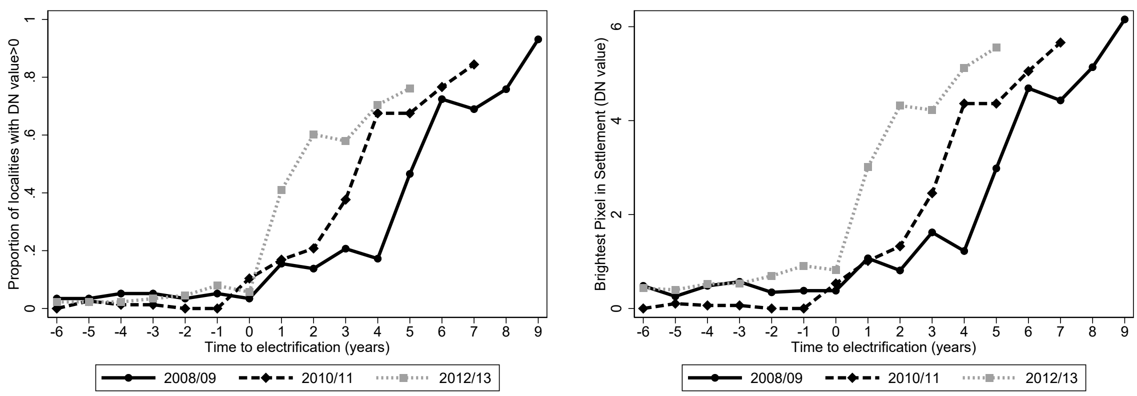
Figure B.3: Overglowing - DMSP versus VIIRS



(a) Night-lights DMSP 2008 (b) Night-lights VIIRS 2012

Notes: This Figure shows the settlement boundaries of the electrified localities overlaid over the raster of night-lights. Subfigure B.3a displays the DMSP raster for 2008, whereas Subfigure B.3b displays the VIIRS raster for 2012. Ouahigouya is a city with more than 70,000 inhabitants in 2006 (Institut National de la Statistique et de la Démographie, 2011); it was electrified in 1964. In contrast, Siguivousse and Sananga are small towns with about 2,000 inhabitants in 2006; both were electrified in 2009. DMSP is prone to over-glowing: light that probably originated from the large city of Ouahigouya extend into the settlement boundaries of Siguivousse and Sananga, about 2 km from Ouahigouya.

Figure B.4: Electrification and Luminosity Using the Harmonized Data by Li et al. (2020)



(a) Luminosity, Extensive Margin (b) Luminosity, Intensive Margin

Notes: Balanced panel. This Figure shows "harmonized" DN values in the run-up and after electrification. The data was prepared and made available by Li et al. (2020) and is available for the years 1992-2018. Their data integrates DMSP and VIIRS in the year 2014. The graphs plot three types of localities: those that received access to electricity in i) 2008/2009, ii) 2010/11 and iii) 2012/2013. Subfigure B.4a shows the proportion of lit localities (extensive margin). N(localities electrified in 2008/2009, 2010/11, and 2012/13)=58, 81 and 90. Subfigure B.4b shows the mean of the brightest pixel within the settlements (intensive margin). The patterns point to an artificial increase in luminosity due to switching from DMSP to VIIRS in 2014.