When does electrification work? Evidence from Sub-Saharan Africa*

Torsten Figueiredo Walter** Niclas Moneke**

June 30, 2026

Half a billion people in rural Sub-Saharan Africa remain without electricity. We document that many grid-connected villages in the region have either near-zero or near-full electricity adoption, indicating an indivisibility in local electrification. We highlight one source of indivisibility, the local low-voltage network, and how anchor customers help overcome it. Exploiting the roll-out of a country-wide electrification program in Zambia, we find that widespread electricity adoption upon grid arrival and ensuing development effects are confined to villages with pre-existing energy-dependent firms. Further evidence suggests that both cost and demand externalities of connecting these firms explain the local success of electrification.

*We thank Ben Faber, Paco Buera, Gharad Bryan, Taryn Dinkelman, Kevin Donovan, Doug Gollin, Marco Gonzalez-Navarro, Karen Macours, Craig McIntosh, Pritchard Mukuka, Simon Quinn, Lupupa Siluapwa, Daniel Sturm, Anant Sudarshan, Chris Woodruff and audiences at PSE, Oxford, Linkages in Labour Economics Workshop, LSE Environment Week, UEA Washington DC, Y-Rise Kingston, AFES Nairobi, Hitotsubashi, NUS, SSE, STEG Annual Conference Abu Dhabi, EEA Barcelona, Infra4Dev Guangzhou, Harvard Urban Workshop, WB, Berkeley ARE and BQSE for helpful comments. Gregory Cuff, Felix Johnson, Hassane Meite, Ana-Maria Radu, Mackenzie Tan Phei Huey, Kristina Sisiakova and Likun Tian provided outstanding research assistance. Financial support from the IGC is gratefully acknowledged.

**Figueiredo Walter: t.f.walter@nyu.edu, Social Science Division, New York University Abu Dhabi, https://sites.google.com/site/torstenfigueiredowalter/. Moneke: niclas.moneke@economics.ox.ac.uk, Department of Economics, University of Oxford, https://niclasmoneke.com.

1 Introduction

Sub-Saharan Africa is the least electrified region in the world. Despite large-scale grid expansion in recent decades, more than half a billion people remain without access to electricity. Meanwhile, electrification efforts across the continent persist. Mission 300, an ongoing initiative by the World Bank and the African Development Bank, aims to connect 300 million people in the region to electricity by 2030 – with an explicit objective to enable "jobs, business, [...], and opportunity" (World Bank, 2026b).

Whether electrification can generate such economic benefits remains contested. The literature is divided between papers documenting seemingly transformative effects of electrification (Dinkelman, 2011; Rud, 2012; Lipscomb et al., 2013; Fried & Lagakos, 2021; Kassem, 2024) and others finding moderate effects at best (Lenz et al., 2017; Lee et al., 2020b; Burlig & Preonas, 2024; Masselus et al., 2025; Kassem et al., 2026). In light of the evidence supporting either perspective, it seems natural to ask: When does electrification work?

This paper provides a new perspective on this question. We unveil stark differences in the impact of rural electrification across locations and highlight the importance of anchor customers for adoption and development. First, we document that many grid-connected villages in Sub-Saharan Africa have either near-zero or near-full electricity adoption, pointing to an indivisibility in local electrification. Second, we highlight one source of indivisibility, the local low-voltage network, and the role of anchor customers in overcoming it. We do so by exploiting the roll-out of a country-wide electrification program in Zambia to show that widespread electricity adoption upon grid arrival and ensuing development effects are confined to villages with pre-existing energy-dependent firms as anchor customers. Finally, we provide suggestive evidence that both cost and demand externalities of connecting these firms explain the local success of electrification.

A stunning 75 million people in rural Sub-Saharan Africa live in locations with grid access, but their households do not have an electricity connection. While existing work points to the low willingness or ability to pay of marginal households in these locations (Lee et al., 2020b; Masselus et al., 2025), we propose an alternative explanation: the absence of local low-voltage networks in many villages that are connected to the medium-voltage network – a phenomenon sometimes referred to as 'ghost electrification'. Consistent with this hypothesis, we document a bimodal adoption pattern across grid-connected villages in nine Sub-Saharan African countries at today's electrification frontier. Electricity adoption rates in such villages are often either close to zero or close to 100%.

Why would a local low-voltage network exist in some villages, but not in others? We argue that anchor customers are crucial for the emergence of local low-voltage networks. Their construction represents a high fixed cost for the utility, especially in rural areas where low density at least doubles fixed costs per-user compared to urban areas (Trimble et al., 2016). Without anchor customers that consume large amounts of electricity, utilities struggle to recover these high fixed cost. This is particularly true in rural Sub-Saharan Africa where the majority of potential customers are households with minimal demand for electricity. Hence, utilities are much more inclined to invest in local low-voltage networks in villages with potential anchor customers.

To provide empirical support for this conjecture, we examine the effects of electrification in rural Zambia. In this context, grain mills, which are highly energy-dependent firms, serve as anchor customers. We document that adopting electricity is highly profitable for millers, who commonly rely on diesel to power their operation in the absence of the grid. As a result, mills contribute a disproportionate share of electricity demand compared to all other end-users in rural Zambia. Exploiting the particular design of the Zambian Rural Electrification Master Plan (REMP) and the resulting gradual expansion of the grid across villages from 2009 until today, we estimate the causal effect of electrification on electricity adoption and economic development, separately in villages with and without grain mills at baseline.

We develop a novel approach to accurately track the arrival of the medium-voltage grid across villages over time. Since the REMP mandates the connection of public facilities in villages reached by the grid, we can reconstruct the expansion of the grid from annual reports of headmasters on the power source of their school in the school census. By linking and geo-locating information on more than 14,300 schools over 16 years, we obtain a highly granular picture of the expansion of the grid across rural Zambia. We verify the accuracy of our electrification measure using several alternative ones, including a subset of grid construction records and a partial cross-section of the medium-voltage grid at endline from the utility.

Does connecting to the grid generate growth and economic development in rural Zambia? To measure the effects of electrification on electricity adoption and local economic development, we combine satellite-based data of nighttime luminosity and built-up area with population census and household survey data on electricity usage, education and sectoral employment that we geo-identify to the level of the village. Additionally, we conduct an endline survey in 2025 with key informants across villages targeted by the REMP to obtain endline information on the number of households, number of businesses and the extent of the low-voltage network within villages. Crucially, to differentiate between villages with and without a mill at baseline, we leverage a complementary dataset collected alongside the 2010 population census that captures the universe of businesses and public facilities across the country.

Our empirical strategy makes use of the particular design of the REMP in 2009. Following the identification of 1,200 villages as so-called rural growth centres (RGCs), these RGCs were prioritised based on expected demand for electricity and subsequently divided into 180 project packages. Each package was formed by a high-priority RGC as the primary target and the surrounding RGCs falling between said target and the nearest existing electric substation. Packages themselves were prioritised for construction based on basic financial indicators. As a result, RGCs of similar expected demand located between target RGCs and substations that belong to different packages were electrified at different times due to differences in priority across packages. Our analysis focuses only on this subset of 784 'inconsequential' RGCs which are spread throughout all ten provinces of Zambia. Of these, 186 inconsequential RGCs had at least one grain mill in their vicinity at baseline in 2010, whereas the remainder did not. Finally, at the end of our sample, almost half of inconsequential RGCs did receive a connection to the grid, with electrification progressing at comparable rates across villages with and without mills.

We employ a staggered difference-in-differences approach to compare changes in outcomes between inconsequential RGCs that are electrified at different points in time. To assess the difference in effects between RGCs with and without a mill at baseline, we do so separately in the two types of RGCs.

We provide three findings: first, we show that the average effects of rural electrification on adoption of electricity are mixed at best, confirming the existing evidence of modest to nil effects on household adoption, usage and economic benefits derived from usage in Kenya (Lee et al., 2020b; Kassem et al., 2026), Rwanda (Lenz et al., 2017; Masselus et al., 2025), India (Burlig & Preonas, 2024) and elsewhere.

Second, we highlight that these average effects mask stark heterogeneity: RGCs with pre-existing mills display large, positive increases in nighttime luminosity, while no statistically significant increase can be detected in RGCs without pre-existing mills.[1] We confirm this core result also in an event study design, highlighting that RGCs with mills have an almost five times larger luminosity response than RGCs without mills when compared over time against their respective not-yet electrified RGC counterparts. We fail to detect pre-trends for either the group of RGCs with mills or the one without.

The adoption documented in satellite-based nightlights can be documented at the household-level, too: households in RGCs with mills are 23pp more likely to report using grid-electricity for lighting and 6pp more likely to use electricity as source of energy for cooking. This adoption response is mirrored by households reporting statistically significantly less reliance on solar lighting and wood or charcoal usage for cooking. The substitution of solar for grid-electric lighting resonates well with previous findings in the literature on households' strong relative preference for grid-electricity relative to any off-grid ('stepping stone') technology when the former is available via a low-voltage network (Burgess et al., 2023). No comparable increase in adoption can be detected in RGCs without mills.[2]

Third, we find large development effects of rural electrification that are largely confined to villages with pre-existing mills, i.e., productive uses of electricity. We leverage our endline survey of RGCs to provide a fifteen-year, long difference comparison between 2010 population and housing census measures of local economic development and our primary data collected in 2025. We find that electrified RGCs with mills have grown significantly in terms of their number of households and number of businesses compared to the not-yet treated RGCs. While electrified RGCs without mills likewise see an increase in the number of households and businesses (in line with Fried and Lagakos's (2021) findings in rural Ethiopia), RGCs with mills grow significantly more, both statistically and economically: they receive approximately 389 additional households and approximately 23 additional businesses. Exploiting four rounds of satellite-based imagery on built-up area until 2020, we can also confirm that, at least over this shorter time horizon, only RGCs with mills add significantly to their building footprint. Finally, the school census data we leverage to track grid arrival also

[1]This pattern holds for both mean and maximum nighttime luminosity. Whereas the former is geared towards detecting changes in (average) nightlights in places that already are somewhat lit, the latter is geared towards picking up genuine extensive margin increases in formerly dark places.

[2]In fact, the adoption for electric lighting appears weakly negative, which may indicate self-electrification by some households in the control group of not-yet electrified RGCs without mills.

allows us to construct an additional RGC-level outcome at annual frequency for all RGCs: the number of schools per RGC. Reassuringly, we fail to detect any pre-trends in either group of RGCs with or without mills. However, the former on average add 0.62 new schools to their village, whereas the latter add none.

Moving to household-derived outcomes, we use the long-difference between the 2010 census and several rounds of household surveys around 2020 to examine the effects of electrification on occupational change, education and housing amenities. In line with the village-wide effects using our own survey as endline, we find significant long-run effects on structural transformation out of agriculture and into services in RGCs with mills, but not in those without. In line with Dinkelman's (2011) seminal contribution, these effects are highly gendered, with women leading the occupational change out of agriculture and into services. These findings neatly provide supply-side effects that directly mirror our demand-side results from the number of businesses discussed above. Educational attainment also increases significantly in RGCs with mills, again driven by women, who significantly improve literacy (+14pp) and years of schooling (+0.7). However, literacy for men also improves significantly by 10pp.[3] Finally, housing amenities improve in RGCs with mills, with more households using taps or boreholes and fewer using wells as their source of drinking water.

Our results are robust to various choices of sample criteria, the construction of the treatment variable, different time horizons and alternative outcome variables.

One open question centres around the role of mills in the villages in our sample. As we highlight in summary statistics, villages with mills are different from those without, which is expected and does not *per se* undermine the comparison of separate staggered difference-in-differences regressions of electrification across RGCs with mills and RGCs without mills. In particular, villages with mills are weakly larger, denser and have more public facilities and businesses. In order to flexibly account for these alternative differences, we run a 'horse race' regression interacting all relevant sources of baseline heterogeneity with electrification. The effect of mills survives and barely changes in magnitude. Therefore, although size, density and the number of public facilities clearly matter, the mill deserves further scrutiny.

To this end, we provide suggestive evidence that mills amplify the effects of vil-

[3]Since we document both an increase in the number of schools, a proxy for the supply of skilled labour, and an increase in non-agricultural employment, a proxy for its demand, we cannot distinguish if households increase educational attainment in response to demand- or supply-side effects.

lage electrification through both cost and demand externalities. First, we show that household adoption is higher in villages with longer predicted low-voltage networks, derived from the spatial arrangement of public buildings the utility has to connect by government mandate. If a larger proportion of the village lies in close proximity to the low-voltage lines connecting, say, the school, health facility and post office, more households adopt. We interpret this findings as suggestive evidence of a cost externality, where the presence of the mill as anchor customer convinces the utility to provide a longer low-voltage network. Accordingly, mills and public buildings appear to be substitutes in terms of enabling electricity adoption for other end-users.

However, whereas RGCs with mills appear to create occupational change, no such economic effects can be detected in RGCs with public facilities, but no mill. We interpret this result as suggestive of the mill also providing something beyond a pure cost externality. Pushing this line of argument even further, we also provide indicative evidence that, indeed, villages with mills see substantially more nighttime luminosity adoption in parts of the village that we predict to fall outside of the low-voltage network that *should* be provided if only the school, public facilities and mill were connected. No such nightlights growth in the 'uncovered' parts of villages can be detected in locations with an equivalently long low-voltage network, but no mill. Finally, our key informant survey provides ample evidence of why mills may matter: millers report of substantial investment that they engage in in the aftermath of electrification given the windfall profits derived for their original business.

This paper contributes to the literature on electrification in three ways. First, we highlight the importance of firms as anchor customers that make electrification financially viable for utilities. While the existing literature has largely studied firm (Rud, 2012; Kassem, 2024) and household electrification (Lenz et al., 2017; Lee et al., 2020b; Masselus et al., 2025) in isolation from each other, this paper sheds light on the interplay of the two through cost and demand externalities. Second, we show that moderate average effects of rural electrification mask stark differences in the impact of electrification across locations, affirming that complementary factors are key to the success of electrification, in line with Moneke (2023), Fetter and Usmani (2024), and Ren (2024), and discussed prominently in Lee et al.'s (2020a) evidence review. Third, we offer a new perspective on the lack of household connections in the vicinity of the medium-voltage grid. We provide suggestive evidence that the absence of the local low-voltage network is an important barrier in many villages in Sub-Saharan Africa.

The remainder of this paper is organized as follows. Section 2 examines electricity adoption patterns in Sub-Saharan Africa, revealing a bimodal distribution of household connections across grid-connected villages in nine countries. Section 3 provides background information on rural electrification in Zambia and the role of grain mills as anchor customers. In section 4, we discuss our empirical strategy and the data we use. Results are presented in Section 5. Finally, Section 6 provides evidence the underlying mechanisms before Section 7 concludes.

2 The last mile of electrification in Sub-Saharan Africa

Sub-Saharan Africa remains the least-electrified region in the world, despite rapid electrification over the last two decades nearly doubling electricity access in the region from 25% in 2001 to 49% in 2021 (IEA, 2023b). Of the approximately 600 million unelectrified people in the region today, nearly 500 million live in rural areas (IEA, 2023a). Therefore, the renewed call for universal access to modern energy, as per Sustainable Development Goal 7, can be interpreted as a call for rural electrification in Sub-Saharan Africa. In line with this, Mission 300, an ongoing initiative by the World Bank and the African Development Bank, aims to connect 300 million people in the region to electricity by 2030 (World Bank, 2026a). The World Bank describes the program's relevance by highlighting that, in their view, "electricity is the bedrock of jobs, opportunity, and economic growth"(World Bank, 2026b).

A puzzling aspect of electrification in rural Sub-Saharan Africa is that around 75 million people live in locations with grid access, but their households do not have an electricity connection (Afrobarometer, 2024). Research has sought to explain this by the low willingness (Lee et al., 2020b; Masselus et al., 2025) and the low ability to pay for electricity (Ruhinduka et al., 2025) of marginal households in grid-connected locations. We propose an alternative explanation: households in grid-connected villages remain unconnected because the local trunk network is missing. While a medium- or high-voltage line reaches the village, there is no local network of low-voltage lines that villagers can connect to – a phenomenon sometimes referred to as "ghost electrification." As a result, connections are prohibitively expensive for most end-users because they involve the construction of a private line with dedicated poles and cables from the closest transformer rather than a simple drop-down from a nearby low-voltage line.

Prima facie evidence from nine countries in Sub-Saharan Africa is consistent with ghost electrification. Leveraging survey microdata from nearly 10,000 households, we document a bimodal distribution of the share of connected households across grid-connected villages. As Figure 1 shows, we observe two modes in each country, one close to a connection rate of zero and another at a much higher connection rate, frequently close to 100%. This suggests that local trunk networks exist in some villages, where they lead to widespread electricity adoption. But in other villages, their absence precludes almost the entire village from adoption.[4]

Why would a local trunk network be constructed in some villages, but not in others? The construction of low-voltage networks represents a high fixed cost for utilities. They need to pay for the installation of poles and cables as well as network maintenance independent of electricity use by villagers. This creates a strong incentive for utilities to concentrate their network construction efforts on villages with high demand for electricity. They can only recover the initial investment and operate profitably if sufficient end-users connect to the grid and consume electricity. Since villages in rural Sub-Saharan Africa are small in population and household demand is limited, potential heavy users, so-called anchor customers, are the key factor differentiating demand between villages. It is their presence that makes it attractive for utilities to construct a local low-voltage network.

In the remainder of the paper, we provide detailed empirical evidence of the key role of anchor customers for electricity adoption in villages and subsequent economic development, focusing on the context of Zambia.

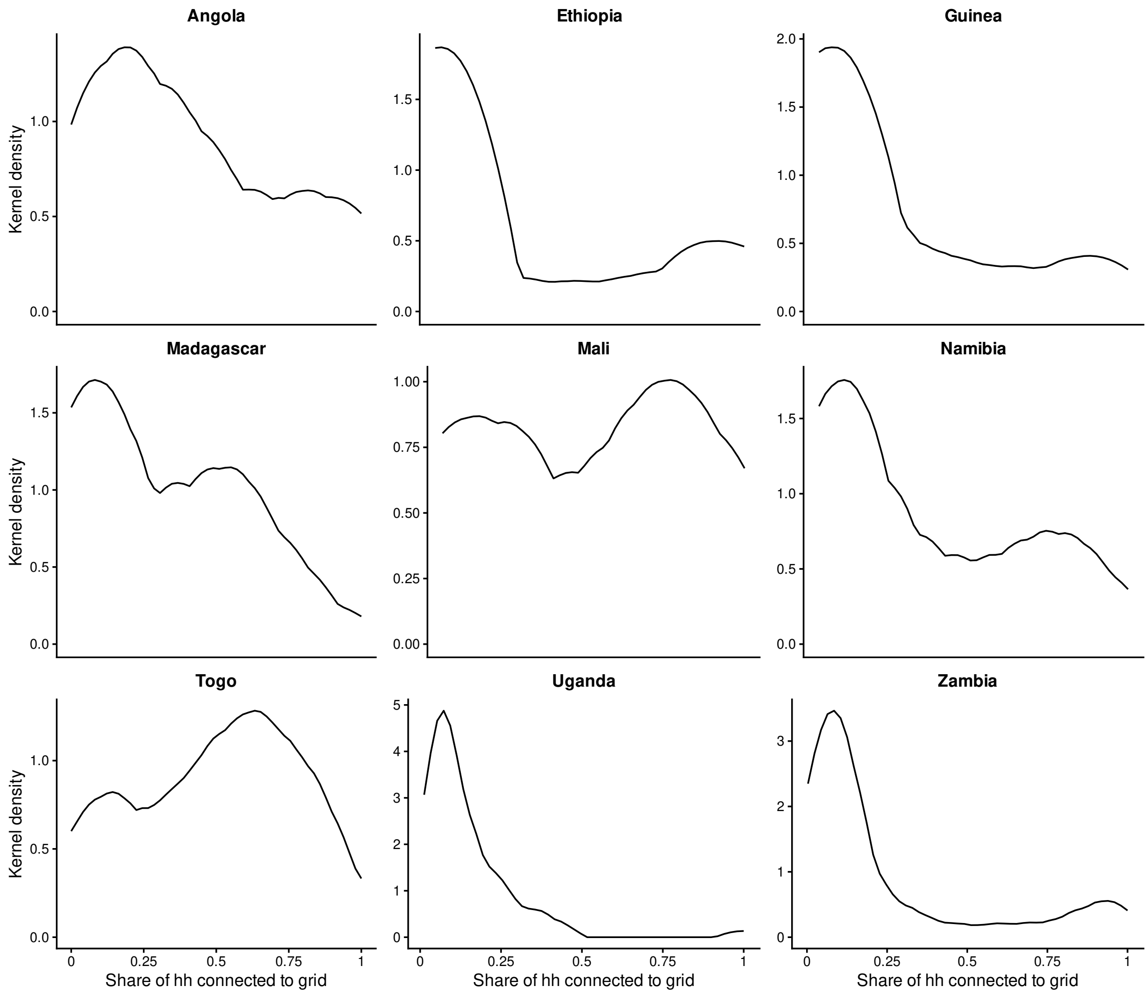
3 Rural electrification in Zambia

3.1 Rural grid expansion

Rural electrification in Zambia started in the early 2000s. In 2003, the Rural Electrification Authority (REA) was established. Subsequently, the first Rural Electrification Master Plan (REMP) was developed to guide grid expansion into rural areas. It was officially launched in 2009 and the first projects were completed in 2012. Since then the share of villages connected to the grid, as proxied by the number of schools connected to the grid, has increased from 10.5% to 23.7%, as illustrated in Figure 2.

[4]Appendix A.I provides details on the underlying data and demonstrates the robustness of the bimodal pattern to the use of alternative data sources.

Figure 1: Electricity adoption in grid-connected villages across Sub-Saharan Africa

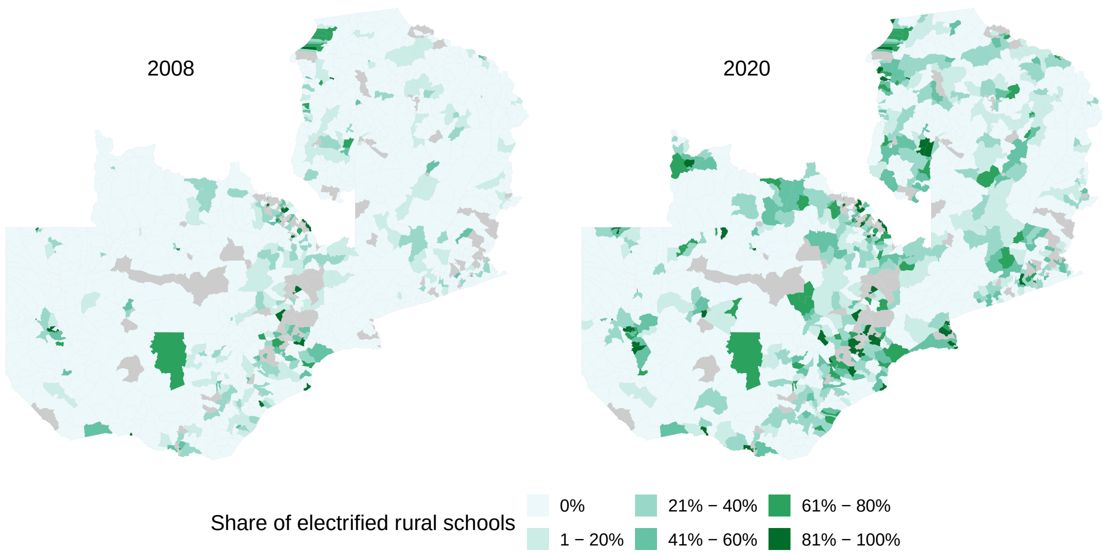


This figure shows the distribution of the share of connected households across grid-connected villages in nine Sub-Saharan African countries. Details on the underlying data are provided in Appendix A.I.

The REMP is a sophisticated expansion plan generated with the assistance of the Japanese International Cooperation Agency using state-of-the-art network planning tools and algorithms.[5] Grid expansion to targeted villages is implemented in two steps. In the first step, the REA builds out the medium-voltage network to a targeted village, installs a transformer near a local public facility, typically the primary school, and connects this facility to the transformer. Subsequently, the installed infrastruc-

[5]The REMP is similar to universal access programs implemented in other low- and middle-income countries, such as the Ethiopian Universal Electricity Access Program (2005), the Kenyan Rural Electrification Master Plan (2009) and the Rwandan Universal Electricity Access Program (2017). Further details on the REMP are provided in Section 4.1 and Appendix A.II.3.

Figure 2: Rural electrification in Zambia



This figure illustrates the growth in rural electrification across Zambian wards. The left- and right-hand map show the share of electrified schools in 2008 and 2020, respectively. Darker shades of green indicate higher electrification rates. Grey shaded areas indicate urban areas and national parks.

ture is handed over to the national utility ZESCO to connect private end-users.[6]

In the second step, the utility receives applications for individual end-user connections from households and firms. Any individual or group application involves standard engineering ground surveys and costing exercises to determine the location-specific fixed cost of erecting a local low-voltage network, including poles, overhead lines, line drops to individual end-users, installation of meters and internal wiring. If the resulting connection fee offered by the utility is accepted by the applicant, the utility will construct the necessary infrastructure to connect the end-user.

Two cases are frequently observed in rural Zambia: either no local low-voltage network is constructed by the utility and the school remains the only user of electricity in the village, or a local low-voltage network does get constructed, involving low-voltage lines, poles, drop-downs to individual end-users such as households and businesses for at least parts of the village. We provide a sketch of both cases in Figure A3. Our guiding hypothesis is that the presence of anchor customers in some villages but not others separates the two cases.

[6]Figure A2 showcases the arrival of a medium-voltage line at a school in rural Zambia alongside the transformer and the low-voltage network spanning from there (Panel a). It also displays drop-downs from the low-voltage network to individual households (Panel b).

3.2 The role of grain mills

In Zambia, grain mills are used to crush maize into coarse flour, locally often referred to as mealie meal. This flour is the main ingredient in Zambia's staple food, nsima.

Grain mills can be powered by diesel or by electricity, but operating an electric mill is more profitable.[7] In a survey of 160 millers, 98% of grid-connected millers stated that their business earned higher profits since they had switched to an electric mill. The reasons behind this are manifold. First and foremost, electricity is cheaper than diesel, hence operating costs are lower. Additionally, maintenance costs are lower and downtime is reduced. Finally, millers indicated that electric mills allowed them to produce higher quality output and attract more customers because they could produce "mealie meal free of diesel scent."[8]

Connecting mills to the electricity grid is not only attractive from the millers' point of view. It is also attractive from the utility's point of view. In fact, the engineers developing the REMP had identified millers as key customers, not only because they would be eager to connect, but also because they would be "one of the major electricity users, probably the largest consumers" (Rural Electrification Authority, 2009). Surveys of recently electrified villages at the time demonstrated the high electricity usage of grain mills compared to all other end-users in rural areas: in all surveyed villages, grain mills accounted for 50% to 90% of total electricity usage. Figure 3 illustrates this pattern using load curves for the villages of Kalobwa, Kambwali and Luamfumu village. Even though grain mills (light blue) do not consume any electricity at night because they are not operating, they clearly account for the bulk of total electricity usage (red) in all three villages.

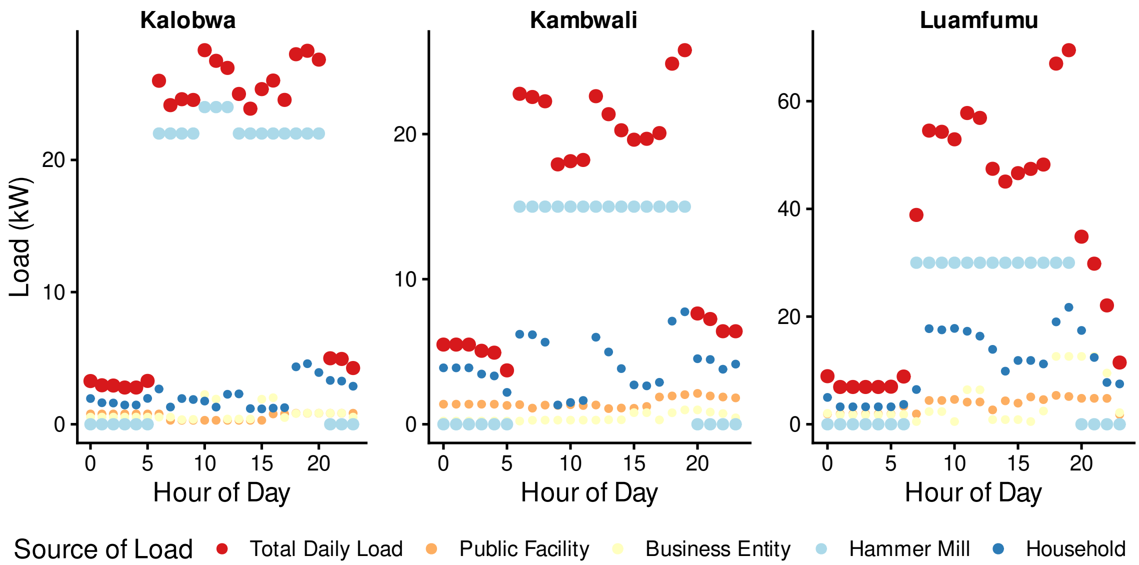
In short, as heavy electricity users, mills are important anchor customers for the utility, and they themselves have a vested interest in connecting to the grid. In line with this, we observe that most grain mills in grid-connected villages are electric. In our own survey, 76% of the mills in electrified villages report operating an electric milling machine. Moreover, all unconnected millers say they would like a connection, but cite the long distance to the low-voltage network as the key barrier.

Given their role as dominant customers for the utility, mill connections have im-

[7] Solar-powered mills have become available recently, but did not feature during our study period.

[8] Figure A4 shows an electric mill and a diesel mill in the same village. A full breakdown of the advantages of electrification listed by millers is provided in Figure A5. Details on the miller survey are provided in Section 4.2.2 and Appendix A.III.2.3.

Figure 3: Electricity consumption by end-user in grid-connected villages



This figure shows daily load curves by end-user type in three grid-connected Zambian villages. Data derived from Rural Electrification Authority (2009), Table 5-6.

plications for electricity adoption in the village more broadly. In villages without anchor customers the utility cannot reasonably expect to recover the high fixed cost of constructing the local low-voltage network over the lifetime of the investment. Households, public facilities and other businesses simply consume too little electricity to fully utilise the low-voltage network. Even worse, poorer villages often have lower population density than richer, more urbanised (and thus: compact) villages, meaning fixed cost for the utility and electricity demand can be inversely linked. In contrast, anchor customers such as mills guarantee steady electricity demand and tariff income for the utility that helps amortise their low-voltage network investment. However, once the distance from the transformer at the school to the mill is covered by low-voltage lines and poles, connecting additional, nearby end-users such as households and businesses also plummets. It is exactly in these villages that widespread adoption and potential development effects are possible. In the following, we test this hypothesis empirically exploiting the Zambian REMP over the last two decades.

4 Research design

4.1 Empirical strategy

The REMP determined 1,200 so-called rural growth centres (RGCs) as primary targets for rural electrification. At baseline, these were unelectrified centres of rural economic activity with a high concentration of residential settlements, as identified through a consultative process with district officials from the entire country.

RGCs were prioritised for electrification based on forecasted demand. While the engineers at the time had recognized the importance of mills as anchor customers, a national database of these was not available, so they could not be explicitly taken into account. Instead, demand was simply modelled as a linear function of baseline population estimates in 2006.

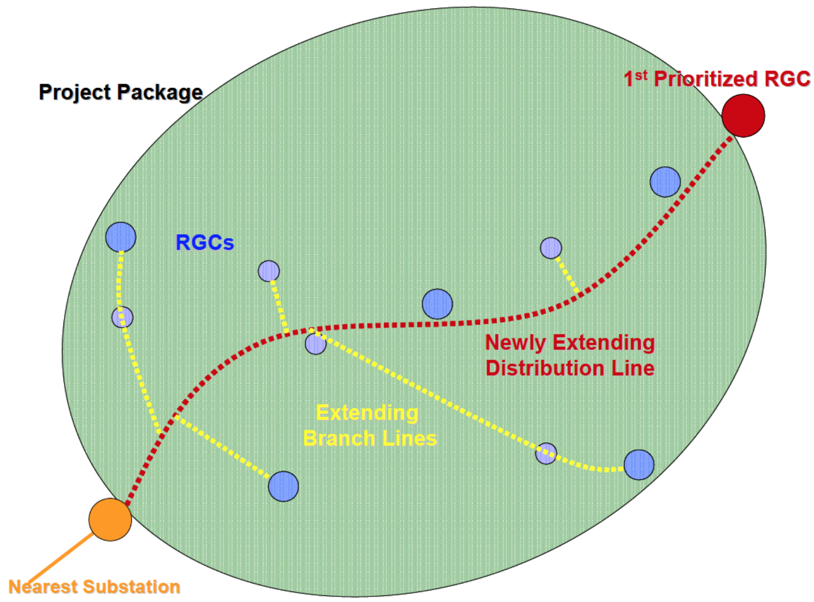
Following prioritisation, RGCs were grouped into 180 project packages. Each package was formed by a high-priority RGC as the primary target and the surrounding RGCs falling between said target and the nearest existing substation. Low-ranked RGCs that did not fall on the path between any high-ranked RGC and a substation were earmarked for off-grid electrification. A schematic drawing of the packaging is shown in Figure 4: the sketch highlights the existing substation (orange circle), the target RGC (red circle) as well as the resulting inconsequential RGCs (blue circles) that are of lower rank but accidentally electrified alongside the target RGC.[9]

Our empirical strategy exploits the fact that some inconsequential RGCs were electrified earlier than others, independent of their individual priority rank, because they were located on the path between a substation and a higher ranked target RGC. In fact, inconsequential RGCs that were connected to the grid during our study period do not systematically differ in their expected electricity demand or rank from the ones that remain unconnected at the end of our study period – neither in RGCs with a mill at baseline nor in RGCs without a mill, as documented in Table A2. Therefore, our empirical strategy compares relatively fortunate inconsequential RGCs that were electrified early to relatively unfortunate inconsequential RGCs that were electrified late or not yet at all due to budget limitations.

Given the dominant role the REMP awarded to grain mills as anchor customers, we compare changes in outcomes associated with village electrification between in-

[9] Additional details on the identification and prioritisation of RGCs as well as their grouping into packages are provided in Appendix A.II.3. We also provide maps of real-world project packages.

Figure 4: Inconsequential RGCs in electrification packages



This figure displays a schematic sketch of a project package composed of a target RGC (red) and several inconsequential RGCs (blue). Source: Rural Electrification Authority (2009).

consequential RGCs with and without a grain mill at baseline. To this end, we adopt the following regression specification:

$$Y_{itp} = \beta Grid_{itp} + \delta Grid_{itp} * Mill_{it_0p} + \gamma_i + \nu_{ct} + \epsilon_{itp} \quad (1)$$

where i denotes an inconsequential RGC, t denotes year, c denotes the constituency that contains RGC i and t_0 denotes the baseline year of 2010 in which we measure presence of mills in location i. γ_i denotes RGC fixed effects, while ν_{ct} denotes constituency-year fixed effects that control for time-varying effects of the relevant parliamentarian's efforts in affecting the release of funds for electrification in their constituency. All standard errors are clustered at the electrification package-level denoted by p.

The coefficient β captures the effect of grid connection in villages without a mill at baseline, and δ the differential effect in villages with a mill at baseline. The identification of the causal effect of electrification in each type of village requires the assumption of parallel trends within each subset.

We report results in two different ways, dictated by the frequency of the outcome data. For annually observed outcomes, we present state-of-the-art dynamic staggered difference-in-difference event studies, separately for villages with and without mills. For outcomes that are only observed at the beginning and at the end of our study period, we report standard two-way fixed-effects regressions in line with equation (1).

4.2 Data

4.2.1 Tracking the expansion of the electric grid

In the absence of administrative data on the medium-voltage network at baseline as well as its subsequent roll-out over our study period, we exploit a prominent feature of rural electrification in Zambia to track the expansion of the grid: the legal mandate for the REA to connect all forms of public infrastructure in villages selected for grid electrification. In practice, this mandate mostly pertains to the electrification of public schools as the most common and widespread type of public infrastructure. As a result, the medium-voltage line is typically routed to the local school where a transformer is installed to connect the school as the first end-user. Thus, tracking school electrification provides an excellent way to track grid expansion.

We leverage the annual School Census from the Ministry of General Education to build a novel village-year panel of electrification. The School Census requires the headmasters of all Zambian schools to fill a paper form with basic information about their school, including its access to electricity. The answer options differentiate between grid access, off-grid access, and no access. Combining annual school census data from 2005 to 2020, we first construct a school-year panel of electrification covering a total of 14,381 schools. Then we geo-identify schools and spatially link them to RGCs through an intricate process described in detail in Appendix A.III.1.1. Ultimately, we successfully match 4,186 public schools that existed in 2008 to 742 out of 784 inconsequential RGCs.[10]

We cross-validate our school-based village electrification measure using partial information from four alternative sources: a 2023 map of the medium-voltage network from the utility, our own 2025 village survey, REA projects records and a series of health facility censuses. Reassuringly, our electrification measure is highly correlated with all of these alternative measures, both in the cross-section and within RGCs over

[10] Unmatched RGCs remain because we cannot geo-identify all schools in the school census due to missing information. Further details on the data extraction, cleaning and geo-identification are provided in Appendix A.III.1.1.

time, as detailed in Appendix A.III.1.2.

The described, school-based approach to recovering the history of grid expansion at the micro-level is key for our study. Beyond our context, it also offers a blueprint for other low- and middle-income countries plagued by a lack of detailed grid expansion records, especially countries where the electrification of public facilities is governed by a similar mandate as the Zambian one.

4.2.2 Outcome data

We use three types of outcome data. First, we leverage remotely-sensed data on nighttime luminosity between 2012 and 2020 (Elvidge et al., 2021), built-up area in 2005, 2010, 2015 and 2020 (European Commission, Joint Research Centre, 2023) and vegetation between 2008 and 2020 (VITO, 2026) at fine spatial resolution over time. To generate RGC-level panel data from this, we overlay the respective remotely-sensed outcome variable raster with two-kilometre buffer polygons around RGC centroids.

Second, we combine geo-identified microdata from the quarterly national labour force surveys 2018-2022 with the geo-identified full count data from the 2010 Population and Housing census (PHC) to measure long-difference changes in household electricity adoption, housing conditions, education and sectoral employment from the beginning of our study period in 2010 to its end around 2020. This allows us to assess changes in these outcome measures against village electrification between 2010 and 2020.[11]

Third, we conducted a key informant survey across 661 out of 784 inconsequential RGCs in 2025 in order to gather both quantitative and qualitative evidence on the village electrification process. In the quantitative analysis, we primarily use the survey data to estimate the long-run effects of electrification on the number of households and businesses. The rich qualitative evidence is used alongside descriptive statistics in Section 6 to support the mechanisms underlying electricity adoption and development effects in villages with mills. A detailed description of the endline survey is provided in Appendix A.III.2.3.

[11] Appendix A.III.2.2 provides further details on the labour force survey.

4.2.3 Baseline heterogeneity

We use geo-identified full count data from the 2010 Population and Housing census (PHC) to examine the heterogeneous effects of electrification across villages by baseline characteristics. In particular, we leverage three separate datasets that were produced as part of the census enumeration as described below.[12]

First, we use a geo-identified cross-section of 58,500 points of interest collected as a by-product of the PHC. This dataset contains coordinates and a classification of any given 'point of interest', including facilities as diverse as grain mills, churches, commercial enterprises, public facilities, wells, bus stops, grain depots and sewage systems. To the best of our knowledge, this dataset is the most accurate and holistic source of ground-truthed facility location data across Zambia at baseline. For our analysis, we group the universe of points of interest into five categories – grain mills, commercial facilities excluding mills, public facilities, religious facilities, large-scale farm operations and large-scale grain depots – and spatially link them to RGCs using two-kilometre buffer zones around RGC centroids. This allows us to study heterogeneity in electrification effects by the baseline presence of grain mills and other facilities. Additionally, it enables us to estimate long-run effects on business and household counts as highlighted above.

Second, we exploit household and individual-level data from the population census to examine baseline differences between RGCs with and without mills, and to estimate the long-run effects of electrification in conjunction with the LFS 2018-2022 endline data described above.

Third, we leverage the universe of buildings recorded in the housing census to provide additional detail and robustness on baseline differences between RGCs with and without mills. In particular, we focus on non-residential buildings that are part of commercial operations as an alternative measure of the presence of potential anchor customers at baseline.

4.3 Descriptive statistics

The number of grid-connected RGCs increased steeply over our study period. While 122 out of 742 RGCs already had a grid-connected school within 2km at baseline,

[12]While the REMP was launched in 2009, the first projects were only completed in 2012. Therefore, the 2010 census data serves as our baseline. Additional details on all three datasets are provided in Appendix A.III.3.

only 10 had all schools in their vicinity connected to the grid at the time. Until 2020, the number of at least partially and fully connected RGCs rose to 349 and 125, respectively, as shown in Figure A15. Among RGCs with a mill at baseline, the number of RGCs with any connected school increased by 115%. In RGCs without a mill, it increased by 237%. Similarly, growth in fully electrified RGCs was somwehat more pronounced among villages without a mill at baseline.[13]

RGCs with and without a mill also differ along other dimensions at baseline. As documented in Table A4, RGCs with mills are larger and more densely populated. As a result, their expected electricity demand is higher and they are given higher priority for electrification in terms of their rank. They also have more public facilities, more commercial and religious activity, and are closer to the road network. However, they are not located in wards that have experienced differential population growth since 2000.[14] Nor do they differ from RGCs without a mill in terms of household electricity usage, sectoral employment shares or education levels.[15] Our comparison of the impact of grid connection in RGCs with and without a mill will capture the joint effect of all the documented differences between the two types villages. However, we provide ample evidence in Sections 5.3.1 and 6 supporting the central role of mills as anchor customers that shape the effect of rural electrification in this context.

Interestingly, prima-facie evidence is consistent with our conjecture that the presence of a mill is an important determinant of electrification success. Breaking down the bimodal distribution of household electrification rates across grid-connected RGCs in 2015 – halfway through our study period – into RGCs with and RGCs without mills at baseline, we find all the mass in the high electrification mode to be accounted for by RGCs with a mill, as shown in Figure A16.

5 Results

5.1 Electricity adoption

We estimate a small positive average effect of village electrification on nighttime luminosity. As shown in Table 1 (columns 1 and 3), electrification is associated with

[13]We illustrate the spatial pattern of the increase in electrification in Figure A10.

[14]The ward is the most granular level at which we can measure local population growth.

[15]We observe higher mean nighttime luminosity in RGCs with a mill in 2012, the first year of the VIIRS nightlights data. However, it is unclear to what extent this captures a baseline difference or a treatment effect because the first electrification projects were already completed by 2012.

an increase in mean (maximum) luminosity by 0.03 (0.21) nanoWatts per square centimeter per steradian.[16] These estimates are in line with recent literature on rural electrification which reports small to moderate effects on adoption (Lenz et al., 2017; Lee et al., 2020b; Burlig & Preonas, 2024; Masselus et al., 2025; Kassem et al., 2026).

Table 1: Nightlights on RGC connection by mill status

	Mean nightlights		Max. nightlights	
	(1)	(2)	(3)	(4)
RGC connected	0.03**	0.01	0.21**	0.05
	(0.02)	(0.01)	(0.10)	(0.10)
RGC connected * mill		0.08**		0.64**
		(0.04)		(0.27)
RGC FE	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓
R^2	0.96	0.96	0.88	0.88
No. of years	9	9	9	9
No. of RGCs	742	742	742	742
No. of const-years	1215	1215	1215	1215
Observations	6678	6678	6678	6678

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC connected denotes percent of schools in RGC reporting grid electricity connection in a given year, derived from balanced sample of schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Mean and max. nightlights represent the mean and maximum, respectively, zonal statistics of annual nighttime luminosity values derived from the VIIRS VNL sensor for each RGC 2km buffer polygon. Const-year FE denotes constituency-year fixed effects.

In line with our hypothesis, these moderate average effects are entirely driven by villages with grain mills at baseline. In such villages, the effect of electrification is approximately three times as large as the average effect, amounting to an increase in mean (maximum) luminosity by 0.09 (0.69) units. In contrast, in villages without a mill, we estimate negligible and statistically insignificant effects (columns 2 and 4).

In Figure 5, we document the temporal and spatial signature of the estimated effects on mean luminosity. As shown in Panel (a), we cannot detect any pre-trends

[16] This is the unit in which spectral radiance is measured. It captures the amount of light energy emitted from a specific area on the ground and captured by the satellite's sensor.

in villages with or without mills. Following electrification, luminosity increases in villages with a mill (blue) within a year and subsequently remains at an elevated level. In villages without a mill (grey), changes in luminosity are much more muted.

The spatial signature in Panel (b) reveals that effects of electrification are concentrated in the direct proximity of RGCs with mills, with significantly higher mean nightlights responses than in no-mill RGCs detectable up to 3km distance from the RGC centre. Although the effects naturally decay with distance to the RGC, it is noticeable that electrified RGCs with mills show a much more prevalent spatial signature than their electrified no-mill counterparts. It is noteworthy that even in villages without a mill, we find a positive effect in the very centre, albeit at much smaller magnitude than for villages with a mill. This is consistent with the fact that at a minimum, the schools in these villages are electrified.[17]

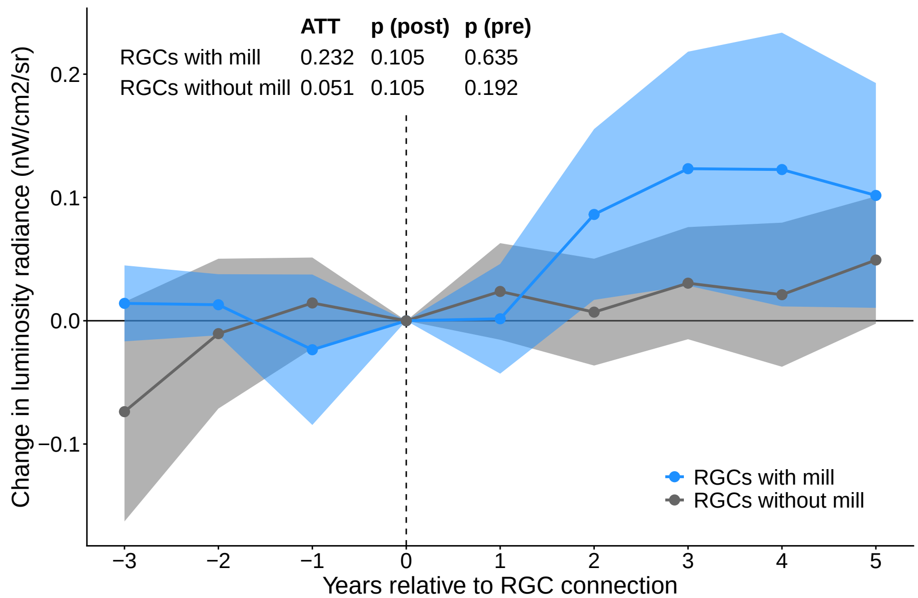
Long difference estimates comparing changes in electricity use for lighting and cooking associated with village electrification between villages with and without mills corroborate our findings. In villages without a mill at baseline, we do not detect any statistically significant effects. In villages with a mill, we document an increase in the use of electric lighting by 23pp and a simultaneous decrease in the use of solar lighting by 16pp, suggesting that households switch from solar to electric lighting. Similarly, we find evidence of an increase in the use of electricity for cooking (6pp) and a corresponding decrease in the use of wood (and charcoal) of the same magnitude. Importantly, estimates are significantly different from those for villages without a mill for all main outcomes.[18]

To conclude our findings on electricity adoption, we also provide cross-sectional results using much richer questions from our key informant survey on the electrical equipment present in villages at endline. As reported in Table A5, electrified RGCs with mills have almost double the number of electric poles (+27.6 poles), at least half a transformer more (+0.6 transformers, with p-value of 0.1005) and more than

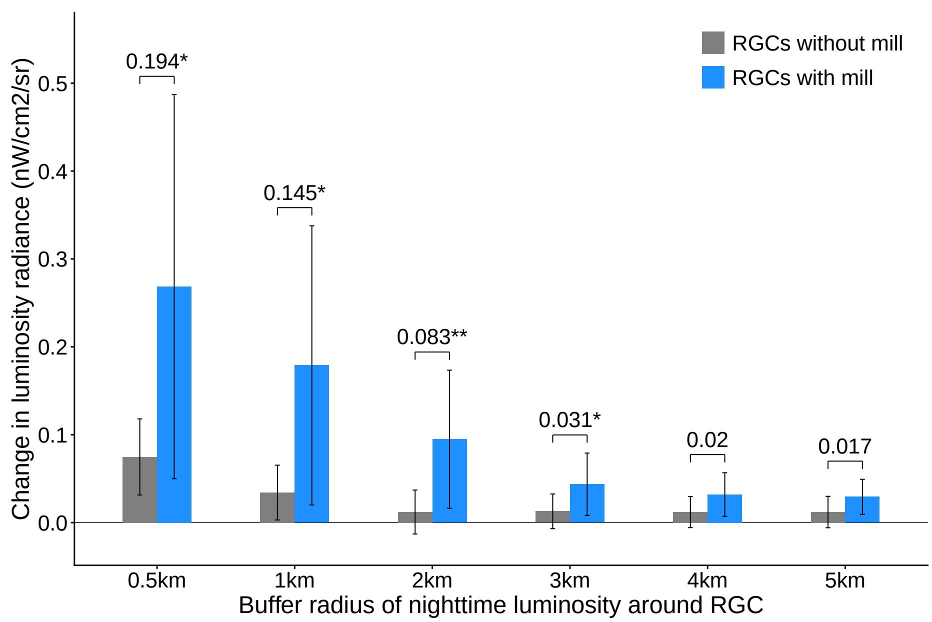
[17]While the schools themselves do not operate at night, it is likely that the teacher houses and outdoor lights on the school premises emit light upon electrification. Additionally, it is conceivable that businesses or households in the direct proximity of the school are also able to connect to the transformer via the school without further investment into low-voltage infrastructure by the utility.

[18]Note that given the small number of RGCs that intersect with labour force survey enumeration areas at endline, the binding constraint for the long-difference sample, and the shape of enumeration area polygons, we use any enumeration area that intersects within 5km of the RGC for the long difference analysis. Robustness to the much smaller sample using enumeration areas that intersect within 2km are provided in Table A6, which are qualitatively unchanged.

Figure 5: Effect of electrification on nightlights: temporal and spatial signature



(a) Event study of the effect of village electrification on nightlights



(b) Spatial signature of the effect of village electrification on nightlights

Panel (a) displays an event study of the effect of village electrification on mean nighttime luminosity in RGCs with and without a mill at baseline. Time is displayed in years relative to the year of village electrification. Panel (b) shows how estimates of the effect of village electrification change as the buffer radius around the RGC centroid increases for both RGCs with and without mills at baseline.

double the share of households reporting an electricity connection at endline than electrified RGCs without mills. Although clearly suggestive given the lack of baseline information on electric equipment, these findings paint an overall highly consistent picture: locations with mills experience a dynamic nightlights response that is four to five times the magnitude of no-mill RGCs. This adoption response extends far into the village, reaching up to 3km from the RGC centre. In addition, households report switching from solar to grid-electric lighting at a rate that matches the difference in cross-sectional difference in the share of households connected almost perfectly. Finally, the broader adoption results are supported by suggestive evidence of disproportionately more low-voltage network equipment being present at endline.

Table 2: Adoption on RGC connection by mill status

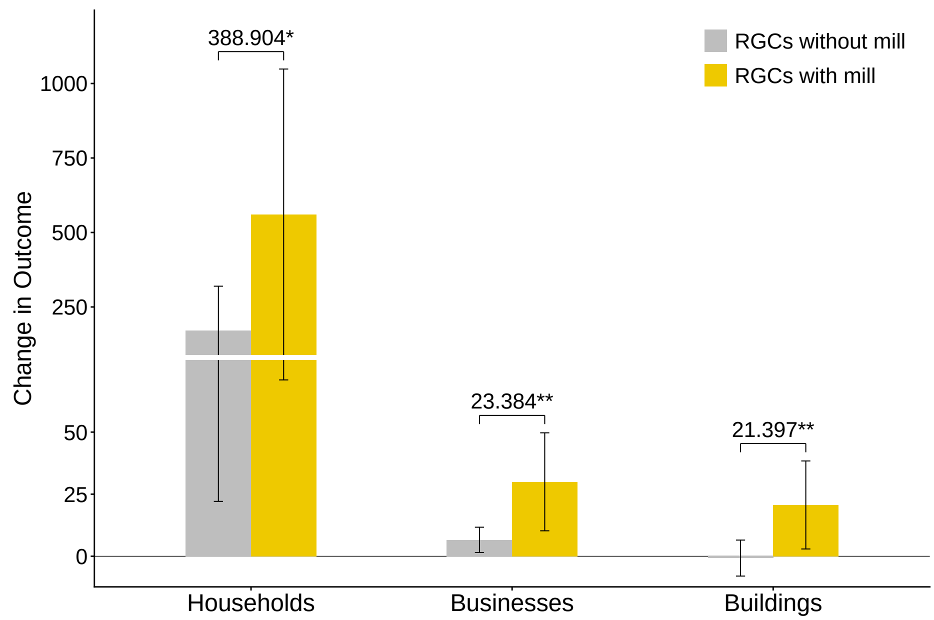
	Lighting			Cooking		
	Electric	Solar	Other	Electric	Wood	Other
RGC connected	-0.08^{**}	0.04	0.04	-0.01	0.00	0.00
	(0.03)	(0.04)	(0.04)	(0.01)	(0.02)	(0.01)
RGC connected * mill	0.23^{***}	-0.16^{***}	-0.09	0.06^{*}	-0.06^{*}	0.00
	(0.09)	(0.05)	(0.09)	(0.03)	(0.03)	(0.01)
RGC FE	✓	✓	✓	✓	✓	✓
Prov-year FE	✓	✓	✓	✓	✓	✓
R^2	0.84	0.79	0.87	0.83	0.82	0.53
No. of years	2	2	2	2	2	2
No. of RGCs	226	226	226	226	226	226
No. of prov-years	20	20	20	20	20	20
Observations	452	452	452	452	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Prov-year FE denotes province-year fixed effects. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) that intersect with any sample enumeration area (SEA) sampled in both PHC 2010 at baseline and any LFS round between 2018-2022 at endline. RGC-level actual electrification derived from schools existing since 2008. Mill in RGC denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Lighting: Electric denotes share of households in sample enumeration area (SEA) that report using grid-electric lighting. Lighting: Solar denotes share that report using non-grid electric lighting. Cooking: Electric and Cook: Wood denote share of households that report cooking using grid-electricity or wood/charcoal, respectively.

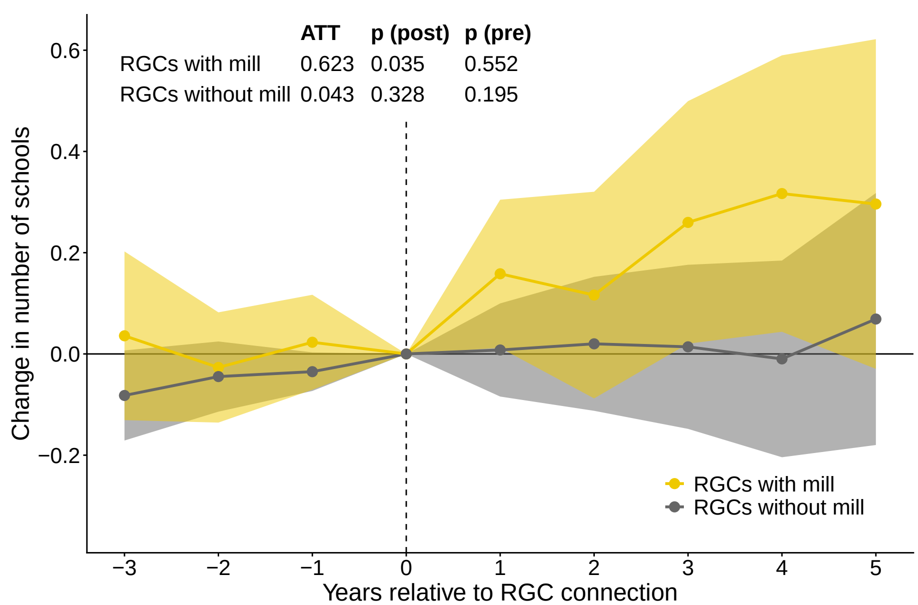
5.2 Local economic development

In light of the stark differences in adoption, we turn to the potential economic effects of grid electrification across villages with and without mills. Do the factors that boost adoption also contribute to potential development effects of electrification? To address this question, we combine data from various sources to assess the effect of village electrification on local economic development in RGCs with and without a mill at baseline. First, we compare the change in the number of households and businesses between baseline (PHC 2010) and endline (key informant survey 2025). As shown in Panel (a) of Figure 6, we detect a significant increase in both businesses and households in both RGCs with and without a mill at baseline. However, the increase is significantly larger in RGCs that had a mill at baseline: the latter report a large increase in population relative to RGCs without a mill, with an additional 389 households added after fifteen years. These stark results are in line with the migration results in Fried and Lagakos (2021), who find a dramatic decrease in out-migration of 26% and 9% more in-migration in recently electrified Ethiopian villages, a high-fertility context comparable to ours. In fact, population growth rates at the ward-level between the 2000 and 2010 population censuses indicate that the parts of rural Zambia home to RGCs in our sample have grown their population by 35% on average (cf., Table A4), equivalent to a 3% growth rate per annum and almost exclusively driven by high fertility. Projecting this rate forward to 2025, RGCs with mills would be expected to have grown to a population of approximately 5,650 people, ceteris paribus, or approximately 1,150 households in the absence of electrification. Our estimates thus indicate a substantial differential increase in agglomeration on top of an already rapid background population growth.

Figure 6: Effect of electrification on local economic development



(a) Effect of electrification on households, businesses and buildings



(b) Event study of the effect of village electrification on number of schools

Panel (a) shows the effect of village electrification on the number of households, businesses (both PHC 2010 vs. village survey 2025) and buildings (GHSL built-up area) in RGCs with and without a mill at baseline. Panel (b) displays an event study of the effect of village electrification on the number of schools in RGCs with and without a mill at baseline. Time is displayed in years relative to the year of village electrification. Note that the y-axis in Panel (a) is broken at a value of 80 and compressed above to aid readability, as indicated by the white gap in the household bars.

Second, in line with these findings on population, we also highlight how electrified RGCs with mills at baseline appear to develop into true 'rural growth centres', hosting 23.3 additional small-scale businesses by 2025 compared to electrified RGCs without mills, which see the entry of 6.6 new businesses after electrification over the fifteen year time span. Anecdotal evidence, including from our own miller survey, highlights that especially small-scale retail and services establishment are responsible for the brunt of this large entry effect.

It is also interesting to note that the spatial signature of business entry may differ across RGCs: qualitative evidence suggests that the increases in business activity in RGCs without mills are limited to the direct vicinity of schools. As a survey respondent in a village where only the school was connected explained, "electrification allowed [him] to request for a place at the school to open a grain mill and a grocery store were [he] sell[s] cold drinks and other groceries." He also mentioned unprompted that "if the community gets connected to the grid it would promote more business such as welding, salons and others."

Third, we estimate the effect on built-up area within a 2km buffer around the RGC centre using remotely sensed information available every five years, spanning the 2005-2020 time period. We find a large increase in RGCs with a mill at baseline, but no effect in RGCs without a mill (Figure 6, Panel (a)). In RGCs with a mill the increase amounts to an additional average 21.4 square metres of built-up area per 10,000 square metre pixel. In other words, the average electrified RGC with mill at baseline experiences one new small-scale building (or hut) of approximately 20 square metres for every 100m*100m pixel within its vicinity – which directly confirms the large effects on the number of households and number of businesses reported above.[19]

Fourth, the annual school census provides us with an additional source of high-frequency outcome data available at the village-year level, which yields two advantages for our analysis. It enables us to test for the absence of pre-trends for both RGCs with and without mills before grid connection using a second high-frequency outcome, but one that is obtained by headmasters on the ground instead of remotely-sensed. It also provides us with another high-frequency outcome to confirm the large-scale development effects documented above in RCGs with mills: the number of schools in RGCs. Panel (b) of Figure 6 provides event study estimates of the effect of grid

[19] To confirm robustness of our results to the 2km buffer choice, we provide spatial signature plots varying this buffer for remotely-sensed built-up area and vegetation outcomes in Appendix A.V.2.

connection on the number of schools in the RGC. A clear, significant upward trajectory is visible in RGCs with mills, which experience the entry of 0.62 new schools, significant at the 5% level. This effect starts to emerge within one year of a headmaster of an existing school reporting electrification in the RGC.[20] No increase in the number of schools is detectable in RGCs without mills at baseline. Reassuringly, neither RGCs with nor those without mills exhibit any pre-trends in the number of schools, which confirms our identifying assumption that parallel trends between early and late-electrified inconsequential RGCs within each group appear to hold.[21]

Fifth, exploiting our long difference in household-level data from the 2010 PHC and the 2018-2022 Labour Force Surveys, we find evidence of structural information out of agriculture into services. As shown in Panel A of Table 3, in RGCs with a mill, grid connection is associated with a decrease in the agricultural employment by 20pp, an increase in services sector employment by 14pp and a small, statistically insignificant increase in manufacturing employment. Effects are highly gendered, with structural transformation effects being almost exclusively driven by women moving out of agriculture into services, in line with the existing literature (Dinkelman, 2011; Dinkelman & Ngai, 2022). This finding also echoes our earlier result on the large increase in small-scale businesses in connected RGCs with mills. In contrast, we do not find any evidence of structural transformation in RGCs without a mill at baseline.

Sixth, we document an increase in educational attainment in response to grid connection, with the effect restricted to RGCs with a mill once again. As Panel B of Table 3 shows, grid connection is associated with an increase in literacy (12pp) and years of schooling (0.71) in RGCs with a mill. We do not observe a statistically significant increase in enrolment, however. Again, effects are mostly driven by women, although male literacy also improves significantly (column 8).[22]

Exploiting the rich microdata underlying our analysis, we provide further corroborating evidence on educational attainment by sex and age group. Reassuringly, as shown in Figure A19, Panel (a), the differential effect of grid connection on literacy in RGCs with mill at baseline is most pronounced for youths under the age of 15, while

[20] Arguably fast, school entry in rural Zambia often occurs bottom-up, with schooling reacting quickly to grass-roots demand, with official incorporation by the Ministry allowed to follow later.

[21] Throughout our main analysis we employ a balanced panel of schools that existed continuously from 2008-2020 to derive RGC connection status. However, for the above analysis of school entry, we use the full sample of all schools instead to allow for school entry and exit in the panel.

[22] We provide results on occupation and education using the smaller sample from the 2km intersection with enumeration areas in Table A7 for robustness. Results are qualitatively similar.

Table 3: Sectoral employment and education on RGC connection by mill status

	Overall			Female (♀)			Male (♂)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Occupation									
	Agr.	Man.	Ser.	Agr.	Man.	Ser.	Agr.	Man.	Ser.
RGC connected	0.02	0.01	−0.03	0.03	0.01	−0.03	−0.01	0.02	−0.02
	(0.09)	(0.07)	(0.04)	(0.09)	(0.07)	(0.05)	(0.09)	(0.08)	(0.05)
RGC connected * mill	−0.20*	0.06	0.14**	−0.27**	0.07	0.20***	−0.13	0.06	0.07
	(0.11)	(0.09)	(0.06)	(0.12)	(0.09)	(0.07)	(0.11)	(0.09)	(0.07)
R^2	0.70	0.60	0.79	0.68	0.59	0.81	0.71	0.62	0.74
Observations	452	452	452	451	451	451	452	452	452
Panel B: Education									
	In edu.	Lit.	Years	In edu.	Lit.	Years	In edu.	Lit.	Years
RGC connected	0.01	0.01	−0.19	−0.01	0.00	−0.09	0.02	0.02	−0.28
	(0.02)	(0.02)	(0.16)	(0.02)	(0.02)	(0.15)	(0.02)	(0.02)	(0.20)
RGC connected * mill	0.04	0.12***	0.71**	0.06**	0.14***	0.73**	0.03	0.10***	0.64
	(0.03)	(0.03)	(0.32)	(0.03)	(0.04)	(0.32)	(0.03)	(0.03)	(0.40)
R^2	0.79	0.90	0.88	0.80	0.89	0.89	0.72	0.86	0.83
Observations	452	452	452	452	452	452	452	452	452
RGC FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Prov-year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
No. of years	2	2	2	2	2	2	2	2	2
No. of RGCs	226	226	226	226	226	226	226	226	226
No. of prov-years	20	20	20	20	20	20	20	20	20

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$. All standard errors clustered at electrification package-level. Prov-year FE denotes province-year fixed effects. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) that intersect with any sample enumeration area (SEA) sampled in both PHC 2010 at baseline and any LFS round between 2018-2022 at endline. RGC-level actual electrification derived from schools existing since 2008. Mill in RGC denotes an indicator value that takes the value of one if at least one mill situated within 5km of RGC. Sectoral Employment: Agr. denotes share of individuals in sample enumeration area (SEA) that report agriculture as their primary occupation. Man. and Ser. denote the equivalent occupation shares for manufacturing and services sectors, respectively. Education: In edu. denotes share of individuals in sample enumeration area (SEA) that report being in school. Lit. denotes share that are literate. Years denotes the years of education reported by individuals.

weaker but insignificant for 15-29 year old women, and absent for 15-29 year old men and any adults aged 30-44. Regarding educational attainment in terms of years of education shown in Panel (b), 15-29 year old women display by far the largest effect, adding 1.3 years of schooling in connected RGCs with a mill at baseline. Effects on the young are symmetric by sex, with both girls and boys under 15 adding approximately half a year of education in grid-connected RGCs with mills. These education effects by age and sex resonate closely with the female home-to-market transition documented above.

In summary, we find interesting labour market dynamics of grid connection in RGCs with a mill, which experience an increase in educational supply due to new schools and an increase in the demand for skill from the emergence of businesses. Especially women change occupations from agriculture to services, and especially young women and adolescents increase literacy and educational attainment. Although we cannot disentangle if the education effects are driven by changes in the demand or supply of education, the assembled evidence provides a very coherent picture of structural transformation in grid connected RGCs with mills at baseline.

Finally, moving to household amenities, we detect an impact of grid connection on one measure of household amenity available in both census and labour force surveys: the source of water supply. This effect is limited to RGCs with a mill at baseline, however. As a respondent to our survey explained, "electricity also helped address water challenges as it allowed villagers to drill boreholes and install electric water pumps." Consistent with this qualitative evidence, we see a reduction in the use of water obtained from wells (-14pp) and an increase in the use of water obtained from boreholes or through a tap (+18pp), as highlighted in Table A8.[23]

It is important to note that, throughout our analysis, we find hardly any effects of grid connection in RGCs without a mill. In other words, development effects are largely confined to RGCs with energy-dependent firms at baseline, which also see dramatically higher adoption and usage of electricity beyond the initial transformer.

5.3 Robustness

5.3.1 Alternative explanations

RGCs with and without a mill at baseline also differ along other observable dimensions, as documented above in Section 4.3. To understand whether these other differences can explain the difference in the impact of grid connection, we interact RGC grid connection with all other relevant factors in addition to the interaction with mills

[23] We group boreholes and tap water since both rely on pumps, which make them electrifiable.

in a single regression. Our core result is robust to including a wide array of interacted baseline observables in this way, as shown in Table A9. This lends further support to our hypothesis that grain mills are crucial anchor customers that enable village-wide electrification and local development upon grid arrival.

In particular, we additionally include interactions of our main treatment variable if a RGC received a grid connection with the following variables describing heterogeneity across villages at baseline: if the village has at least one public facility, at least one commercial facility other than the mill, at least one large-scale farm, at least one large-scale grain depot, as well as the RGC's minimum distance to a road, the population density of the census enumeration areas lying within 2km of the RGC, the enumeration areas' population, their wealth index derived from real consumption proxies, as well as the enumeration areas' share of households that use electricity for lighting, all measured in 2010. To ensure comparability with the estimated coefficients on the indicator variables for baseline heterogeneity variables, e.g., has a mill, has a public facility, etc., we transform all covariates relevant for baseline heterogeneity into indicator variables that reflect covariate values above the median value of a given covariate in our sample. For example, the interaction coefficient for distance to a road provides the estimate of an RGC getting electrified while being situated further from the nearest road in 2010 than the median RGC.

To exploit the longest possible panel dataset for this 'horse race' analysis of alternative explanations that the mill interaction effect could pick up, we employ either the mean (in columns 1-2) or maximum (in columns 3-4) nighttime luminosity as outcome variables, which are available annually from 2012-2020.

Although some interaction terms are statistically significantly different from zero, such as the indicator for the share of households using electric lighting being higher than in the median RGC, if the RGC has a public facility (both positively), or if it far from the nearest road, or if it has a place of worship (both negatively related with mean nightlights), the interaction effect on mill presence at baseline remains virtually unchanged, although statistical power decreases slightly. For maximum nighttime luminosity, population also exhibits a significantly positive interaction, but the interaction effect with mill presence at baseline likewise remains unchanged.

Therefore, although we cannot rule out alternative explanations to also help explain successful electrification in rural Zambia, it is noticeable that including various alternative interaction effects does not decrease the interaction effect with mills or reduce it to zero. In addition, no alternative explanation appears quantitatively more important than mills. One source of baseline heterogeneity, if the RGC has a public facility at baseline, appears very similar qualitatively and quantitatively to the mill, which is why we investigate the different subsamples of RGCs with mills, public facilities and those with both more closely in Section 6 below.

5.3.2 Alternative measures of treatment and baseline heterogeneity

Our core results are highly robust to various permutations of the empirical analysis, including alternative definitions of the treatment variable, alternative sample restrictions, alternative controls or alternative measures for baseline heterogeneity.

In particular, we show that our main result, reproduced in Table A10, Panel B, is robust to alternative definitions of the treatment variable, either as an indicator variable that takes a value of one if any school in a RGC reports grid arrival in a given year (Panel A), or a more extreme indicator that takes a value of one only if all schools in a given RGC are reporting grid arrival (Panel C). Although statistical significance varies, the interaction effect remains quantitatively unchanged and significant when using any school electrified, and qualitatively unchanged when using all schools electrified, although effects in terms of both mean and maximum hover around conventional levels of statistical significance.

Likewise, we can construct the panel of schools that inform RGC-level electrification status in three different ways: as balanced panel of all schools that existed at baseline in 2008 (preferred specification), as balanced panel for all schools that exist in 2020, with electrification values back-filled as zero for those school-years in which the school does not report (i.e. did not exist yet), or as unbalanced panel of all schools that ever exist. Since school construction may respond endogenously to electrification (or: school and grid construction occur at the same time) during our study period, we prefer the most conservative balanced panel of only pre-existing schools at baseline. Results are unchanged with less strict school panels underlying the electrification measure.

Schools are matched to RGCs in a laborious process described in briefly in Appendix A.III.1.1. To test robustness against this tailored process, we provide two additional checks: in Table A11 we compare our main result against a sample split of RGCs into only those villages that matched with a singled school (columns 3-4) or those that matched with multiple schools (columns 5-6). Results for both subsamples are qualitatively unchanged, although the subsample with multiple schools is driving the main effect. Second, in Table A12, we vary the strictness of the match quality between schools and RGCs, leading to ever smaller estimating samples from columns 2 to 5 (mean) and columns 7 to 10 (maximum), respectively. Results are surprisingly robust even for the extreme case of only considering school-to-RGC matches that lie within one kilometre of each other and have an exact name match to each other.

A recurring issue in the analysis of place-based treatments is uncertainty about the exact spatial extent of each village, especially if this extent is expected to change in response to treatment as indicated in Figure 6, Panel (a). Our approach attempts to remain as agnostic as possible, and let the data inform our choice of a reasonable average spatial extent of RGCs in Zambia. To this end, the spatial signature in Panel (b) of Figure 5 and the equivalent spatial signature for built-up area in Figure A17 indicate how the treatment effect of electrification in RGCs with mills slowly peters out across space, and is usually confined to around two kilometres around RGCs – confirming our preferred spatial extent of inconsequential RGCs in our sample.[24]

An alternative complication arises from the fact that the RGCs, as their name suggests, were supposed to be rural locations. Either due to mismeasurement, incomplete data, human oversight or political interference, the REMP proposes several RGCs for electrification that appear urban or peri-urban even at baseline. Our preferred approach restricts the sample of RGCs to only those that had below 1,500 people per square kilometre population density at baseline in 2010, a common classification of 'urban' settlements. However, in Table A13 we confirm robustness to three alternative sample definitions of 'rural' locations: population restrictions based on surrounding ward population, or two other maximum density restrictions.

Similarly to the issue of non-rural RGCs, our main analysis focuses on the subset of inconsequential RGCs, while we do not expect either high-priority ('target') RGCs or off-grid RGCs to adhere to the quasi-experimental ideal of inconsequentially-electrified locations. Nonetheless, the sign and magnitude of treatment effects in the full sample is still of interest: in Table A14 we confirm that inclusion of either target or off-grid RGCs does not materially affect results. Put differently, the planners do not appear to have predicted economic potential of electrification well.[25]

[24] Conversely, vegetation should approximately mirror this spatial signature which we report in Figure A18 for rings of vegetation around RGCs.

[25] Given the small sample sizes of the pools of consequential and off-grid inconsequential RGCs, we lack the degrees of freedom to perform analysis for each subsample alone using our main specification.

One concern that pertains to our empirical strategy is the possibility that other spatial shocks or treatments are co-determining outcomes since they may share the spatial or temporal signature of rural electrification in Zambia. Moneke (2023) highlights how large infrastructure expansions often arrive bundled, and that complementarities exist. Therefore, we control for time-varying phenomena such as road access, main road access (as key determinants of market access) and rainfall patterns (as key determinant of agricultural productivity and seasonality). Table A15 confirms that our results are qualitatively unchanged for market access (despite the large drop in sample due to only three time periods), and quantitatively unchanged for rainfall.

Finally, this paper focuses on energy-dependent firms and their potential contribution to construction of the local low-voltage network. The REMP engineers themselves pointed out the likely relevance of grain mills as energy-dependent, and the evidence on mills' load curves confirms their assumed disproportionately high expected benefits from electrification. However, the points of interest dataset allows us to test in an agnostic fashion for other categories of potential pre-existing characteristics of settlements that may result in the construction of a low-voltage network. In Table A16, we separately test for interaction effects of alternative facilities, such as commercial entreprises (e.g. shops, workshops, etc.), large farms, churches or places of worship, public facilities and large grain depots.[26] With the exception of public facilities (discussed in Section 6) below, none of the other categories of facilities present at baseline are significantly determining the effects of electrification.

6 Mechanisms

Through which channels do grain mills enable widespread electricity adoption and local economic development following grid arrival? We argue that both cost and demand externalities play an important role, and provide evidence supporting both of these mechanisms.

First, we examine the extent to which mills may lower connection costs for other villagers, thereby facilitating widespread adoption. In our endline survey, the primary barrier to electricity adoption reported by households and businesses is a high connection cost resulting from a large distance to the closest transformer or low-voltage

[26]Figure A13 provides the correlation matrix of these facilities in RGCs, highlighting the spatial overlap between, for example places of worship and commercial facilities such as shops.

line. Mills can lower the connection cost for others by contributing disproportionately to the construction of the local low-voltage network. After all, they benefit greatly from an electricity connection and have a higher ability to pay for it because they hold more capital than the average rural household. In extreme cases where other villagers are reluctant to get connected, they may pay for a private low-voltage line from the transformer to their mill. In other cases, they have been reported to help overcome coordination problems by lobbying their neighbours to join a group application in order to achieve a lower connection fee per end-user.

If mills enable electricity adoption through a cost externality, they should matter less in RGCs that have a low-voltage network for other reasons. To test this, we exploit the fact that the REA is mandated to connect all public facilities in RGCs reached by the grid. As a result, REA builds at least parts of the local low-voltage network in RGCs that have more than one public facility before handing the RGC over to the utility for any private end-user connections. In such RGCs, private connection costs are lower to begin with and the scope for a mill to reduce them further are limited. In line with this hypothesis, we find that both the presence of mills and the presence of public facilities beyond the school are associated with larger effects of grid connection on nighttime luminosity (Table 4, columns 1 and 2 as well as 4 and 5). Interestingly, there is no additional benefit of having a mill in the presence of multiple public facilities. In fact, in columns 3 and 6, the point estimate on the triple interaction between grid connection, the presence of a mill and the presence of multiple public facilities is negative (although not statistically significant), suggesting that mills and public facilities act as substitutes in facilitating electricity adoption.[27]

We corroborate the above finding of mills being substitutes with public facilities by showing that adoption does not only increase with the presence of a mill, or the presence of a public facility, but also in the number of these facilities at baseline. Table A17 confirms that every additional facility contributes to more adoption, and also highlights that mills and public facilities may be true substitutes, indicated by the negative and significant triple interaction effect in column 6.

Second, we provide evidence consistent with demand externalities and demonstration effects of mill connections. In particular, the impact of a cost externality should be most concentrated directly along the path from the transformer next to

[27] In column 3, the coefficient on the interaction of RGC connection and mill is not statistically significant at conventional levels, albeit nearly so with a p-value of 0.1203.

Table 4: Nightlights on RGC connection by mill, public and their interaction

	Mean nightlights			Max. nightlights		
	(1)	(2)	(3)	(4)	(5)	(6)
RGC connected	0.01	−0.01	−0.01	0.05	−0.04	−0.13
	(0.01)	(0.02)	(0.02)	(0.10)	(0.15)	(0.16)
RGC connected * mill	0.08**		0.05	0.64**		0.66***
	(0.04)		(0.03)	(0.27)		(0.25)
RGC connected * public		0.14***	0.14*		0.85**	0.94*
		(0.05)	(0.07)		(0.38)	(0.54)
RGC connected * mill * public			−0.03			−0.66
			(0.08)			(0.58)
RGC FE	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓
R^2	0.96	0.96	0.96	0.88	0.88	0.88
No. of years	9	9	9	9	9	9
No. of RGCs	742	742	742	742	742	742
No. of const-years	1215	1215	1215	1215	1215	1215
Observations	6678	6678	6678	6678	6678	6678

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC connected denotes percent of schools in RGC reporting grid electricity connection in a given year, derived from balanced sample of schools existing since 2008. Mill (public) denotes an indicator value that takes the value of one if at least one mill (public facility) situated within 2km of RGC. Mean and max. nightlights represent the mean and maximum, respectively, zonal statistics of annual nighttime luminosity values derived from the VIIRS VNL sensor for each RGC 2km buffer polygon. Const-year FE denotes constituency-year fixed effects.

the school to any public facility and the mill. It is under such a spanning tree that cost should be reduced most dramatically for any other end-user willing to connect. However, beyond the low-voltage network required to connect the school, public facilities and the mill, any increase in adoption likely reflects a genuine extension of the low-voltage network to connect other private end-users. Do mills also help increase adoption beyond the minimal low-voltage network required to connect them?

We test this hypothesis by exploiting the rich micro data at hand on the exact locations of schools, public facilities and any mills present at baseline. We construct hypothetical minimum spanning trees that would be required to connect the school as point of grid arrival, all public facilities and the mill in each RGC in the most costefficient fashion. We coin these minimum spanning trees the *predicted* low-voltage network for each RGC. Figure 7 provides two examples of RGCs and their predicted low-voltage networks if the school (as location of transformer, in the centre of the RGC), all public facilities and, in Panel (a), also the mill were connected to each other in cost-minimising fashion. We plot a 50m buffer around the predicted low-voltage network for illustration inside the 2km buffer around the RGC.

Darker colours of blue represent greater nighttime luminosity, either at baseline (left) or endline (right) in an illustrative RGC with a mill in Panel (a) and an illustrative RGC without a mill in Panel (b). In line with our cost externality interpretation, Panel (b) clearly shows the increase in nightlights in a tight area around the predicted low-voltage network in the centre of the RGC. In contrast, the RGC with mill experiences much broader nighttime luminosity responses much further from the predicted low-voltage network. In other words, given the realised luminosity increases, the unobserved, *real* low-voltage network that emerged between baseline and endline must be of substantially larger size than predicted. Even when accounting generously for potential 'blossoming' of nighttime lights, the RGC with mill displays vastly more nightlights in the uncovered part of the village, indicating effects beyond a pure cost externality in these locations. Figure A20 provides a second illustration of a different pair of RGCs confirming the same intuition.

With predicted low-voltage networks for every RGC at hand, we can systematically examine the effect of grid connection separately on the parts of the RGCs that are and are not covered by this minimum-spanning tree. We document three findings from this exercise, as shown in Table 5: first, our predicted low-voltage networks are highly predictive of nightlights increases in RGCs without mills, again strongly suggesting a cost externality to exist in the last mile of rural electrification. In these RGCs, any effect of electrification is solely confined to locations covered by the path required to connect public facilities, as confirmed in the null effects in the uncovered parts. Second, connected RGCs with mills see weakly more adoption in the parts of the RGC covered by the predicted low-voltage network, albeit not statistically significantly at conventional levels with p-values of 0.112 and 0.128, respectively. Third, connected RGCs with mills experience strong nightlights responses in the large areas of the RGC not covered by the predicted low-voltage network.[28] This strong result

[28]This result is robust to restricting the sample to only those RGCs with at least one mill or one public facility beyond the school, cf., Table A18, and to definitions of covered/uncovered parts that

from the uncovered part of RGCs with mills suggests that mills spur further electricity adoption, whether through demonstration effects or demand externalities.[29]

Table 5: Nightlights in parts of RGC covered or uncovered by low-voltage network (including mills) on RGC connection by mill status

	Entire RGC		Covered part		Uncovered part	
	Mean	Max	Mean	Max	Mean	Max
RGC connected	0.01	0.05	0.07***	0.12***	0.01	0.10
	(0.01)	(0.10)	(0.02)	(0.04)	(0.01)	(0.09)
RGC connected * mill	0.08**	0.64**	0.18	0.36	0.08**	0.58**
	(0.04)	(0.27)	(0.11)	(0.23)	(0.04)	(0.25)
RGC FE	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓
p-val $\delta = 0$	0.041	0.021	0.112	0.128	0.043	0.021
p-val $\beta + \delta = 0$	0.019	0.008	0.031	0.059	0.020	0.008
R^2	0.96	0.88	0.95	0.94	0.96	0.89
No. of years	9	9	9	9	9	9
No. of RGCs	742	742	742	742	742	742
No. of const-years	1215	1215	1215	1215	1215	1215
Observations	6678	6678	6678	6678	6678	6678

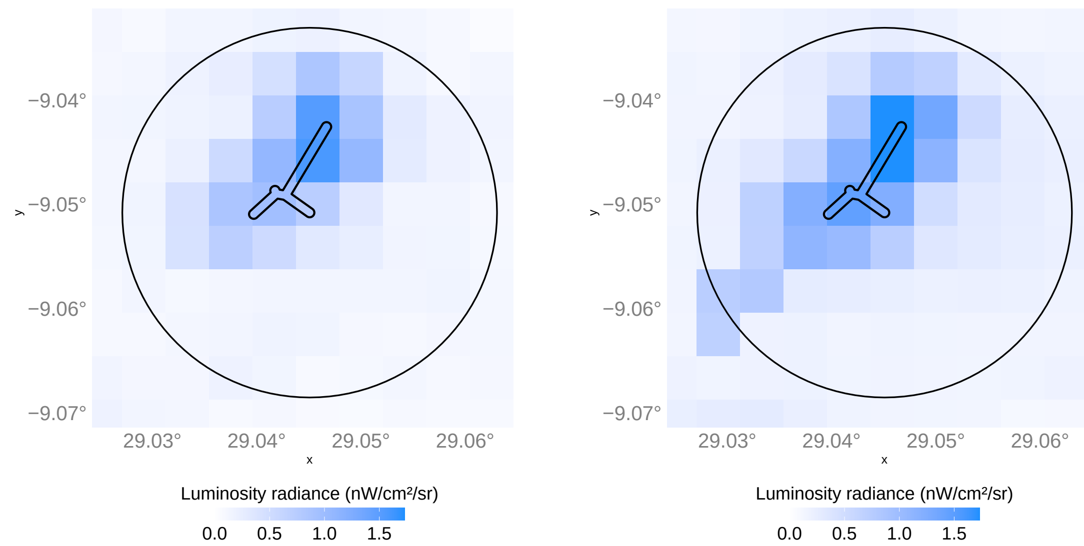
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC connected denotes percent of schools in RGC reporting grid electricity connection in a given year, derived from balanced sample of schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Covered part denotes the area within the 2km buffer around the RGC that is covered by a 25m buffer around the minimum-spanning tree. Uncovered part denotes the area within the 2km buffer around the RGC that is not covered by a 25m buffer around the minimum-spanning tree. Minimum-spanning trees represent a single, cost-minimising low-voltage network required to connect all schools, health facilities, public facilities and mills in a RGC to the grid. p-val δ reports the p-value on the interaction term with mill, whereas p-val $\beta + \delta$ reports the same for the overall effect for RGCs with mills.

In line with this conjecture, we also find that effects of grid connection on structural transformation are limited to RGCs with a mill at baseline and do not extend to RGCs with multiple public facilities – although both mills and public facilities contribute to wider adoption. As Table 6 highlights, effects on electricity adoption are nearly identical in RGCs with mills and multiple public facilities, but we only

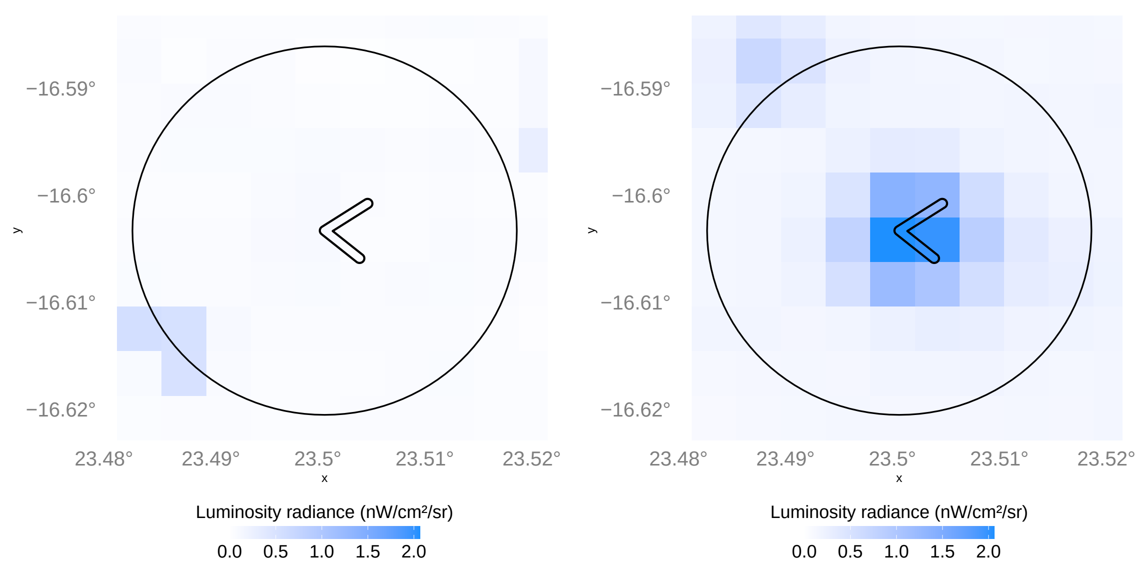
exclude the mill from the construction of the predicted low-voltage network, cf., Table A19.

[29]We discuss robustness of our uncovered part results against blossoming of nightlights in Appendix A.VI.2. Our results hold even when masking pixels up to 250m from the predicted network.

Figure 7: Effect of electrification on nightlights in parts of RGC covered and not covered by local low-voltage network



(a) RGC with mill



(b) RGC without mill

This figure illustrates the increase in nighttime luminosity in buffers of two kilometre radius around two grid-connected RGCs at the beginning and end of our samples. The figure also plots the minimum spanning tree of the low-voltage network in each RGC. This tree represents path a low-voltage network would have to take to connect all public facilities, schools, health facilities and mills, if present, in a cost-minimising fashion. We plot a 50m buffer around each RGC's minimum-spanning tree for illustration.

observe a shift of employment out of agriculture into services in RGCs with a mill.

Table 6: Adoption and sectoral employment on RGC connection by mill and public status

	Adoption				Sectoral employment					
	Light	Light	Cook	Cook	Agr.	Agr.	Man.	Man.	Ser.	Ser.
RGC connected	-0.08^{**}	-0.07^{**}	-0.01	-0.01	0.02	-0.03	0.01	0.04	-0.03	-0.01
	(0.03)	(0.04)	(0.01)	(0.02)	(0.09)	(0.11)	(0.07)	(0.08)	(0.04)	(0.05)
RGC connected * mill	0.23^{***}		0.06^{*}		-0.20^{*}		0.06		0.14^{**}	
	(0.09)		(0.03)		(0.11)		(0.09)		(0.06)	
RGC connected * public		0.18^{**}		0.05^{*}		-0.01		-0.04		0.05
		(0.07)		(0.03)		(0.14)		(0.09)		(0.07)
RGC FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Prov-year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.84	0.83	0.83	0.83	0.70	0.70	0.60	0.60	0.79	0.79
No. of years	2	2	2	2	2	2	2	2	2	2
No. of RGCs	226	226	226	226	226	226	226	226	226	226
No. of prov-years	20	20	20	20	20	20	20	20	20	20
Observations	452	452	452	452	452	452	452	452	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Prov-year FE denotes province-year fixed effects. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) that intersect with any sample enumeration area (SEA) sampled in both PHC 2010 at baseline and any LFS round between 2018-2022 at endline. RGC-level actual electrification derived from schools existing since 2008. Mill in RGC denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Adoption: Light denotes share of households in sample enumeration area (SEA) that report using grid-electric lighting. Adoption: Cook denotes the equivalent share of households that report using electricity as source of energy for cooking. Sectoral Employment: Agr. denotes share of individuals in sample enumeration area (SEA) that report agriculture as their primary occupation. Man. and Ser. denote the equivalent occupation shares for manufacturing and services sectors, respectively.

Our endline survey results allow us to shed more light on this transformation. When asked in 2025, millers report that connecting to the grid allowed them to expand their businesses. For example, a miller in Kamyunga recounted how grid arrival "allowed [him] to expand [his] already existing business by buying one more mill." Consistent with this, we find mills in grid-connected villages to be larger at endline than their peers in unconnected villages. As shown in Table A20, connected mills operate 0.26 more milling machines, have 0.37 more employees and produce weakly more output (columns 1-3). They also charge 12% lower milling fees (column 4), consistent with the impression of a connected miller that lower prices attracted more customers, and the expectation of an unconnected miller in our survey who argued that "being connected to power would also benefit the community members in such a way that they would pay less money for milling their produce."[30]

Finally, we find echoes of millers' qualitative answers indicating that electric mills break down less and require fewer repairs (cf., Figure A5): although not statistically significant due to the very small sample, some millers appear to run their mills for longer, and had to shut down for fewer weeks in the 2024 season (columns 5-6).

However, millers did not not only expand their existing milling businesses. They also set up new businesses in other sectors. In our survey, 38% of connected millers reported that switching to electricity for their mill allowed them to open new businesses (as shown in Figure A5) in the village.

Taken together, the evidence presented in this section suggests that mills enhance the local effects of grid connection through a combination of cost and demand externalities. On the on hand, their presence reduces connection costs for other villagers. On the other hand, they reinvest their profits from electrification in the local economy, spur competition in the milling sector and may exhibit demonstration effects for would-be productive users of electricity. Jointly these externalities do not only lead to widespread electricity adoption, but also to economic development.

7 Conclusion

When does electrification work at today's frontier in rural Sub-Saharan Africa? We find that grid-connected villages with pre-existing energy-dependent firms experience widespread adoption of electricity, and large medium-run development effects of electrification. Productive users, in our context grain mills, act as anchor customers for the utility, leading the utility to invest in local low-voltage infrastructure. Thanks to anchor customers, the construction of the 'last mile' low-voltage network generates large benefits of rural electrification for their villages.

In particular, we find that villages with pre-existing grain mills experience economically significant growth in the number of households, businesses and schools, and an accompanying increase in the area of the village covered by buildings. In line with the growth in businesses, occupational change out of agriculture and into services ensues, especially for women. Educational attainment, both in terms of literacy and years of schooling, improves significantly, especially for women and the young.

[30]Interestingly, the REMP engineers also foresaw this development, expecting that "price reduction of milling fees, as a result of market competition, may also occur with entrance of mill owners, which is also expected to trigger other secondary impacts" (Rural Electrification Authority, 2009).

Finally, households switch to cleaner water sources such as boreholes and tap water, pointing to increased employment of electric pumps.

How do anchor customers help translate the arrival of the grid to widespread adoption and development effects of electrification? Exploiting rich spatial information within villages and primary data collected from grain millers themselves, we provide suggestive evidence on two mechanisms: electrified grain mills appear to exhibit both a cost externality, decreasing the cost of connection for households and other businesses in their neighbourhood, and a demand externality, re-investing windfall profits from their switch to electricity into the village and spurring wider income effects through lower milling prices and greater competition.

In summary, this paper provides new insights into when rural electrification can work and why. By focusing our analysis on both firms and households, we provide a nuanced perspective of why electrification may generate transformative development effects in some, and near-zero adoption and development effects in other locations in the same empirical context. Therefore, our findings fill an important gap in the literature, while being highly relevant to the ongoing, accelerated grid expansion policy agenda that Mission 300 initiated across Sub-Saharan Africa.

Our key takeaway is that rural electrification can work in causing positive development effects in rural Sub-Saharan Africa – but probably not everywhere. We show that the targeting of rural grid expansion to villages with pre-existing productive uses is crucial, both due to the profits energy-dependent firms will reap from electrification, and the cost externalities they exhibit on their neighbours. Any utility should benefit from targeting villages with pre-existing energy-dependent firms first.

In contrast, of the large number of villages without productive use before grid arrival, many will take decades (Dinkelman et al., 2024) before maturing into economically viable customers for the utility that would warrant the high fixed cost outlay of building the local low-voltage network. In the meantime, off-grid technologies provide a valuable stepping-stone (Burgess et al., 2023). However, it remains an open question if the education and public health benefits derived from the isolated electrification of schools and health facilities could already warrant the cost of extending the grid to these villages – even in the absence of any private benefits. We leave a rigorous analysis of these social returns to grid electrification, as hypothesised by Lee et al. (2020b), for future research. Similarly, an investigation into the political returns of electrification as an alternative motivation for the (lack of) targeting of grid expansion is also beyond the scope of this work.

Ultimately, the double fixed cost hurdle of opening a business that could use electricity and contributing to the construction of the local low-voltage network means future work should investigate the conditions under which grid expansion to locations without anchor customers at baseline can create, not just amplify, economic activity. Energy-dependent firms on the cusp of entry may only be able to profitably enter once cheap grid electricity was available, but not without. This possibility highlights the value of taking potential complementarities of infrastructure provision into account.

References

Afrobarometer. (2024). Energy gaps: Slight, uneven progress still leaves many Africans without electricity. (Afrobarometer Dispatch No. 793).

Burgess, R., Greenstone, M., Ryan, N., & Sudarshan, A. (2023). Demand for electricity on the global electrification frontier. *London School of Economics mimeo*.

Burlig, F., & Preonas, L. (2024). Out of the darkness and into the light? Development effects of rural electrification. *Journal of Political Economy*.

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

Dinkelman, T., Collinson, M., Twine, W., & Wittenberg, M. (2024). Leaping or creeping up the energy ladder: Technology adoption in Africa. *Notre Dame U mimeo*.

Dinkelman, T., & Ngai, L. R. (2022). Time use and gender in Africa in times of structural transformation. *Journal of Economic Perspectives*, *36*(1), 57–80.

Elvidge, C. D., Zhizhin, M., Ghosh, T., Hsu, F.-C., & Taneja, J. (2021). Annual time series of global viirs nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sensing*, *13*(5), 922. https://doi.org/10.3390/rs13050922

European Commission, Joint Research Centre. (2023). *Ghs-built-s r2023a: Ghs built-up surface grid, multitemporal (1975–2030)*. https://doi.org/10.2905/9F06F36F-4B11-47EC-ABB0-4F8B7B1D72EA

Fetter, T. R., & Usmani, F. (2024). Fracking, farmers, and rural electrification in India. *Journal of Development Economics*, *170*.

Fried, S., & Lagakos, D. (2021). Rural electrification, migration and structural transformation: Evidence from Ethiopia. *Regional Science and Urban Economics*.

IEA. (2023a). Electricity grids and secure energy transitions. *IEA Publications*.

IEA. (2023b). SDG7: Data and projections. *IEA Publications*.

Kassem, D. (2024). Does electrification cause industrial development? grid expansion and firm turnover in Indonesia. *Journal of Development Economics*, *167*.

Kassem, D., Zane, G., & Uzor, E. (2026). Revisiting the last mile: The development effects of a mass electrification program in kenya. *Energy Economics*, *154*, 109098. https://doi.org/10.1016/j.eneco.2025.109098

Lee, K., Miguel, E., & Wolfram, C. (2020a). Does household electrification supercharge development? *Journal of Economic Perspectives*, *34*(1), 122–144.

Lee, K., Miguel, E., & Wolfram, C. (2020b). Experimental evidence on the economics of rural electrification. *Journal of Political Economy*, *128*(4), 1523–1565.

Lenz, L., Munyehirwe, A., Peters, J., & Sievert, M. (2017). Does large-scale infrastructure investment alleviate poverty? impacts of Rwanda's Electricity Access Roll-out Program. *World Development*, *89*, 88–110.

Lipscomb, M., Mobarak, A. M., & 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.

Masselus, L., Ankel-Peters, J., Gonzalez Sutil, G., Modi, V., Mugyenyi, J., Munyehirwe, A., Williams, N., & Sievert, M. (2025). Adoption of Electricity in Rural Rwanda 10 Years after Connection. *Nature Communications*, *16*, 10942. https://doi.org/10.1038/s41467-025-66986-0

Moneke, N. (2023). Can Big Push infrastructure unlock development? Evidence from Ethiopia. *University of Oxford mimeo*.

Ren, L. (2024). *Transportation, electricity, and the birth of firms* (Mimeo). University of Notre Dame.

Rud, J. P. (2012). Electricity provision and industrial development: Evidence from India. *Journal of Development Economics*, *97*(2), 352–367.

Ruhinduka, R. D., Bensch, G., Selejio, O., & Lokina, R. (2025). What Could Explain Low Uptake of Rural Electricity Programs in Africa? Empirical Evidence from Rural Tanzania. *Environment and Development Economics*, 1–19. https://doi.org/10.1017/S1355770X25000099

Rural Electrification Authority. (2009). *Rural Electrification Master Plan for Zambia 2008–2030: A Blue Print for Providing Electricity to All Rural Areas* (tech. rep.). Government of the Republic of Zambia. Lusaka, Zambia.

Trimble, C., Kojima, M., Perez Arroyo, I., & Mohammadzadeh, F. (2016). Financial viability of electricity sectors in Sub-Saharan Africa. *World Bank Policy Research Working Paper*, (No. 7788).

VITO. (2026). *Spot-vgt s1 daily synthesis surface reflectance and ndvi - v2*. Flemish Institute for Technological Research. https://doi.org/10.71570/gffd-7284

World Bank. (2026a). Mission 300 — Electricity to Power Africa and its Economy [Accessed on July 3].

World Bank. (2026b). Under mission 300, a new way of doing business connects over 50 million people to electricity across Africa. *World Bank Group*, (Press Release No. 2026/130/AFE).

Online Appendix

A.I The last mile of electrification in Sub-Saharan Africa

A.I.1 Data

To examine local electricity adoption patterns across Sub-Saharan Africa, we leverage microdata from nine different nationally representative household surveys, as shown in Table A1. In each survey, our analysis focuses on in grid-connected villages in rural areas. Survey clusters are defined as grid connected if surveyors indicate the presence of an electricity grid in the village in the community questionnaire. In surveys where this measure is not available, we instead define a village as grid connected if at least one household in the rural survey cluster reports a connection to the national electricity grid, as indicated in column 4.[31] The number of sampled grid-connected villages ranges from 32 to 234 across surveys (column 5). The share of connected households in a cluster is the number of connected households over the total number of surveyed households, where the latter varies between 8 and 20 across surveys (column 6).

Table A1: Data sources for bimodality in electricity adoption

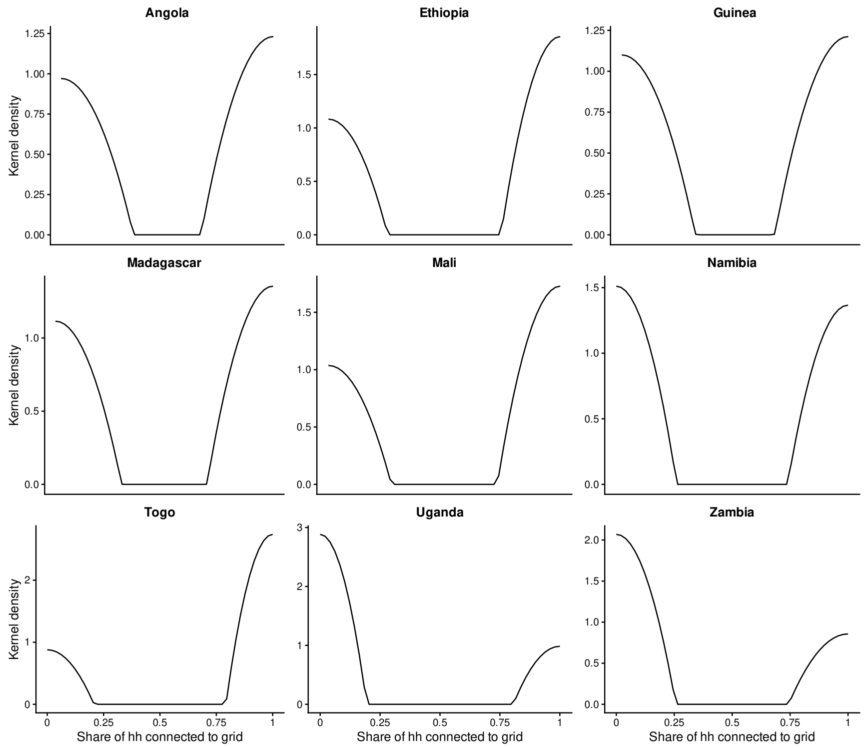
Country	Survey	Year	Village connection measure	Connected villages	Hhs/village
AGO	IDREA	2018	Any household connected	52	8
ETH	ESS	2018	Grid observed	124	11
GIN	DHS	2018	Any household connected	50	20
MDG	ADEIE	2011	Any household connected	33	13
MLI	EHCVM	2018	Grid observed	32	12
NAM	HIES	2015	Any household connected	234	12
TGO	EHCVM	2018	Grid observed	110	12
UGA	NPS	2015	Any household connected	64	11
ZMB	LCMS	2015	Any household connected	100	15

Connected villages (col. 5) indicates the number of rural, grid-connected survey clusters in the sample. Hh/village (col. 6) indicates the median number of households surveyed per rural cluster.

[31] Under this alternative definition, we misclassify grid-connected villages in which no surveyed household reports a connection as unconnected. As a result, we underestimate the share of villages with low connection rates.

A.I.2 Robustness

Figure A1: Electricity adoption in grid-connected villages across SSA: Afrobarometer



This figure shows the distribution of the share of connected households across grid-connected villages in nine Sub-Saharan African countries. All underlying data is from Afrobarometer Rounds 8 to 10.

To check the robustness of the bimodality in electricity adoption across grid-connected villages documented in Figure 1, we replicate our results using alternative source data from Afrobarometer. For each survey cluster, surveyors record whether they observe the electricity grid. Moreover, from round 8 onwards, households are asked whether or not they are connected to the national electricity grid. Since Afrobarometer surveys are much smaller in sample size than the household surveys used for Figure 1, typically including 150 survey clusters in total (including grid-connected and unconnected clusters in both rural and urban areas) with 8 households per cluster, we pool data from rounds 8 to 10. Figure A1 illustrates a similar bimodal pattern as shown in Figure 1 even though patterns somewhat differ for individual countries. One potential reason for this is the passage of time between the household survey data used for Figure 1 (2011-2018) and the Afrobarometer data used for Figure A1 (2019-2024).

A.II Rural electrification in Zambia

A.II.1 Grid expansion

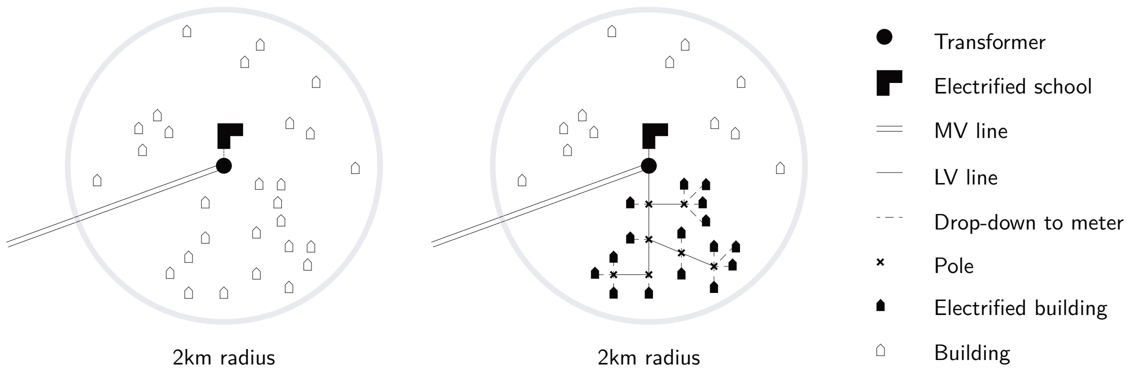
Figure A2: Pictures of local electric trunk network in rural Zambia

(a) Transformer under medium-voltage line, spanning low-voltage trunk network

(b) Low-voltage trunk network pole and end-user connections

This figure shows different parts of local trunk networks in rural Zambia. The pictures were published on the official Facebook account of the Rural Electrification Authority.

Figure A3: Sketch of village with/without local electric trunk network in rural Zambia



This figure shows a sketch of a grid-connected village without (left panel) or with (right panel) a local low-voltage network. In both scenarios, the medium-voltage (MV) line reaches the village and, by government mandate, the school has to be electrified using a transformer to step-down voltage. In the right panel, the utility builds additional poles, low-voltage lines and drop-downs to electric meters to electrify various buildings, including households and businesses.

A.II.2 The role of grain mills

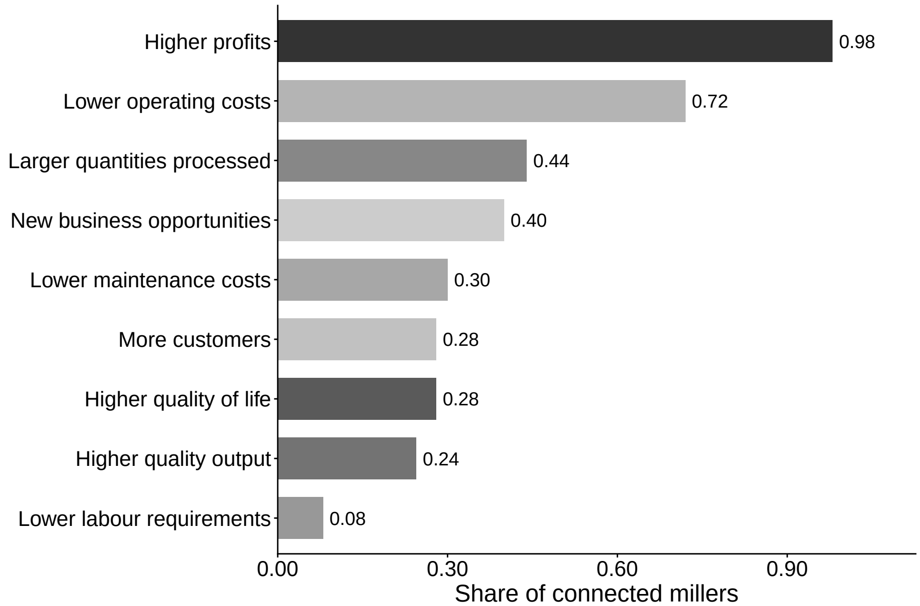
Figure A4: Pictures of standard electric and diesel grain mills in rural Zambia

(a) Electric grain mill

(b) Diesel grain mill

This figure shows an electric and a diesel grain mill in rural Zambia. The pictures were taken by the authors.

Figure A5: Benefits of electrification reported by grid-connected millers



This figure displays the most commonly mentioned advantages of mill electrification cited by connected millers in our 2025 village and miller survey (Appendix A.III.2.3). We code open-ended responses to the question "What were the top three advantages of getting connected to the electricity grid?" into the categories shown.

A.II.3 The Rural Electrification Master Plan

A.II.3.1 Identification and prioritisation of RGCs

The REMP identified 1,200 villages that were deemed important centres of economic activity in rural areas, yet remained unelectrified. This list of villages was created in a consultative process with district officials from the entire country. The villages on the list are referred to as rural growth centres (RGCs).

RGCs were prioritised for electrification based on their expected total electricity demand in 2030. Demand was essentially assumed to be a linear function of population in 2006 even though demand forecast followed a somewhat more elaborate process, as described in detail by the Rural Electrification Authority (2009). As illustrated in Figure A6, predicted demand directly determined RGC priority, with the exception of the top eight RGCs. The top eight RGCs were given priority over all other RGCs independent of their expected demand because they were deemed particularly important by their respective province administrations.

Figure A6: Prioritisation of RGCs for electrification

Table 5-11 Temporary Electrification Priority of RGCs Based on Demand Criteria (1/13)

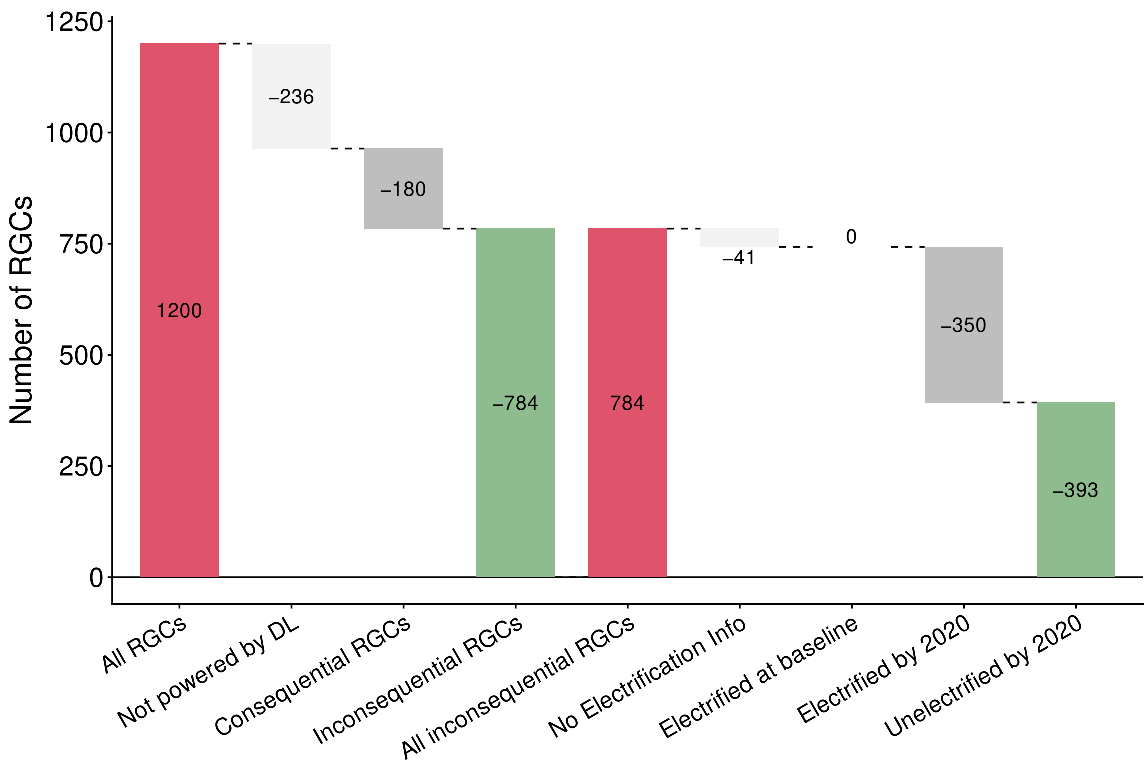
Ranking	RGC	District	Priority	Province	# of HHs (2006)	# of HHs (2030)	Daily Max Load
1	Mpulungu Central	Mpulungu	4	Northern	2,000	3,972	2,200,731
2	Mwinilunga Boma	Mwinilunga	0	North-Western	1,900	3,774	2,093,225
3	Shangombo	Shangombo	1	Western	1,100	2,185	1,277,541
4	Boma	Luangwa	1	Lusaka	580	1,152	752,118
5	Chienge	Chienge	6	Luapula	560	1,113	642,046
6	Mpongwe	Mpongwe	22	Copperbelt	441	876	499,270
7	Nsama Sub Boma	Kaputa	1	Northern	441	876	499,270
8	Talayi	Milenge	1	Luapula	202	402	241,663
9	KPG Market	Kapiri Mposhi	22	Central	7,400	14,697	8,141,484
10	Chisanga	Kasama	9	Northern	5,000	9,930	5,530,061
11	Chindenza School	Katete	3	Eastern	5,000	9,930	5,515,225
12	Mtandaza RHC	Katete	5	Eastern	5,000	9,930	5,509,327
13	Kagoro	Katete	1	Eastern	4,000	7,944	4,401,462
14	Sikapila	Mporokoso	2	Northern	3,646	7,241	4,012,551
15	Kauwe	Kazungula	14	Southern	3,511	6,973	3,927,399
16	Palace Chipepo Mukuni-Ngombe	Kapiri Mposhi	23	Central	3,500	6,951	3,847,529
17	Twapia	Ndola	2	Copperbelt	3,333	6,620	3,735,864
18	Nchembwe	Kapiri Mposhi	20	Central	3,100	6,157	3,416,570
19	Kapungwe	Petauke	1	Eastern	3,084	6,125	3,407,622
20	Nyamphinga	Petauke	10	Eastern	3,084	6,125	3,405,327
21	Chikalawa	Petauke	12	Eastern	3,084	6,125	3,402,990
22	Mwanjawanthu	Petauke	3	Eastern	3,036	6,030	3,345,943
23	Kasenengwa Rural Centre	Chipata	5	Eastern	3,000	5,958	3,321,933
24	Kamphambe	Katete	7	Eastern	3,000	5,958	3,308,596
25	Sikatengwa	Lundazi	22	Eastern	2,949	5,857	3,246,409
26	Mushili	Samfa	18	Luapula	2,751	5,464	3,032,798
27	Madimawe Rural Health Centre	Chipata	8	Eastern	2,667	5,297	2,953,674
28	Matonje	Petauke	13	Eastern	2,587	5,138	2,851,518
29	Mumbi	Petauke	5	Eastern	2,503	4,971	2,762,099
30	Kawama East	Mufulira	6	Copperbelt	2,448	4,862	2,732,408
31	Chimutende,	Katete	2	Eastern	2,472	4,910	2,728,967
32	Thendere	Isoka	4	Northern	2,400	4,767	2,698,020
33	Ntipo	Isoka	7	Northern	2,420	4,807	2,685,921
34	Chinkhombe	Katete	13	Eastern	2,385	4,737	2,634,234
35	Mushindomo	Solwezi	21	North-Western	2,380	4,727	2,628,469
36	Mulekatembo	Isoka	5	Northern	2,350	4,667	2,618,045
37	Kaula	Kabompo	14	North-Western	2,350	4,667	2,585,437
38	Lukulu Township	Lukulu	1	Western	2,012	3,996	2,512,048
39	Nande	Senanga	8	Western	2,148	4,266	2,393,483
40	George Camp	Ndola	4	Copperbelt	2,165	4,300	2,383,973

This figure shows the top 40 RGCs prioritised for electrification in the REMP. Source: Rural Electrification Authority (2009).

A.II.3.2 Project package creation

The 1,200 RGCs are grouped into 180 project packages based on geographical proximity and existing infrastructure. Each package contains one highly ranked RGC that serves as the primary target to be powered by distribution line extension from the nearest substation, and other unelectrified RGCs with lower priority that fall along the electrification route are also included in the same package. Package size varies between 1 and 22, with an average of nine RGCs per package.

Figure A7: Electrification treatment status across (inconsequential) RGCs by 2020

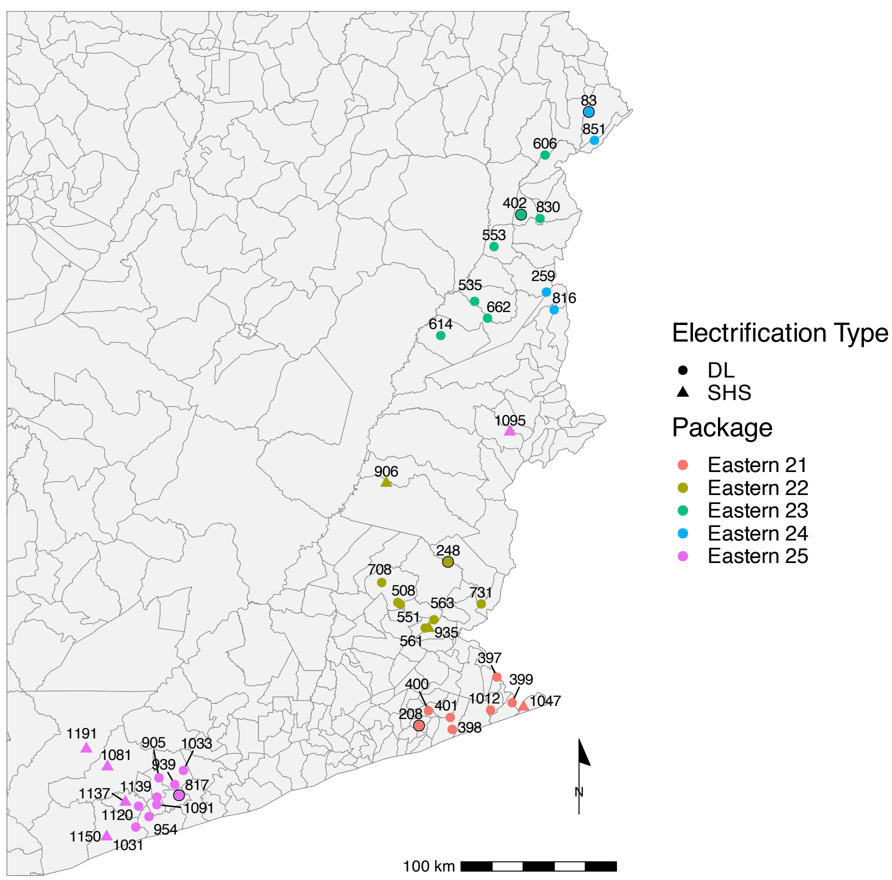


This figure illustrates the breakdown of the full set of RGCs into those to be electrified by distribution line (DL) and those to be electrified off-grid. Among the ones to be electrified by DL, it distinguishes between target RGCs and inconsequential RGCs. Finally, it visualises how many inconsequential RGCs receive a grid connection by the end of our sample in 2020.

Projects with numerous unelectrified RGCs are further divided into components to be electrified via either transmission/distribution line extension or stand-alone electrification modes (e.g. solar home system, mini-hydro power development, or diesel generator) in order to cut down inefficient distribution line construction and optimize resource utilization. JICA conducted case studies for all project packages to determine the optimal electrification method for each RGC based on the unit life time cost in net present value (US$/kWh). As a result of this, 79.9% of RGCs, corresponding to 94.5% of households, were scheduled to be electrified by distribution line extension, whereas 19.8% of RGCs fell under SHS electrification and only 0.3% (4 RGCs) were planned for mini-hydro electrification. Figure A7 illustrates this breakdown.

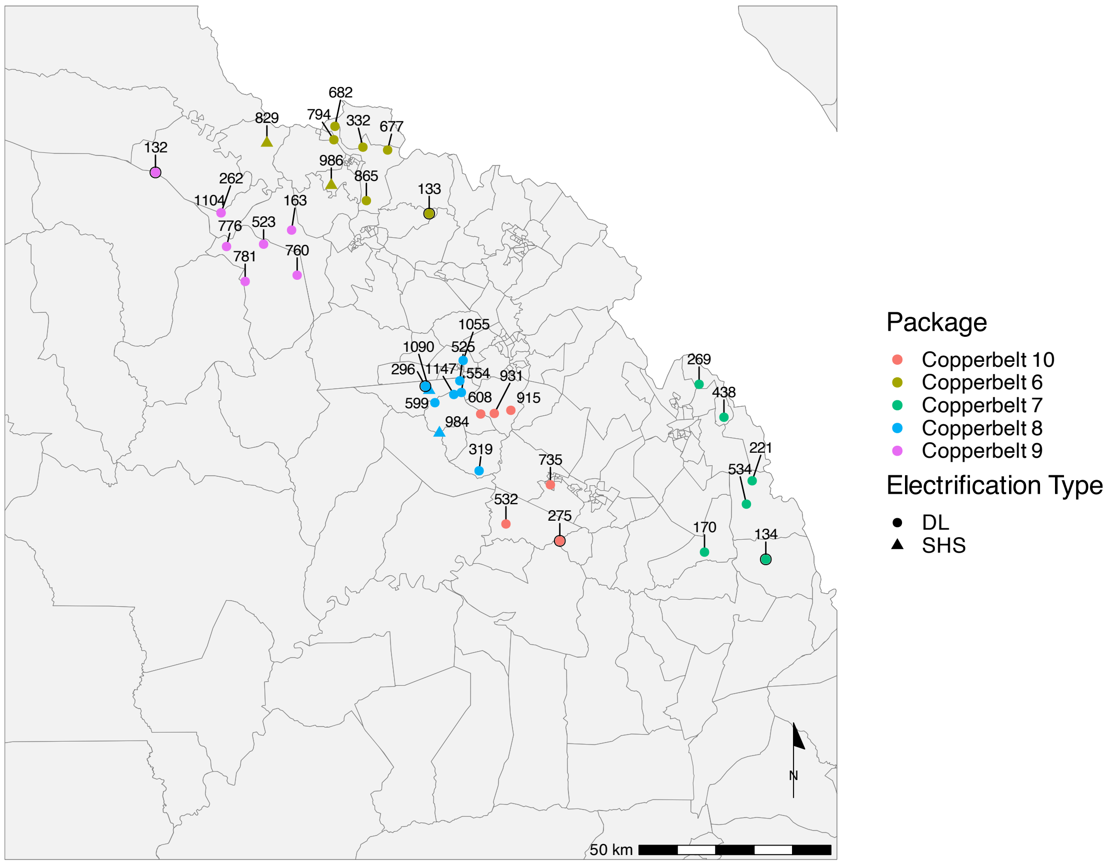
The resulting allocation of RGCs to project packages and electrification modes is illustrated in Figures A8 and A9, focusing on five packages in Eastern Province and Copperbelt, respectively. It highlights the highest ranked RGC in each package (black circle around coloured dot), the inconsequential RGCs (coloured dots), and un-selected RGCs that are allocated to off-grid solutions and exit our sample (triangles). Ranking numbers are displayed as labels next to each RGC, confirming that inconsequential RGCs are of vastly lower priority (i.e., higher rank) than the target RGC.

Figure A8: Inconsequential units in practice: packages #21–#25, Eastern Province



This figure illustrates the grouping of RGCs into five electrification packages in Eastern Province, and indicates the electrification mode chosen for each RGC.

Figure A9: Inconsequential units in practice: packages #6–#10, Copperbelt Province



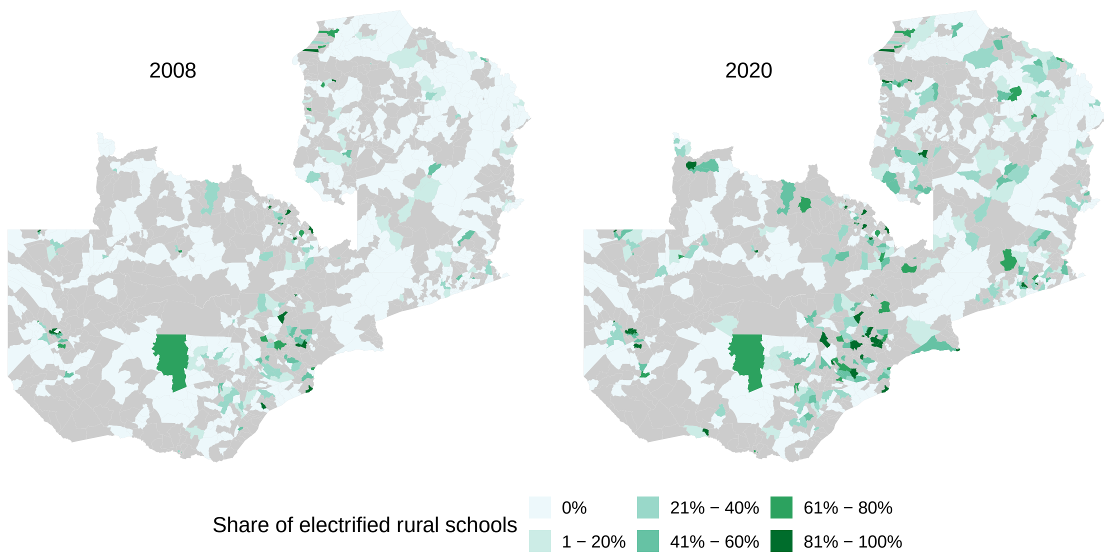
This figure illustrates the grouping of RGCs into five electrification packages in Copperbelt Province, and indicates the electrification mode chosen for each RGC.

After the determination of the project packages, the packages themselves were prioritised for electrification priority based on financial indicators (Financial Internal Rate of Return and Economic Internal Rate of Margin).[32] Figure A10 illustrates the resulting RGC electrification process, showing the share of electrified schools in inconsequential RGCs across wards in 2008 and 2020.

As Table A2 shows, the design of the REMP implies that, within constituency, inconsequential RGCs that are connected during our study period do not systematically differ from the ones that are only connected after our study period in terms of priority rank, baseline size or expected electricity demand.

[32]The derivation of these rates assumes a constant growth rate of 2.9% for the number of customers across all types. An implication of this is that electrification will not create new customers of a given type if an unelectrified RGC starts with zero customers of this type.

Figure A10: Electrification of inconsequential rural growth centres



This figure illustrates the electrification of inconsequential RGCs across Zambian wards. The left- and right-hand map show the share of electrified schools in inconsequential RGCs in 2008 and 2020, respectively. Darker shades of green indicate higher electrification rates. Grey shaded areas indicate wards that do not contain inconsequential RGCs.

Table A2: REMP baseline predictors of RGC electrification by mill status

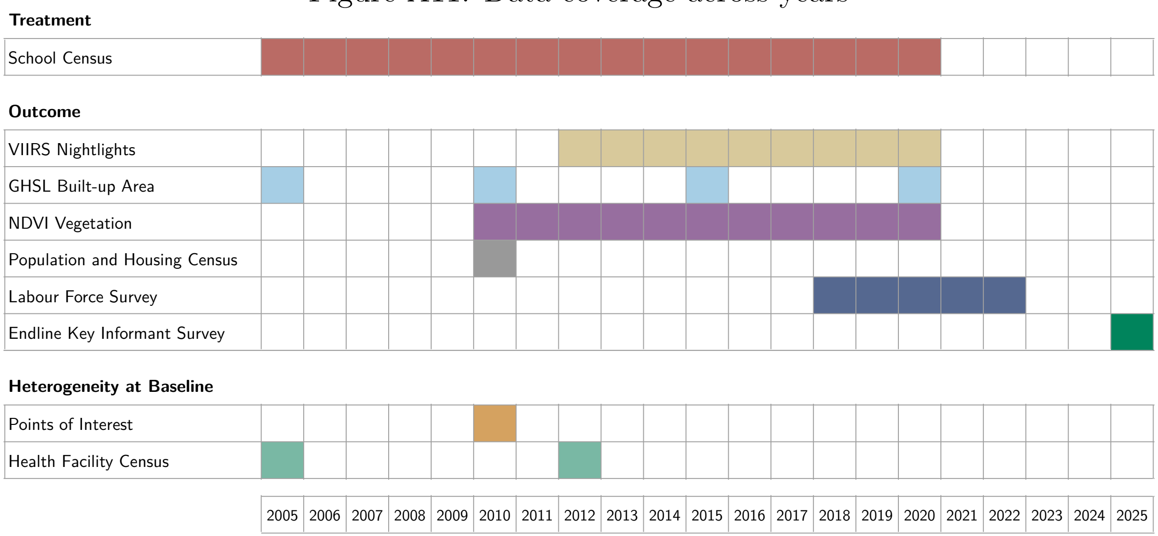
	RGCs with mill			RGCs without mill		
	Treatment	Std. Err.	N	Treatment	Std. Err.	N
Number of households, #	−27.15	74.60	186	26.90	34.70	556
Expected peak electricity load, kW	−23.39	82.01	186	30.71	38.14	556
RGC priority, rank	−10.34	45.58	186	16.23	31.14	556

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at the electrification package-level. Outcome variables at baseline sourced from Rural Electrification Authority (2009), including estimated number of households in Rural Growth Centres (RGCs) at baseline, expected peak electricity load at baseline as a function of, among others, number of households and, resulting ranking of RGC at baseline according to the master plan. Treatment denotes the coefficient of a regression of baseline REMP variables on a dummy variable that takes the value one if RGC ever received an electricity connection during the sample period (2008-2020).

A.III Data

This section provides details on all the data sources used for this paper. Figure A11 provides an overview of all of these, indicating how they are used in the analysis and for which years they are available.

Figure A11: Data coverage across years



This figure shows provides an overview of the data used in this paper. For each dataset, it indicates the years for which it is available and the role it plays in the analysis.

A.III.1 Measuring village electrification

A.III.1.1 Annual school census

We leverage the Annual School Census conducted by the Zambian Ministry of General Education to determine the time of arrival of the grid in each RGC. Below we describe how we extract the school census data from the Ministry's Education Information Management System (EMIS), build a school-year panel, clean all relevant variables and geo-identify all schools.

Data extraction and cleaning. The raw data from the Education Management Information System consists of a Microsoft Access database for each year from 2004 to 2020. Each of these relational databases contains many different data tables, and the encoding of value labels for all variables. To clean the data we proceed as follows.

1. Data extraction: We export the data tables `gdb_school` and `gdb_waterelec` from the Access database for each year from 2004 to 2020 to Microsoft Excel. Additionally, we export the value label tables for all relevant variables and years.

2. Data import: We import each of the data tables into statistical software. Within a given year we merge records using the variable "code" as the unique school identifier. Next, we import the labels associated with the values of each variable from the corresponding label tables. In cases where value labels change across years, we harmonize these. The available information allows us to generate an indicator for each of the following sources of power – grid, generator, solar, hydro – for each year and school. Finally, we append the data from all the years.

3. Harmonizing school identifiers: We clean school identifiers. In principle, there are three sets of variables that can be used to identify schools: (i) school number, (ii) school code, (iii) school location (district) and school name. We begin by examining variation of school code within school number and ensure that changes in code within number do not reflect changes in school identity, judging from the wealth of information we have about each school in each year. Next, we examine variation in school location and school name within school number. To this end, we first harmonize province and district names across EMIS years such that they are aligned with the 2010 population census which featured 72 districts across 9 provinces. Second, we clean up the spelling of school names. Then we ensure that changes in school location and name within school number do not represent changes in school identity. Finally, we check for the presence of duplicates comparing different school numbers in the same location that are associated with similar school names to each other. In the resulting dataset each entry is uniquely identified by school number and year.

4. Adding school coordinates: The EMIS data contains GPS coordinates of a subset of schools that were collected by the Ministry of Education in 2007. We leverage additional school location information from other initiatives that collected school coordinates over the years and merge these to our dataset. The sources of these additional coordinates are: ZEMA, FSDZ, Ministry of Health, IPA (community health worker project), DFID and the World Bank (ZEEP). Each of these sources is a list of schools with GPS coordinates. Using

a combination of computational and manual methods, we match these to our data using district, school name and school number. This leaves us with up to seven sets of coordinates per school.

Geo-identification. We have location information for each school, namely province, district, constituency and ward as reported by the school headmaster. For 4,598 out of 14,395 schools (31.94%), the headmaster also reported a GPS coordinate for their school. To fill in the missing coordinates, we leverage six additional data sources on school coordinates, as detailed above. To resolve contradictions between coordinates from different sources for a given school, we proceed as follows. First, we develop a measure of data quality for each source by counting how often coordinates correspond to a school building visible in satellite imagery in a random subset of the data. Second, we score coordinate pairs as follows. We run pairwise comparisons between all sets of coordinates available for a given school. Coordinate pairs receive 1 point for each other coordinate pair that is within 5.55km. Finally, we use the highest ranking coordinate pair as the school location in our analysis. If two coordinate pairs score equally high, we choose the source with the higher quality, as determined in the first step.

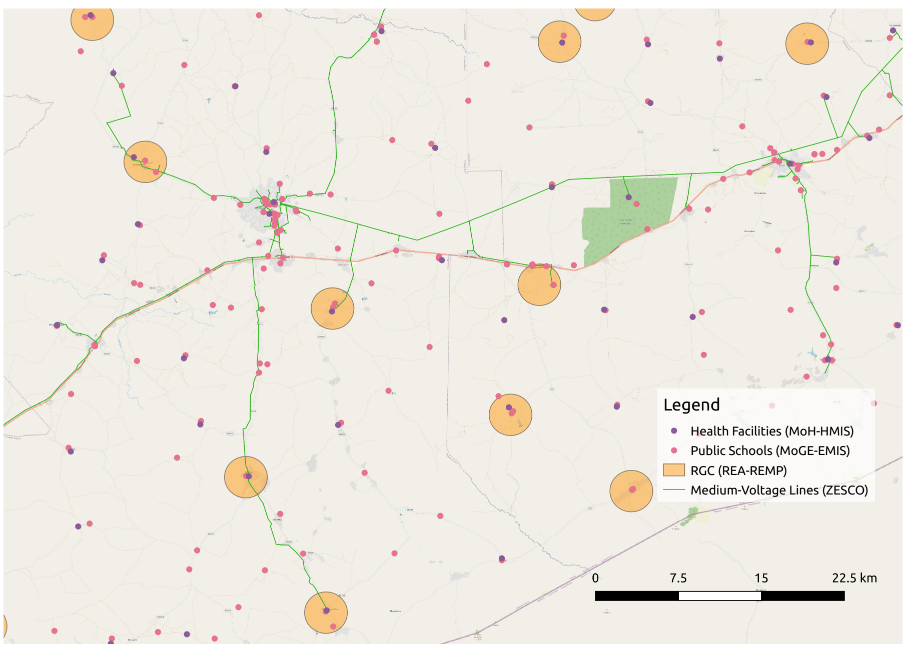
Linking schools to rural growth centres. To find all schools in a given RGCs, we rely on school coordinates as well school names and the name of the ward they are located in. We link each RGC to all schools that fall within 2km of the RGC centroid. Additionally, we match all schools with an identical name as the RGC within up to 10km to account for the uncertainty about school coordinate quality in some parts of the country. Finally, we leverage auxiliary information from GRID-3 Schools, GRID-3 Settlements, Google Maps, Google Earth and Open Street Map to validate our matches. This way, we match a total of 4,186 schools to 1,177 RGCs, out of which 54% are within 1km of the RGC centroid and 98.5% within 5km.

A.III.1.2 Cross-validation using alternative data sources

We use four alternative data sources to cross-validate the village electrification measure constructed from the school census: a ZESCO medium-voltage line map from 2023, our own village survey from 2025, health facility censuses from 2005, 2012 and 2017 and REA project records from 2008 to 2018. None of these alternatives covers the entire sample of RGCs over time, however. Therefore, they cannot be used for our main analysis.

The ZESCO map of medium-voltage lines from 2023 is not available for any earlier time period. Hence, it does not allow us to track changes in RGC electrification over time. We can only use it to cross-validate RGC electrification status at endline by checking whether a medium voltage line intersects with a 2km buffer around RGCs that we code as electrified in 2020 based on the school census, as illustrated in Figure A12. Since there may have been additional electrification between 2020 and 2023, we expect the map to show some villages as electrified even though they are not recorded as such in our data in 2020.

Figure A12: ZESCO medium-voltage lines connecting RGCs in 2023



This figure shows an excerpt of the 2023 ZESCO map of medium-voltage lines from Petauke. RGCs are shown as yellow circles with a 2km radius. Purple and red dots indicate health facilities and schools, respectively.

Our own village survey was conducted in 2025, as described in Appendix A.III.2.3. Similar to the ZESCO map, we can use it to cross-validate RGC grid connection in the cross-section, although this is imperfect because the time gap between the survey and the last year of the school census amounts to five years.

The REA project database contains aggregate information on construction projects that usually involve multiple RGCs. However, the database can typically only be linked to a subset of the involved RGCs due to missing information. We observe the year in which the entire construction project was completed and use this as a proxy for the time of grid arrival in the matched RGCs.

Health facility censuses are only informative about RGCs that contain at least one health facility (669 out of 742). Moreover, health facility censuses are not conducted annually, but were only implemented in 2005, 2012 and 2017 during our study period. We follow a similar procedure as described for the school census above to link health facilities across years, geo-identify them, obtain their electrification status and match them to RGCs.

Table A3 shows how the electrification information from the different alternative data sources correlates with our school-based measure, both across RGCs in the cross-section (OLS) and within RGC over time (DD). In all specifications, the regressor is the share of schools in the RGC that are electrified. In columns 1 to 4, the outcome variable is an indicator that takes the value one if the RGC is electrified based on the indicated data source, and zero otherwise. In columns 5 and 6, it is the share of health facilities in the RGC that is electrified. Reassuringly, we find a strong positive association between the school-based electrification measure and all the alternative measures, both in the cross-section and within RGC over time.

Table A3: School-reported vs alternative measures of RGC electrification

	ZESCO	Village Survey	REA Project		Health Facility	
	2023	2025	2018	2008-2018	2017	2005/2012/2017
School Electrification (%)	0.68***	0.63***	0.40***	0.42***	0.95***	0.67***
	(0.06)	(0.05)	(0.12)	(0.09)	(0.05)	(0.07)
Year FE				✓		✓
RGC FE				✓		✓
R^2	0.29	0.24	0.26	0.78	0.47	0.79
No. of Years				11		3
No. of RGCs				40		669
Observations	743	468	39	427	699	1707

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. School Electrification (%) denotes RGC-level electrification derived from schools within 2km from RGC that exist continuously since 2008, covering the same year(s) as outcome variable, at most 2008-2020. ZESCO transformers' electrification status is reported as single cross-section in 2023 at RGC-level (including a minority of REA transformers not yet handed over to ZESCO). Village survey electrification is measured as key informant reporting the presence of an electric transformer in the village in 2025. REA projects' electrification status is reported at electrification package-level in selected project records from 2006-2018 and matched to RGC-level via schools mentioned in project records. Health Facility electrification is reported at health facility-level and aggregated to RGC-level, available for 2005, 2012 and 2017. In all cross-sectional regressions, only the final year of outcome data is used.

A.III.2 Outcome data

A.III.2.1 Satellite-based data products

Nightlights. We employ a standard VIIRS VNL nighttime luminosity product that provides annual nighttime luminosity radiance in nW/cm2/sr on a scale from 0-73 from 2012-2020. The spatial resolution is 350m*350m, and we extract mean and maximum nighttime luminosity values from the annual average composite product. Data is pre-processed and due to sensor correction marginally negative values are possible in unelectrified areas such as in rural Zambia. Due to the incomplete coverage in the sensor's first year, 2012, and due to the global effects of the Covid-2019 pandemic in 2020, the first and last years the sensor correction is disproportionate in comparison to other years. Year fixed effects in all regressions control for this variance in correction. For the raw data plotted in Figure 7 and Figure A20, we plot 2013 as baseline and 2019 as endline to avoid residualising the raw data on year fixed effects to make the outlier years 2012 and 2020 comparable.

Built-up area. We employ a standard GHSL-S product that provides square metres of built-up surface of a given 100m2 area. We calculate the mean and sum of built-up area zonal statistics for each RGC 2km buffer for the four years, 2005, 2010, 2015 and 2020. The sum of built-up area is divided by the area of the RGC 2km buffer to provide a percentage of total area covered by buildings interpretation.

Vegetation. We employ a standard Normalised Difference Vegetation Index (NDVI), constructed as $NDVI = (NIR - Red)/(NIR + Red)$, where NIR denotes the reflectance in the near infrared bands, and Red to the reflectance in the red bands.

We use the NIR and Red bands from SPOT NDVI, and construct annual composites in the following way. For every year from 2010-2020 we take the average of the daily composite of three separate days in March (March 1st, 11th and 21st). For 2010-2013, this is derived from SPOT-VGT at 1km resolution, with the underlying data originating from SPOT-4, which stopped working in July 2013. For 2014-2020, we use the successor product, PROBA-V. This technically has a 300m resolution, but we use the 1km resolution for consistency with the earlier period, using the same procedure of taking the average of three separate daily composite NDVI values for the same three days in March every year. Zonal statistics for mean and maximum

NDVI calculated at the RGC 2km buffer for every year.

A.III.2.2 Labour force surveys 2018-2022

We use microdata from the quarterly labour force surveys (LFS) 2018-2022 that was kindly provided by Zamstats for this study. Key variables of interest include the use of electricity for lighting and cooking and the main source of drinking water at the household level as well as education and employment outcomes at the individual level. We harmonize all of these variables such that they are comparable to those collected in the 2010 Population and Housing Census (PHC).

We leverage information on the 2010 census enumeration area in which each surveyed household resides to link LFS data to RGCs. We assign an enumeration area to an RGC if it intersects with 5km buffer around the RGC. In a robustness check, we reduce the buffer radius around the RGC to 2km.

A.III.2.3 Village and miller survey

In 2025, we conducted a key informant survey across RGCs to gather qualitative and quantitative evidence on the village electrification process and the role of grain mills. The survey was conducted via telephone and respondents were typically local head teachers. They were primarily asked about household and business electricity connections in the village. In a subset of RGCs, we also identified mill owners and spoke with them about their experience of the grid connection of their village. This section provides details on the design and implementation of the survey.

Sampling. The sampling frame for the survey was constructed from the universe of rural growth centres (RGCs) listed in the 2009 Rural Electrification Master Plan. We excluded all RGCs that were that were the principal target of their respective grid construction package, as well as RGCs earmarked for off-grid electrification, and focused on inconsequential RGCs to be electrified by grid only. From the resulting set of 784 RGCs, we randomly drew 661 RGCs for the survey.

Questionnaire. The survey was composed of two parts, a key informant survey and a miller survey. The core of the key informant survey was implemented in all villages while the miller survey and an extended key informant survey were only implemented in a subset of 186 villages, all of which had a mill at baseline.

The core key informant survey collected information on schools, health facilities and businesses in the village and their grid connection. Additionally, data on village size, electricity infrastructure, such as poles and transformers, and household connections was gathered. In the extended key informant survey, open ended questions were added to learn how electrification had unfolded in the village and what barriers had been encountered. The miller survey asked questions about mill operations and the electrification history of the mill (if applicable).

The survey was approved by the ERES Converge Ethical Review Board in Lusaka on February 27, 2025. The registered study title is "Economic and Social Impact of Rural Electrification in Zambia" and the reference number is 2025-Feb-027.

Identification of respondents. The survey was implemented via telephone. To identify key informants and millers in the sampled villages and obtain their phone numbers, we proceeded as follows.

First, we spatially joined the set of sampled villages with the universe of geo-identified schools listed in the Education Information Management System of the Ministry of General Education. For the schools in the sampled villages, we obtained the phone numbers of the head teachers. Additionally, the Rural Electrification Authority provided a letter of support for the interview process.

Second, we called the identified head teachers and verified their location. If we were able to confirm their location, we proceeded with the interview. If not, we asked them to refer us to the head teacher in the correct village. If they were not able to do so, we called other schools in the surroundings until we found a suitable contact person in the village of interest. We also made sure key informants were not new to the villages of interest, but had spent a significant number of years in the village. In our sample, the minimum number of years respondents had spent in the village was 4 and 56% of respondents had spent 9 or more years in their villages.

Third, to identify millers for the miller interviews, we elicited a list of mills, including their owners and contact numbers, from the key informants. We also asked key informants which mill was the oldest and attempted to call this mill first. If we were not successful, we attempted other millers until we found a respondent.

Fieldwork. The survey was implemented in two phases. The first phase was conducted between May 8 and May 19, 2025, and covered 186 villages. All of these had a mill at baseline (152 in close proximity) and 47% had been electrified by 2020. In each of these village, we conducted both key informant and miller interviews.

The second phase covered 475 villages and consisted of key informant interviews without open-ended questions only. It was conducted between September 20 and December 11, 2025. None of the included villages had a mill at baseline and about half (55%) had been connected to the grid by 2020.

Response rate. The survey response rate was 84.7%, varying between 86.0% in the first phase and 84.2% in the second phase. Non-response is uncorrelated with village electrification in 2020.

A.III.3 Baseline heterogeneity

A.III.3.1 Population and Housing Census 2010

We use the full count sample of the 2010 population and housing census (PHC), as provided by Zamstats, for two purposes. First, we build a two-period panel where the second period data comes from the pooled LFS 2018-2022. To this end, we clean and harmonize all census variables that are also available in the LFS. Second, we leverage census data to provide a detailed description of baseline differences between villages with and without grain mills as well as villages that are and are not electrified during our study period. For this exercise, we use both data from the population census (household and individual characteristics) and data from the housing census (information on the existence and use of buildings).

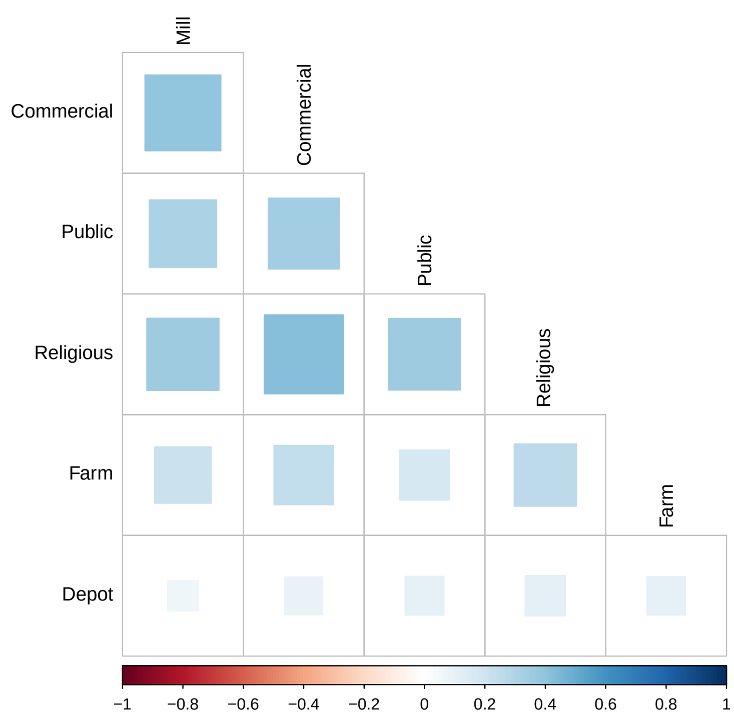
To link census data to RGCs, we proceed as described above for the LFS.

Third, the full Population and Housing Census 2010 also includes unique information on each building and structure in Zambia, totalling 2.1m structures in 2010. We use this data source to provide an alternative measure of productive capacity pre-existing at baseline that may be relevant for electrification in a given location. In particular, we focus on the subset of buildings determined by two key variables: occupancy of structure and structure type. We focus on non-residential buildings that are part of commercial operations as an alternative measure of productive capacity and high expected benefits of electrification in a given location: we obtain non-residential buildings from the 'occupancy' variable (which includes residential, non-residential and vacant occupancy), and (part of) commercial buildings from the 'type' variable (which includes, for example, flat, house, commercial, makeshift or mobile types).

A.III.3.2 Points of Interest

Alongside the PHC 2010, census enumerators collected a geo-identified dataset of the universe of points of interest. Among other things, points of interest include governmental institutions, such as administrative offices, police stations, post offices, and courts, rural institutions such as large-scale commercial farms or large-scale grain depots, religious and cultural institutions such as places of workship, as well as private businesses, such as shops, hospitality, and – importantly – mills. Figure A13 shows how the different points of interest relevant in the rural parts of Zambia are correlated with each other.

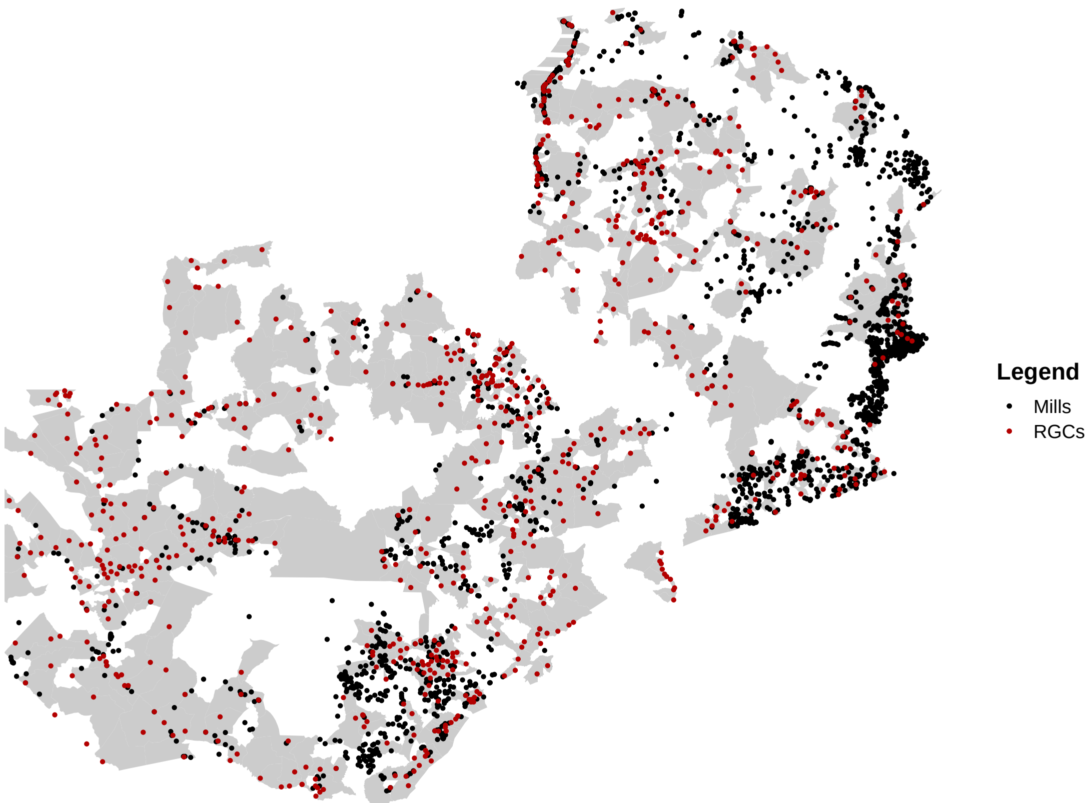
Figure A13: Correlation matrix of points of interest in inconsequential RGCs



This figure displays a matrix that visualises the pairwise correlation between various points of interest observed in inconsequential RGCs at baseline in 2010.

We link points of interest to RGCs based on their distance to RGC centroids. In the main specifications, we work with a 2km radius around RGC centroids, as derived from the spatial signatures of RGCs on nighttime luminosity, built-up area and vegetation. In robustness checks, we also consider alternative radii. Figure A14 illustrates the spatial distribution of (inconsequential) RGCs and mills across Zambia.

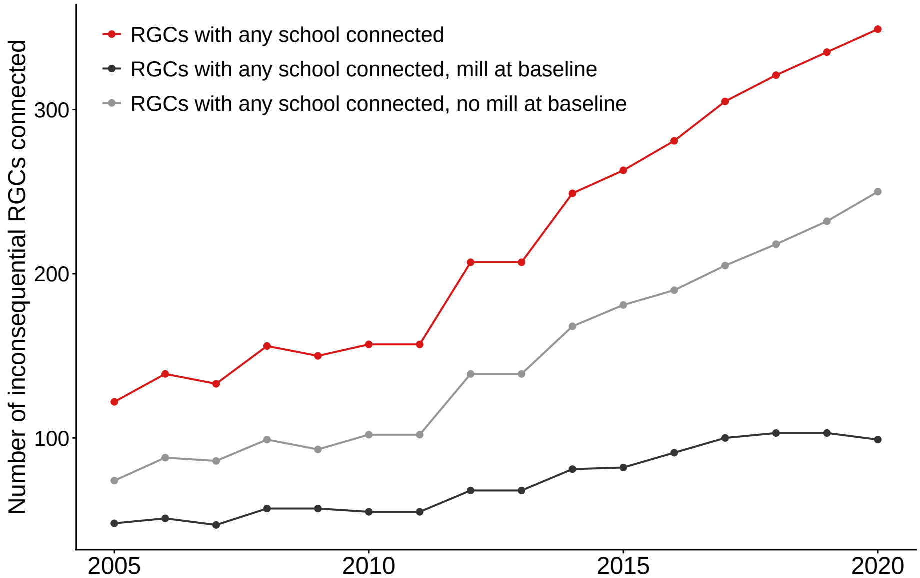
Figure A14: Spatial overlap between mills and inconsequential RGCs



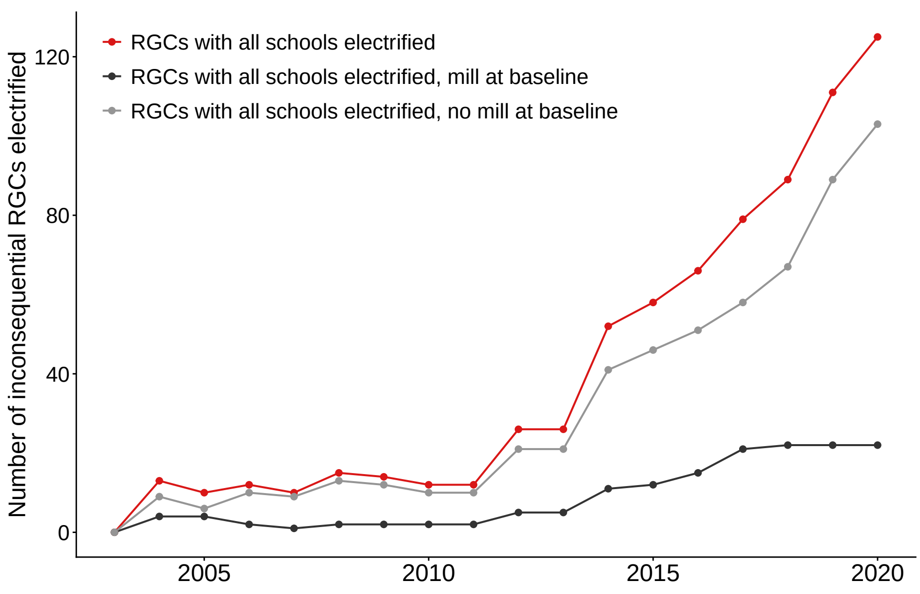
This map displays in black all mills that existed in Zambia at baseline in 2010. It also plots in red all inconsequential RGCs, including their respective wards in gray.

A.IV Descriptive statistics

Figure A15: Grid connection of inconsequential RGCs over time



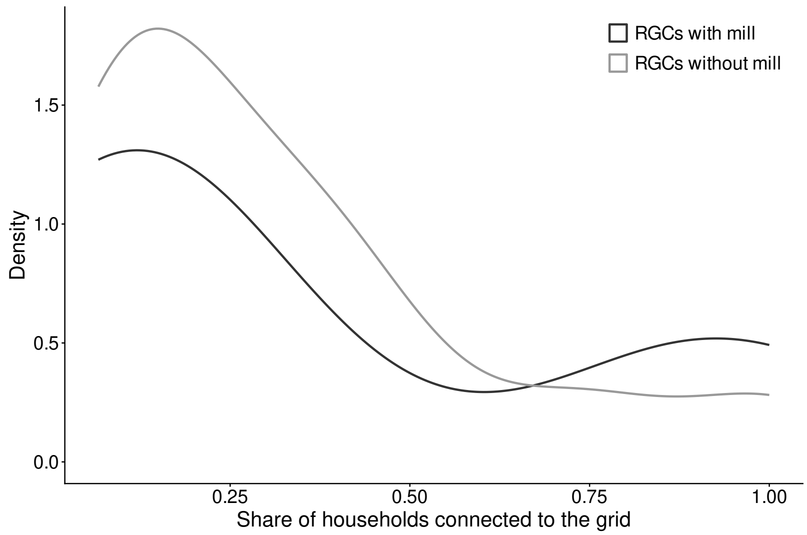
(a) Any school connected



(b) All schools connected

This figure illustrates the increase in grid connected schools across RGCs over time. Panel (a) measures RGC connection based on any school within 2km of the RGC reporting a grid connection, whereas Panel (b) only counts RGCs as connected when all schools within 2km of it report a grid connection.

Figure A16: Electricity adoption in grid-connected RGCs with and without a mill



This figure displays the share of households in LCMS 2015 survey clusters near RGCs in which at least one household reports being connected to the grid. The sample includes all clusters that lie within 2km of an RGC. Pre-existing mill indicates that the RGC had a mill at baseline.

Table A4: Summary statistics across Rural Growth Centres (RGCs) by mill status

	RGCs with mill		RGCs without mill			
	Mean	Std. Dev.	Mean	Std. Dev.	Difference	p-value
Panel A: Rural Electrification Master Plan (2009)						
Number of households, #	422	446	347	336	76	0.035**
Expected peak electricity load, kW	483.99	490.27	397.40	369.60	86.59	0.028**
RGC priority, rank	556	300	611	303	-55	0.032**
Panel B: Education and Health Monitoring Information Systems						
Number of schools, # (EMIS, 2008)	2.17	2.21	1.35	1.57	0.82	0.000***
Number of health posts, # (HMIS, 2012)	0.61	0.59	0.49	0.57	0.12	0.014**
Panel C: Points of Interest (2010)						
Has public facility, 0/1	0.56	0.50	0.21	0.41	0.35	0.000***
Has commercial facility, 0/1	0.75	0.43	0.31	0.46	0.44	0.000***
Has religious facility, 0/1	0.95	0.22	0.57	0.50	0.38	0.000***
Has depot facility, 0/1	0.05	0.22	0.02	0.15	0.03	0.115
Distance to nearest road, km	2.51	5.11	3.70	6.00	-1.19	0.009***
Panel D: Remotely-sensed data						
Mean nightlights, nW/cm2/sr (VIIRS, 2012)	0.28	0.87	0.16	0.63	0.12	0.090*
Max. nightlights, nW/cm2/sr (VIIRS, 2012)	1.58	3.76	1.17	3.77	0.41	0.197
Mean built-up area, m2/pixel (GHSL, 2005)	139.76	208.49	57.07	145.15	82.69	0.000***
Mean vegetation, index (NDVI, 2010)	0.57	0.07	0.57	0.08	-0.00	0.876
Panel E: Population and Housing Census (2010)						
Ward population, #	7150	3785	7178	4056	-28	0.935
SEA population, #	3494	4438	2339	3716	1155	0.004***
SEA density, #/km2	663.50	1400.34	309.75	1487.02	353.75	0.008***
Ward pop. growth, % (2000-2010)	35.26	71.52	36.71	90.08	-1.45	0.840
Wealth, index	0.03	0.01	0.03	0.01	0.00	0.812
Lighting: electricity, %	0.07	0.11	0.08	0.18	-0.00	0.904
Lighting: solar, %	0.03	0.03	0.04	0.05	-0.01	0.283
Cooking: electricity, %	0.04	0.07	0.06	0.16	-0.02	0.383
Cooking: biomass, %	0.95	0.07	0.94	0.16	0.01	0.498
Sector: agriculture, %	0.63	0.24	0.61	0.32	0.03	0.603
Sector: manufacturing, %	0.23	0.16	0.28	0.25	-0.05	0.215
Sector: services, %	0.14	0.14	0.12	0.15	0.02	0.489
Employed, %	0.58	0.12	0.57	0.17	0.01	0.711
Sector: agriculture, % (female)	0.62	0.25	0.62	0.33	0.00	0.984
Sector: manufacturing, % (female)	0.22	0.16	0.25	0.25	-0.03	0.491
Sector: services, % (female)	0.16	0.17	0.13	0.20	0.02	0.486
Employed, % (female)	0.52	0.16	0.53	0.21	-0.01	0.733
In school, %	0.32	0.07	0.33	0.07	-0.01	0.547
Literate, %	0.66	0.11	0.64	0.12	0.02	0.456
Years of schooling, #	6.01	0.79	5.95	1.02	0.06	0.710
In school, % (female)	0.30	0.07	0.30	0.07	-0.01	0.683
Literate, % (female)	0.62	0.13	0.60	0.14	0.02	0.456
Years of schooling, # (female)	5.58	0.85	5.55	1.05	0.03	0.868
Number of Rural Growth Centres, #	186		556			

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All Rural Electrification Master Plan variables sourced from Rural Electrification Authority (2009), including estimated number of households in Rural Growth Centres (RGCs) at baseline, expected peak electricity load as a function of, among others, number of households and resulting ranking of RGC according to the master plan. Education and Health Monitoring Information System data used to construct treatment variables sourced from Ministry of General Education and Ministry of Health, respectively. Points of interest data sourced from 2010 Grid3 Operational Settlements, collected as auxiliary data for the Population and Housing Census 2010. Remotely-sensed data sourced from publicly available sources for VIIRS, GHSL and NDVI. Population and Housing Census data sourced from ZAMSTATS 2010 PHC microdata, using 100% samples of housing, household and individual schedules.

A.V Additional results

A.V.1 Electricity adoption

In our endline survey, we asked key informants to count the number of various pieces of electric equipment in their villages. This includes the number of poles, the number of transformers, the share of households that have a grid connection, and, in the subset of villages that had a mill at baseline in 2010, a qualitative measure of reliability.

Table A5 reports the cross-sectional comparison of electrified and unelectrified villages in 2025, separated by their baseline mill status. Although the instructions for enumeration highlighted the requirement to only count low-voltage poles required for the local low-voltage network (see sketch in Figure A3, we cannot rule out that key informants provided a number of poles that mixed both medium-voltage and low-voltage poles, which look similar to the untrained observer. This would naturally explain why electrified villages without mills report a large number of poles, despite many not seeing any connections beyond the school in main analysis. However, the passing of time from the end of our sample in 2020 until the survey in 2025 may alternatively explain the slow build-out of some form of low-voltage network everywhere.

One important detail is that the p-value on the interaction effect on the number of transformers in column 3 is 0.1005, effectively implying that RGCs with mills have 0.63 additional transformers in the village at endline.

Table A5: Electricity adoption on RGC connection by mill status (2025)

	Elec. index	No. of poles	No. of trafos	Share of hhs	Reliability
	(1)	(2)	(3)	(4)	(5)
RGC connected	0.75***	32.05***	1.50***	0.14***	
	(0.14)	(7.31)	(0.21)	(0.03)	
RGC connected * mill	0.61**	27.59*	0.63	0.17***	0.40
	(0.29)	(15.52)	(0.38)	(0.05)	(0.49)
Const. FE	✓	✓	✓	✓	✓
Obs.	468	463	468	468	185
No. of years	1	1	1	1	1
No. of const-years	124	123	124	124	88
R^2	0.49	0.42	0.57	0.54	0.53

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan that were surveyed in 2025 endline survey. RGC connection derived from closest primary school in endline survey reporting a grid connection. Mill in RGC denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. No. of poles denotes the number of wooden or steel poles reported to be situated in RGC in endline survey. No. of trafos denotes the number of transformers required to step down medium- to low-voltages for end-user consumption in RGC in endline survey. Share of hhs denotes the share of households in RGC reported to have a grid connection in endline survey. Elec. index denotes an index of electrification variables, combining the number of poles, transformers and connected households in RGC. Reliability denotes the numeric score endline survey respondents awarded to the reliability of grid electricity in the endline survey, ranging from 1 (very poor) to 5 (very good). Reliability information is only available in RGC connected locations, such that the column reports the level difference between electrified mill and electrified no-mill RGCs. Const. FE denotes constituency fixed effects.

Table A6: Adoption on RGC connection by mill status

	Lighting			Cooking		
	Electric	Solar	Other	Electric	Wood	Other
RGC connected	−0.07	0.01	0.06	0.01	−0.02	0.00
	(0.05)	(0.05)	(0.06)	(0.03)	(0.03)	(0.00)
RGC connected * mill	0.24**	−0.19***	−0.06	0.04	−0.04	0.01
	(0.12)	(0.06)	(0.12)	(0.05)	(0.05)	(0.01)
RGC FE	✓	✓	✓	✓	✓	✓
Prov-year FE	✓	✓	✓	✓	✓	✓
R^2	0.82	0.80	0.86	0.84	0.84	0.54
No. of years	2	2	2	2	2	2
No. of RGCs	123	123	123	123	123	123
No. of prov-years	20	20	20	20	20	20
Observations	246	246	246	246	246	246

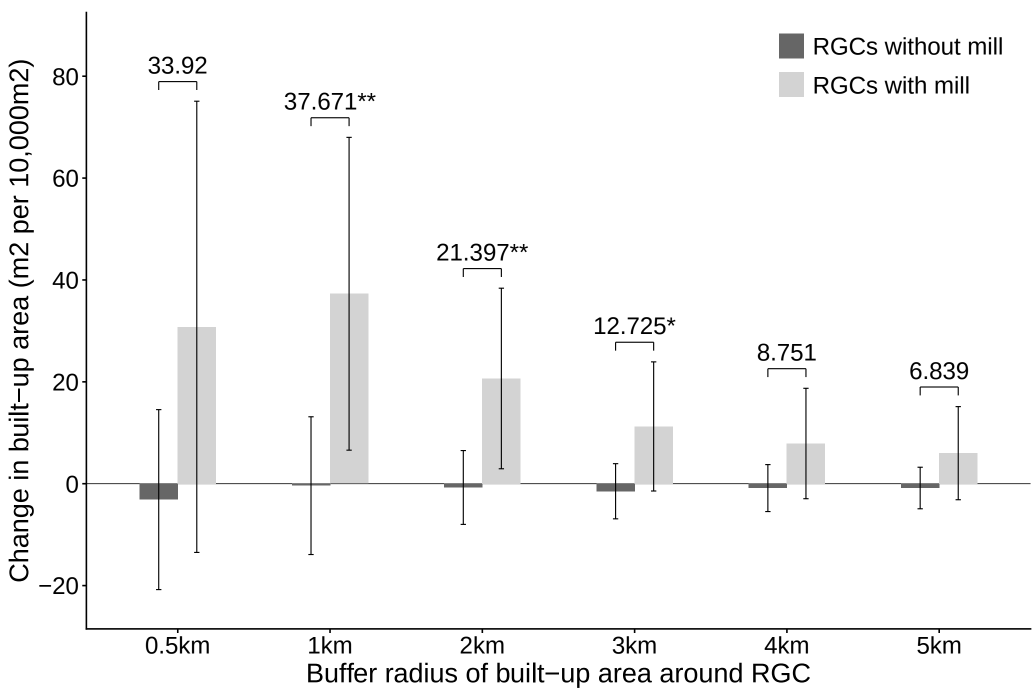
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Prov-year FE denotes province-year fixed effects. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) that intersect with any sample enumeration area (SEA) sampled in both PHC 2010 at baseline and any LFS round between 2018-2022 at endline. RGC-level actual electrification derived from schools existing since 2008. Mill in RGC denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Lighting: Electric denotes share of households in sample enumeration area (SEA) that report using grid-electric lighting. Lighting: Solar denotes share that report using non-grid electric lighting. Cooking: Electric and Cook: Wood denote share of households that report cooking using grid-electricity or wood/charcoal, respectively.

A.V.2 Buildings and greenery

To test for the robustness of our main results employing a 2km buffer around RGCs to different spatial signatures of RGCs, we report results for built-up area and vegetation across varying spatial extents below.

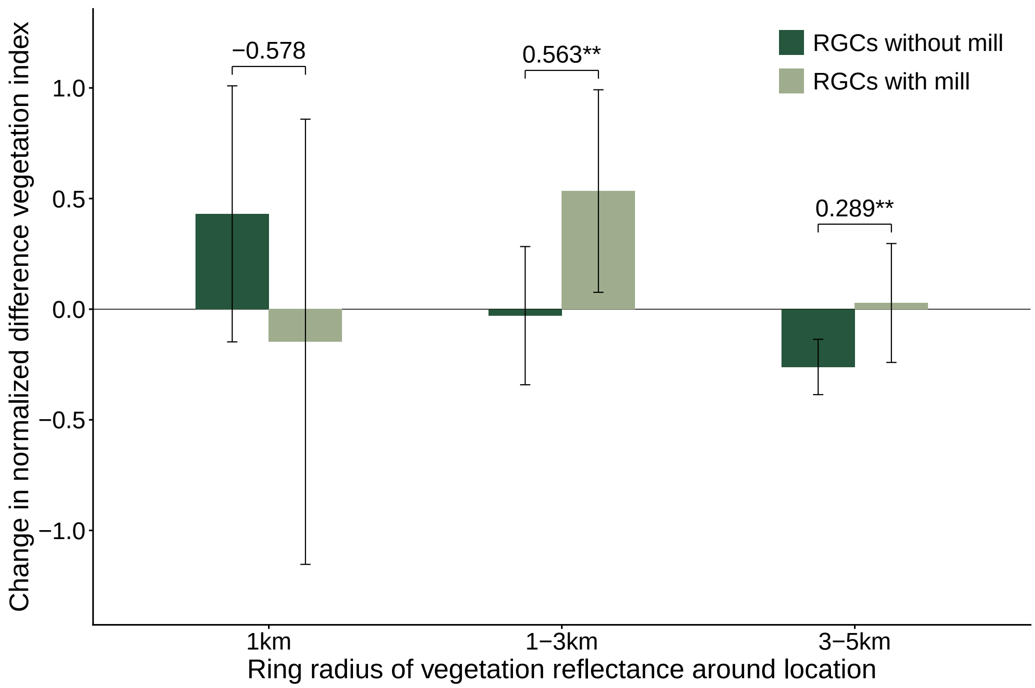
Figure A17 highlights how RGCs with mills at baseline experience significantly more built-up area upon electrification up until 2km radius from the RGC centre, although the difference to RGCs without mills remains statistically significant up until 3km distance. The comparison with RGCs without mills is stark: the effect of RGC connection on the share of a location's area covered in built structures is positive and statistically significantly different from zero within one, two and three kilometres radius around RGCs with a mill, but insignificant throughout around RGCs without one.

Figure A17: Effect of electrification on built-up area: spatial signature



A complementary finding is provided in Figure A18, namely that mean vegetation reflectance in RGCs with a mill is weakly negative within a one kilometre radius, but positive and significant in the outer distance 'halo' rings around the settlement, for

Figure A18: Effect of electrification on greenery: spatial signature



example from one to three and three to five kilometres. One interpretation of these results is that, first, agricultural activity in centre of the location is replaced with built-up structures, whereas the outskirts of electrified locations with mills see more intense, or higher-productivity agricultural activity.[33] Since we suspect the mill to adopt electricity, and given the cost advantage of mills powered by electricity, increases in agricultural productivity appear plausible.

[33] An alternative interpretation of the latter result would be that the cutting of trees for use as sources of energy is reduced, cf. Table 2 which highlights switches out of wood as source of cooking.

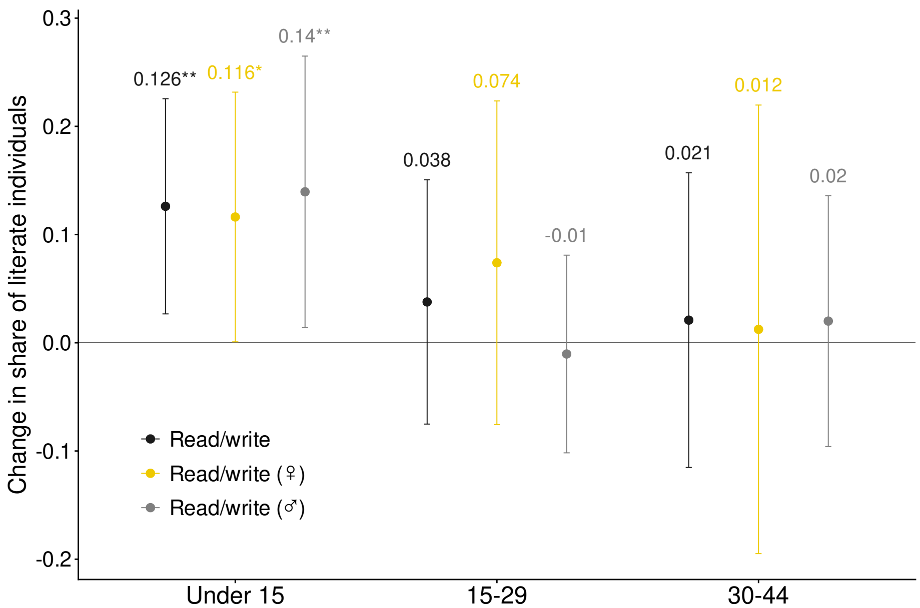
A.V.3 Occupation and education

Table A7: Sectoral employment and education on RGC connection by mill status

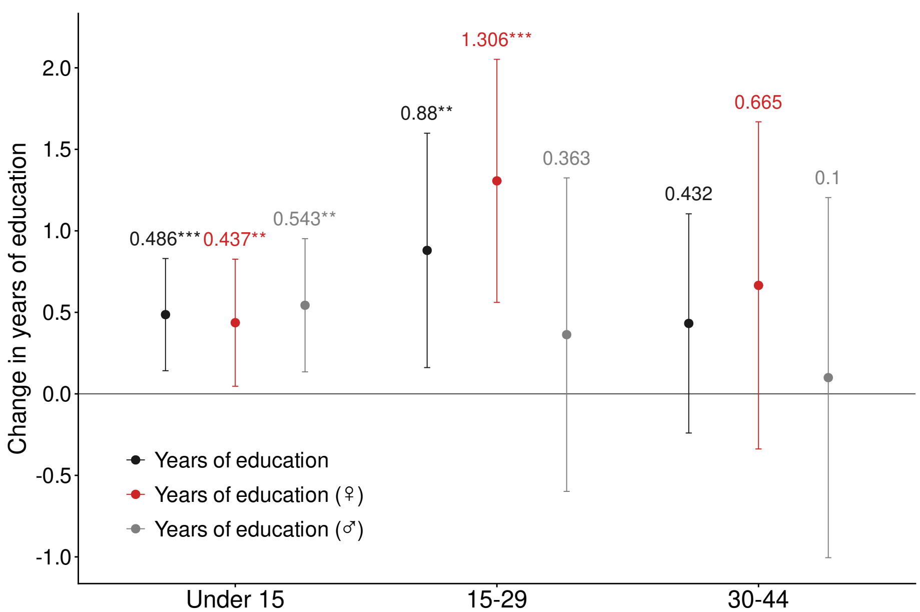
	Overall			Female (♀)			Male (♂)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Occupation									
	Agr.	Man.	Ser.	Agr.	Man.	Ser.	Agr.	Man.	Ser.
RGC connected	0.00	−0.05	0.05	0.01	−0.06	0.05	−0.03	−0.02	0.04
	(0.14)	(0.11)	(0.07)	(0.16)	(0.10)	(0.09)	(0.13)	(0.12)	(0.07)
RGC connected * mill	-0.31^{*}	0.24	0.07	-0.38^{*}	0.26	0.12	−0.25	0.20	0.04
	(0.19)	(0.15)	(0.08)	(0.20)	(0.17)	(0.10)	(0.18)	(0.16)	(0.09)
R^2	0.71	0.63	0.79	0.69	0.64	0.82	0.71	0.62	0.71
Observations	245	245	245	244	244	244	245	245	245
Panel B: Education									
	In edu.	Lit.	Years	In edu.	Lit.	Years	In edu.	Lit.	Years
RGC connected	0.01	0.02	−0.21	−0.01	0.01	−0.18	0.02	0.03	−0.25
	(0.03)	(0.03)	(0.26)	(0.04)	(0.04)	(0.24)	(0.04)	(0.03)	(0.31)
RGC connected * mill	0.05	0.13^{***}	0.33	0.05	0.15^{***}	0.41	0.07	0.11^{***}	0.18
	(0.04)	(0.03)	(0.41)	(0.04)	(0.04)	(0.31)	(0.05)	(0.03)	(0.57)
R^2	0.75	0.86	0.85	0.77	0.86	0.87	0.68	0.82	0.80
Observations	246	246	246	246	246	246	246	246	246
RGC FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Prov-year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
No. of years	2	2	2	2	2	2	2	2	2
No. of RGCs	123	123	123	123	123	123	123	123	123
No. of prov-years	20	20	20	20	20	20	20	20	20

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Prov-year FE denotes province-year fixed effects. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) that intersect with any sample enumeration area (SEA) sampled in both PHC 2010 at baseline and any LFS round between 2018-2022 at endline. RGC-level actual electrification derived from schools existing since 2008. Mill in RGC denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Sectoral Employment: Agr. denotes share of individuals in sample enumeration area (SEA) that report agriculture as their primary occupation. Man. and Ser. denote the equivalent occupation shares for manufacturing and services sectors, respectively. Education: In edu. denotes share of individuals in sample enumeration area (SEA) that report being in school. Lit. denotes share that are literate. Years denotes the years of education reported by individuals.

Figure A19: Effect of RGC connection on educational attainment by age and sex



(a) Literacy in connected mill vs no-mill villages



(b) Years of education in connected mill vs no-mill villages

*This figure displays the interaction effect of RGC connection in RGCs with mill present at baseline on the share of literate individuals in Panel (a) and years of educational attainment in Panel (b). Each coefficient and confidence interval represents the interaction effect of the RGC connected * mill coefficient in separate regressions, restricting by age group and sex. Coefficients in black represent changes in outcomes for all individuals of a given age group, whereas coloured (gray) ones represent changes in outcomes for female (male) individuals of a given age group.*

A.V.4 Water sources

Households in both the 2010 Population and Housing Census and the 2018-2022 Labour Force Surveys report their source of water in the household schedules. We group answers into three categories: tap/borehole, which represents water sources that require some form of pump to function, and could be run on electricity; well, which includes covered and uncovered wells; and other, which includes surface water sources as well as water delivery.

Table A8: Water source on RGC connection by mill status

	Water source		
	Tap	Well	Other
RGC connected	-0.09^{*}	0.09	0.00
	(0.05)	(0.07)	(0.06)
RGC connected * mill	0.18^{**}	-0.14^{*}	-0.04
	(0.08)	(0.08)	(0.07)
RGC FE	✓	✓	✓
Prov-year FE	✓	✓	✓
R^2	0.83	0.82	0.75
No. of years	2	2	2
No. of RGCs	226	226	226
No. of prov-years	20	20	20
Observations	452	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Prov-year FE denotes province-year fixed effects. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) that intersect with any sample enumeration area (SEA) sampled in both PHC 2010 at baseline and any LFS round between 2018-2022 at endline. RGC-level actual electrification derived from schools existing since 2008. Mill in RGC denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Water Source: Tap denotes share of households in sample enumeration area (SEA) that report using tap or borehole as source of water. Well denotes either protected or unprotected wells as source of water, while Other denotes surface water bodies, delivery or other sources of water.

A.V.5 Robustness

Table A9: Nightlights on RGC connection by baseline heterogeneity

	Mean nightlights		Max. nightlights	
	(1)	(2)	(3)	(4)
RGC connected	0.01	−0.03	0.05	−0.09
	(0.01)	(0.02)	(0.10)	(0.18)
RGC connected * mill	0.08**	0.08*	0.64**	0.64**
	(0.04)	(0.04)	(0.27)	(0.28)
RGC connected * SEA density		0.03		0.31
		(0.03)		(0.21)
RGC connected * SEA population		0.04		0.33*
		(0.02)		(0.17)
RGC connected * SEA wealth index		0.01		−0.11
		(0.02)		(0.18)
RGC connected * SEA elec. lighting		0.05**		0.23*
		(0.02)		(0.14)
RGC connected * road distance		−0.04*		−0.15
		(0.02)		(0.16)
RGC connected * public		0.09**		0.56**
		(0.04)		(0.25)
RGC connected * commercial		0.00		0.01
		(0.05)		(0.26)
RGC connected * religious		−0.06*		−0.55**
		(0.03)		(0.22)
RGC connected * farm		0.01		0.13
		(0.03)		(0.18)
RGC connected * depot		0.17		0.51
		(0.18)		(0.65)
RGC FE	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓
R^2	0.96	0.97	0.88	0.87
No. of years	9	9	9	9
No. of RGCs	742	634	742	634
No. of const-years	1215	1044	1215	1044
Observations	6678	5706	6678	5706

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC connected denotes percent of schools in RGC reporting grid electricity connection in a given year, derived from balanced sample of schools existing since 2008. Mill, public, commercial, religious, farm and depot denote indicator values that take a value of one if at least one mill, public, commercial, religious, farm or depot point of interest, respectively, is situated within 2km of RGC. Road distance denotes an indicator value that takes a value of one if the distance between the RGC and its closest road is greater than the sample median across RGCs. SEA density, SEA population, SEA wealth index and SEA electric lighting denote indicator values that take a value of one if the sample enumeration area (SEA) number of people per square metre, number of people, household wealth index and share of households using electric lighting, respectively, is greater or equal to the sample median across SEAs that intersect with any non-urban, non-solar home system RGCs. All baseline heterogeneity variables refer to values in 2010. Mean and max. nightlights represent the mean and maximum, respectively, zonal statistics of annual nighttime luminosity values derived from the VIIRS VNL sensor for each RGC 2km buffer polygon. Const-year FE denotes constituency-year fixed effects.

Table A10: Nightlights on RGC connection by mill status: robustness across treatment measures

	Mean nightlights		Max. nightlights	
	(1)	(2)	(3)	(4)
Panel A: Electrification dummy (any school electrified)				
RGC connected (any school)	0.01	0.00	0.10	0.01
	(0.01)	(0.01)	(0.07)	(0.08)
RGC connected (any school) * mill		0.04^{*}		0.36^{**}
		(0.02)		(0.15)
Panel B: Electrification percentage				
RGC connected (% of schools)	0.03^{**}	0.01	0.21^{**}	0.05
	(0.02)	(0.01)	(0.10)	(0.10)
RGC connected (% of schools) * mill		0.08^{**}		0.64^{**}
		(0.04)		(0.27)
Panel C: Electrification dummy (all schools electrified)				
RGC connected (all schools)	0.01	−0.00	0.13	0.00
	(0.02)	(0.02)	(0.12)	(0.13)
RGC connected (all schools) * mill		0.07		0.62^{*}
		(0.04)		(0.33)
RGC FE	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓
R^2	0.96	0.96	0.88	0.88
No. of years	9	9	9	9
No. of RGCs	742	742	742	742
No. of const-years	1215	1215	1215	1215
Obs.	6678	6678	6678	6678

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC-level connection derived from schools existing since 2008. Mill denotes an indicator that takes the value of one if at least one mill is situated within 2km of the RGC. Panel A: RGC connected denotes an indicator variable that takes value one if any school within 2km of RGC reports being electrified; Panel B: RGC connected denotes the share of schools within 2km of RGC that report being electrified; Panel C: RGC connected denotes an indicator variable that takes value one if all schools within 2km of RGC report being electrified.

Table A11: Nightlights on RGC connection by mill status: robustness to number of schools in RGC

	Full sample		Single school		Multiple schools	
	Mean	Max	Mean	Max	Mean	Max
RGC connected	0.01	0.05	-0.02	-0.08	0.04^{*}	0.26^{*}
	(0.01)	(0.10)	(0.02)	(0.15)	(0.02)	(0.15)
RGC connected * mill	0.08^{**}	0.64^{**}	0.03^{*}	0.26^{*}	0.09	0.70^{*}
	(0.04)	(0.27)	(0.02)	(0.14)	(0.06)	(0.37)
RGC FE	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓
R^2	0.96	0.88	0.85	0.74	0.97	0.90
No. of years	9	9	9	9	9	9
No. of RGCs	742	742	293	293	449	449
No. of const-years	1215	1215	846	846	1116	1116
Obs.	6678	6678	2637	2637	4041	4041

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC connected denotes percent of schools in RGC reporting grid electricity connection in a given year, derived from balanced sample of schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC.

Table A12: Nightlights on RGC connection by mill status: robustness to school-to-RGC match quality

	Mean nightlights					Max. nightlights				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
RGC connected	0.03**	0.01				0.21**	0.05			
	(0.02)	(0.01)				(0.10)	(0.10)			
RGC connected * mill		0.08**					0.64**			
		(0.04)					(0.27)			
RGC connected (1)			0.02**					0.02		
			(0.01)					(0.05)		
RGC connected (1) * mill			0.09*					0.43***		
			(0.05)					(0.15)		
RGC connected (2)				0.02**					0.01	
				(0.01)					(0.05)	
RGC connected (2) * mill				0.09*					0.43***	
				(0.05)					(0.15)	
RGC connected (3)					0.06					0.09
					(0.04)					(0.07)
RGC connected (3) * mill					0.05					0.30**
					(0.06)					(0.15)
RGC FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.96	0.96	0.97	0.97	0.96	0.88	0.88	0.87	0.87	0.86
No. of years	9	9	9	9	9	9	9	9	9	9
No. of RGCs	742	742	646	632	489	742	742	646	632	489
No. of const-years	1215	1215	1197	1188	1116	1215	1215	1197	1188	1116
Obs.	6678	6678	5814	5688	4401	6678	6678	5814	5688	4401

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. Baseline RGC-level connection measure derived from schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Alternative RGC-level connection measure (1) derived from only considering schools that that fall within 2km of RGC and have a good name match with the village name. Alternative RGC-level connection measure (2) derived from only considering schools that that fall within 1km of RGC. Alternative RGC-level connection measure (3) derived from only considering schools that that fall within 1km of RGC and have an exact name match with the village name. As a result, the sample of RGCs decreases as more and more restrictive definitions of RGC-to-school match are employed.

Table A13: Nightlights on RGC connection by mill status: robustness to rural sample restriction

	Mean nightlights				Max. nightlights			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RGC connected	0.01	0.01	0.00	0.00	0.05	0.05	0.01	0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.10)	(0.10)	(0.09)	(0.09)
RGC connected * mill	0.08**	0.08**	0.03**	0.03*	0.64**	0.64**	0.34*	0.37*
	(0.04)	(0.04)	(0.02)	(0.02)	(0.27)	(0.27)	(0.18)	(0.19)
RGC FE	✓	✓	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.96	0.96	0.87	0.93	0.88	0.88	0.65	0.79
No. of years	9	9	9	9	9	9	9	9
No. of RGCs	742	742	681	715	742	742	681	715
No. of const-years	1215	1215	1143	1197	1215	1215	1143	1197
Obs.	6678	6678	6129	6435	6678	6678	6129	6435

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. Model key: (1) and (5) include only RGCs with a 2010 population below 18,849.56 (baseline specification). (2) and (6) include only RGCs with a 2010 population below 20,000. (3) and (7) include only RGCs with a 2010 population density below 1,500 people per square kilometre according to World Pop. (4) and (8) include only RGCs that either only contain intersecting 2010 census enumeration areas marked as rural, or those RGCs with a 2010 population density of the intersecting 2010 census enumeration area below 100 people per square kilometer irrespective of being marked rural. RGC-level connection derived from schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC.

Table A14: Nightlights on RGC connection by mill status: robustness to consequential and off-grid RGCs

	Mean nightlights				Max. nightlights			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RGC connected	0.01	0.01	0.01	0.01	0.05	0.07	0.06	0.08
	(0.01)	(0.01)	(0.01)	(0.01)	(0.10)	(0.08)	(0.09)	(0.08)
RGC connected * mill	0.08**	0.07**	0.06**	0.08**	0.64**	0.37**	0.43**	0.46*
	(0.04)	(0.03)	(0.03)	(0.04)	(0.27)	(0.18)	(0.19)	(0.24)
RGC FE	✓	✓	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓	✓	✓
Inconsequential RGCs	✓	✓	✓	✓	✓	✓	✓	✓
Solar RGCs		✓	✓			✓	✓	
Consequential RGCs		✓		✓		✓		✓
R^2	0.96	0.96	0.96	0.96	0.88	0.84	0.84	0.87
No. of years	9	9	9	9	9	9	9	9
No. of RGCs	742	1134	965	911	742	1134	965	911
No. of const-years	1215	1251	1233	1233	1215	1251	1233	1233
Obs.	6678	10206	8685	8199	6678	10206	8685	8199

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Model key: (1) and (5) use sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan (REMP) (baseline specification). (2) and (6) include all RGCs, including both consequential and solar home system RGCs. (3) and (7) include inconsequential grid and off-grid RGCs in the REMP. (4) and (8) include consequential and inconsequential grid RGCs in the REMP. RGC-level connection derived from schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC.

Table A15: Nightlights on RGC connection: robustness to time-varying controls (roads, main roads, rainfall)

	Mean nightlights				Max. nightlights			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RGC connected	0.01	−0.02	−0.04	0.01	0.05	0.09	−0.24	0.04
	(0.01)	(0.04)	(0.04)	(0.01)	(0.10)	(0.29)	(0.19)	(0.10)
RGC connected * mill	0.08**	0.17	0.19*	0.08**	0.64**	1.05	1.29**	0.62**
	(0.04)	(0.11)	(0.10)	(0.04)	(0.27)	(0.86)	(0.52)	(0.27)
Road control		✓				✓		
Main road control			✓				✓	
Mean precipitation control				✓				✓
RGC FE	✓	✓	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.96	0.95	0.95	0.96	0.88	0.87	0.87	0.88
No. of years	9	3	3	9	9	3	3	9
No. of RGCs	742	742	742	742	742	742	742	742
No. of const-years	1215	405	405	1215	1215	405	405	1215
Obs.	6678	2226	2226	6678	6678	2226	2226	6678

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC-level connection derived from schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Road and Main road controls denote an indicator values that takes the value of one if at least one road or main road situated within 2km of RGC.

Table A16: Nightlights on RGC connection by mill status: robustness to alternative baseline facilities

	Mean nightlights							Max. nightlights						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
RGC connected	0.03**	0.01	0.01	0.03**	0.02**	−0.01	0.02	0.21**	0.05	0.07	0.22**	0.24**	−0.04	0.15
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.10)	(0.10)	(0.18)	(0.11)	(0.11)	(0.15)	(0.13)
RGC connected * mill		0.08**							0.64**					
		(0.04)							(0.27)					
RGC connected * commercial			0.05							0.41				
			(0.04)							(0.33)				
RGC connected * farm				−0.00							−0.03			
				(0.03)							(0.16)			
RGC connected * religious					0.01							−0.04		
					(0.02)							(0.17)		
RGC connected * public						0.14***							0.85**	
						(0.05)							(0.38)	
RGC connected * depot							0.45							2.81
							(0.39)							(2.86)
RGC FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.88	0.88	0.88	0.88	0.88	0.88	0.88
No. of years	9	9	9	9	9	9	9	9	9	9	9	9	9	9
No. of RGCs	742	742	742	742	742	742	742	742	742	742	742	742	742	742
No. of const-years	1215	1215	1215	1215	1215	1215	1215	1215	1215	1215	1215	1215	1215	1215
Obs.	6678	6678	6678	6678	6678	6678	6678	6678	6678	6678	6678	6678	6678	6678

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC-level connection derived from schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Commercial denotes an indicator value that takes the value of one if at least one commercial facility situated within 2km of RGC. Agriculture denotes an indicator value that takes the value of one if at least one agricultural facility situated within 2km of RGC. Religion denotes an indicator value that takes the value of one if at least one religious facility situated within 2km of RGC. Public denotes an indicator value that takes the value of one if at least one public facility situated within 2km of RGC. Depot denotes an indicator value that takes the value of one if at least one large-scale grain depot situated within 2km of RGC.

A.VI Mechanisms

A.VI.1 Robustness of cost externality

Table A17: Nightlights on RGC connection by mill count, public count and their interaction

	Mean nightlights			Max. nightlights		
	(1)	(2)	(3)	(4)	(5)	(6)
RGC connected	0.02	−0.02	−0.03	0.10	−0.01	−0.12
	(0.01)	(0.02)	(0.02)	(0.10)	(0.15)	(0.15)
RGC connected * mill count	0.05^{**}		0.02	0.32^{**}		0.31^{**}
	(0.02)		(0.02)	(0.15)		(0.14)
RGC connected * public count		0.10^{**}	0.11^{*}		0.40^{**}	0.53^{**}
		(0.04)	(0.06)		(0.20)	(0.25)
RGC connected * mill count * public count			−0.01			-0.12^{***}
			(0.01)			(0.04)
RGC FE	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓
R^2	0.96	0.96	0.96	0.88	0.88	0.88
No. of years	9	9	9	9	9	9
No. of RGCs	742	742	742	742	742	742
No. of const-years	1215	1215	1215	1215	1215	1215
Observations	6678	6678	6678	6678	6678	6678

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC connected denotes percent of schools in RGC reporting grid electricity connection in a given year, derived from balanced sample of schools existing since 2008. Mill (public) count denotes the number of mills (public facilities) situated within 2km of RGC. Mean and max. nightlights represent the mean and maximum, respectively, zonal statistics of annual nighttime luminosity values derived from the VIIRS VNL sensor for each RGC 2km buffer polygon. Const-year FE denotes constituency-year fixed effects.

A.VI.2 Robustness of demand externality

It is important to note that blossoming of nightlights could contribute to this finding, although it remains unclear why blossoming should affect the uncovered part of RGCs with mills differently than the uncovered part of RGCs without mills, given that both

Table A18: Nightlights in parts of RGC covered or uncovered by low-voltage network (including mills, at least two nodes) on RGC connection by mill status

	Entire RGC		Covered part		Uncovered part	
	Mean	Max	Mean	Max	Mean	Max
RGC connected	0.03**	0.13	0.12***	0.19***	0.03**	0.15
	(0.02)	(0.13)	(0.03)	(0.06)	(0.01)	(0.12)
RGC connected * mill	0.08*	0.66**	0.17	0.31	0.08*	0.62**
	(0.05)	(0.32)	(0.13)	(0.27)	(0.04)	(0.28)
RGC FE	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓
p-val $\delta = 0$	0.08	0.043	0.212	0.251	0.081	0.029
p-val $\beta + \delta = 0$	0.015	0.007	0.042	0.084	0.015	0.007
R^2	0.97	0.90	0.96	0.95	0.97	0.91
No. of years	9	9	9	9	9	9
No. of RGCs	498	498	498	498	498	498
No. of const-years	1143	1143	1143	1143	1143	1143
Observations	4482	4482	4482	4482	4482	4482

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC connected denotes percent of schools in RGC reporting grid electricity connection in a given year, derived from balanced sample of schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Covered part denotes the area within the 2km buffer around the RGC that is covered by a 25m buffer around the minimum-spanning tree. Uncovered part denotes the area within the 2km buffer around the RGC that is not covered by a 25m buffer around the minimum-spanning tree. Minimum-spanning trees represent a single, cost-minimising low-voltage network required to connect all schools, health facilities, public facilities and mills in a RGC to the grid. Only RGCs included for which minimum-spanning tree includes at least two nodes. p-val δ reports the p-value on the interaction term with mill, whereas p-val $\beta + \delta$ reports the same for the overall effect for RGCs with mills.

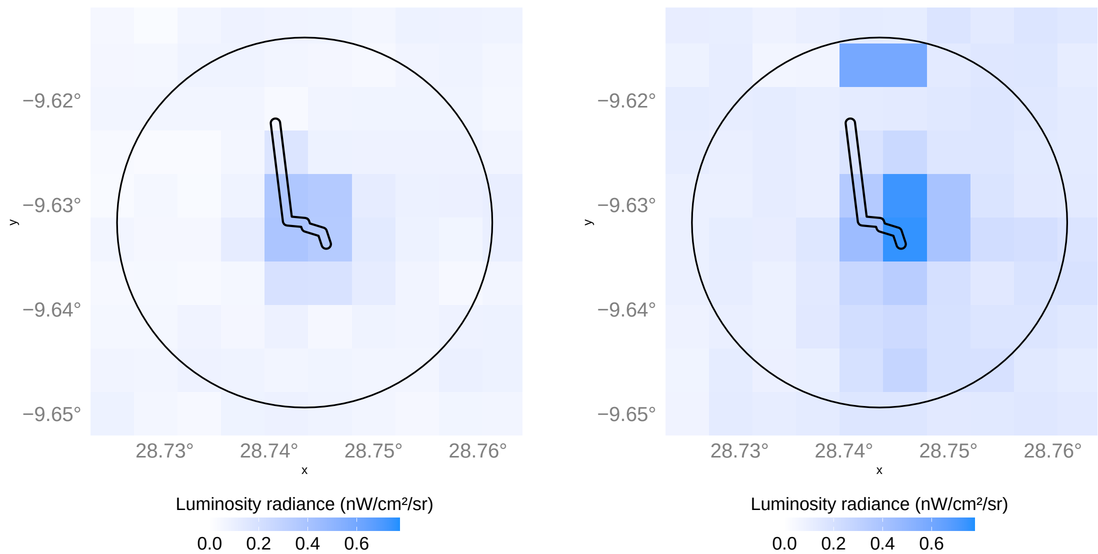
types of locations are significantly brighter in the parts covered by the predicted low-voltage network. To investigate robustness of our finding, we show in Figure A21 how the uncovered point estimate on the interaction effect, i.e., row 2 in columns 5-6 of Table 5, would change if all pixels touched by the predicted low-voltage network were masked entirely from the uncovered part. The figure highlights how this masked point estimate in the uncovered part remains statistically significant for mean nightlights (Panel (a)) and maximum nightlights (Panel (b)) even when masking pixels that lie within 250m from the predicted low-voltage network.

Table A19: Nightlights in parts of RGC covered or uncovered by low-voltage network on RGC connection by mill status

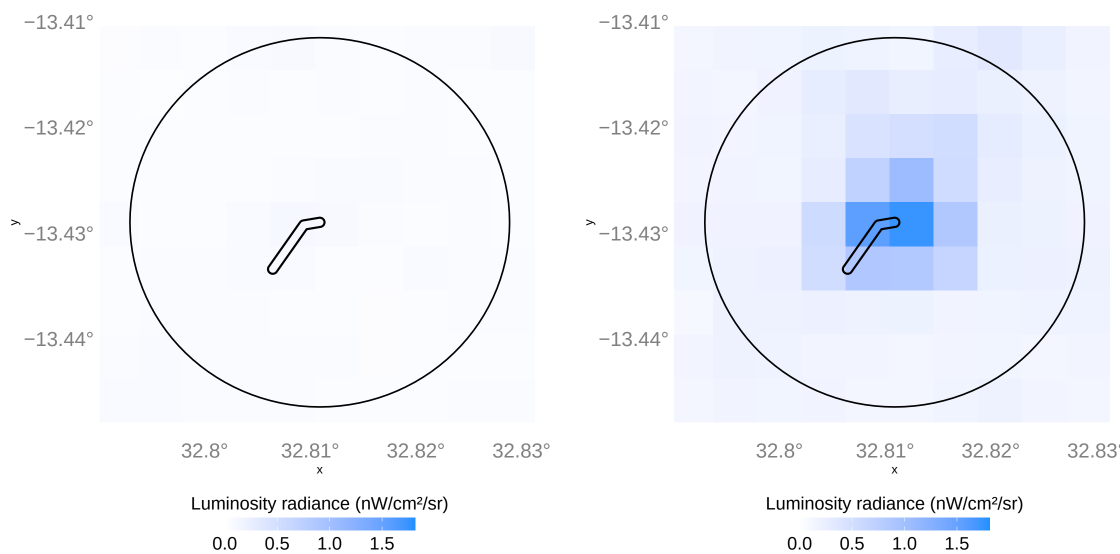
	Entire RGC		Covered part		Uncovered part	
	Mean	Max	Mean	Max	Mean	Max
RGC connected	0.01	0.05	0.07^{***}	0.12^{***}	0.01	0.10
	(0.01)	(0.10)	(0.02)	(0.04)	(0.01)	(0.09)
RGC connected * mill	0.08^{**}	0.64^{**}	0.21^{*}	0.36	0.08^{**}	0.58^{**}
	(0.04)	(0.27)	(0.12)	(0.23)	(0.04)	(0.25)
RGC FE	✓	✓	✓	✓	✓	✓
Const-year FE	✓	✓	✓	✓	✓	✓
p-val $\delta = 0$	0.041	0.021	0.081	0.128	0.043	0.021
p-val $\beta + \delta = 0$	0.019	0.008	0.024	0.059	0.020	0.008
R^2	0.96	0.88	0.95	0.94	0.96	0.89
No. of years	9	9	9	9	9	9
No. of RGCs	742	742	742	742	742	742
No. of const-years	1215	1215	1215	1215	1215	1215
Observations	6678	6678	6678	6678	6678	6678

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan. RGC connected denotes percent of schools in RGC reporting grid electricity connection in a given year, derived from balanced sample of schools existing since 2008. Mill denotes an indicator value that takes the value of one if at least one mill situated within 2km of RGC. Covered part denotes the area within the 2km buffer around the RGC that is covered by a 25m buffer around the minimum-spanning tree. Uncovered part denotes the area within the 2km buffer around the RGC that is not covered by a 25m buffer around the minimum-spanning tree. Minimum-spanning trees represent a single, cost-minimising low-voltage network required to connect all schools, health facilities and public facilities in a RGC to the grid.

Figure A20: Effect of electrification on nightlights in parts of RGC covered and not covered by local low-voltage network



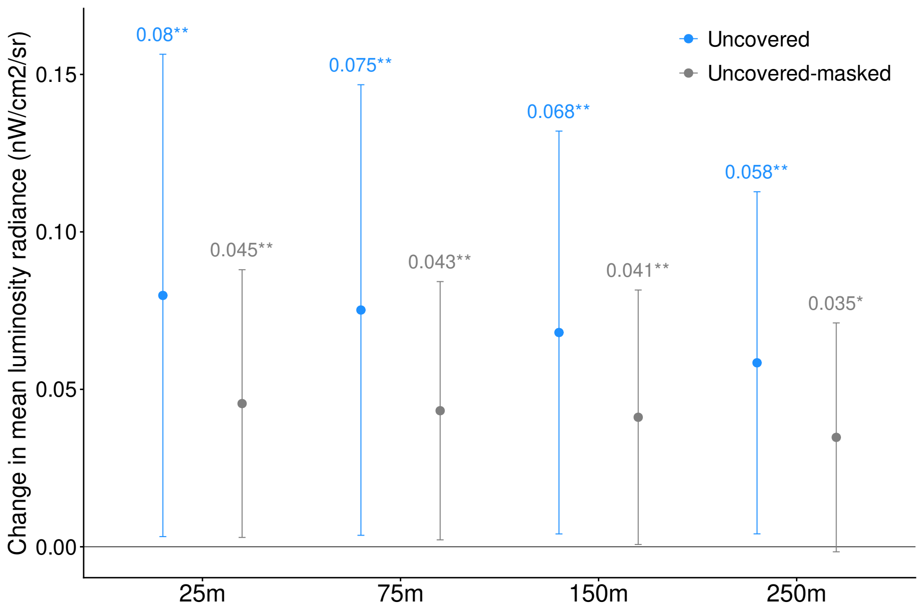
(a) RGC with mill



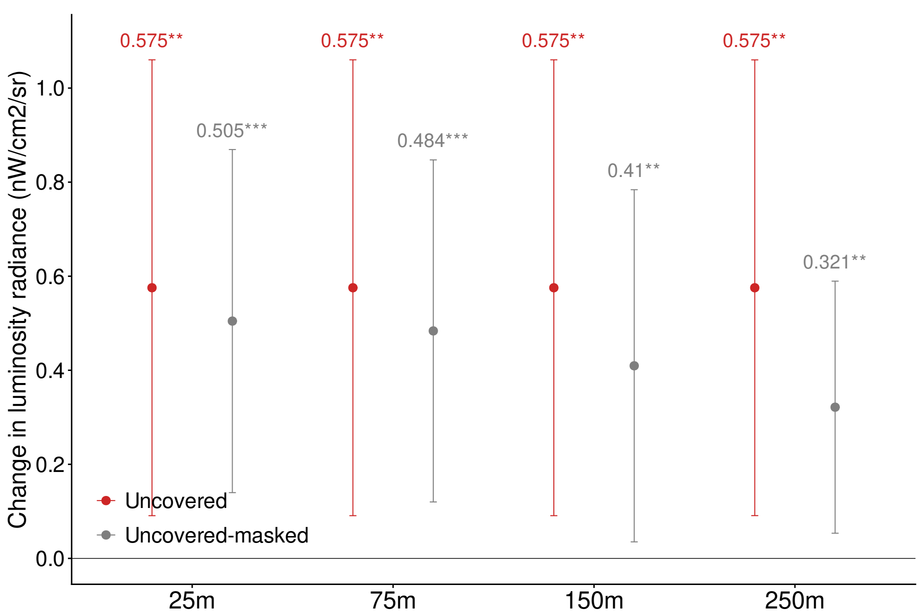
(b) RGC without mill

This figure illustrates the increase in nighttime luminosity in buffers of two kilometre radius around two grid-connected RGCs at the beginning and end of our samples. The figure also plots the minimum spanning tree of the low-voltage network in each RGC. This tree represents path a low-voltage network would have to take to connect all public facilities, schools, health facilities and mills, if present, in a cost-minimising fashion. We plot a 50m buffer around each RGC's minimum-spanning tree for illustration.

Figure A21: Effect of RGC connection on nightlights in uncovered part, by MST buffer



(a) Mean nightlights in uncovered part of connected mill RGC, masked or unmasked



(b) Max. nightlights in uncovered part of connected mill RGC, masked or unmasked

This figure displays the interaction effect of RGC connection in RGCs with mill present at baseline on the mean (blue) and max. (red) nightlights in the part of the RGC uncovered by the local low-voltage network. For robustness, the size of the buffer around the local low-voltage network, i.e., the definition of which part remains 'uncovered' is increased. To safeguard against bleeding of nightlights, we also present in gray the equivalent effect where any pixel touched by the local low-voltage network buffer is masked from the uncovered part of the RGC.

Table A20: Miller operations on mill connection (2025)

	Machines	Employees	Quantity	Price	Op. days	Non-op. weeks
	(1)	(2)	(3)	(4)	(5)	(6)
Mill connected	0.26**	0.37*	28.21	−0.12*	0.24	−0.42
	(0.12)	(0.19)	(22.97)	(0.06)	(0.23)	(1.08)
Miller controls	✓	✓	✓	✓	✓	✓
No. of years	1	1	1	1	1	1
R^2	0.06	0.03	0.01	0.04	0.13	0.13
Obs.	153	153	153	153	153	153

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. All standard errors clustered at electrification package-level. Sample of 160 mills, one mill per inconsequential, non-urban, non-solar home system Rural Growth Centres (RGCs) included in Rural Electrification Master Plan that had a pre-existing mill at baseline in 2010. Solar-powered mills (7) excluded from the sample. Mill connection derived from miller's response in 2025 miller survey. Miller controls include an indicator variable if the miller is female and the miller's tenure in years. Machines denotes the number of machines the miller reports to be operating. Employees denotes the number of employees the miller reports to employ. Quantity denotes the number of medas the miller reports to process per day. Price denotes the logarithm of one plus the price per meda the miller reports to charge for milling services. Op. days denotes the number of operating days per week the miller reports. Non-op. weeks denotes the recalled number of weeks in 2024 during which the miller was unable to operate.