Working paper

Does electrification cause industrial development?

Grid expansion and firm turnover in Indonesia

Dana Kassem

May 2018

When citing this paper, please use the title and the following reference number:
E-89341-IDN-1

DIRECTED BY

FUNDED BY

Does Electrification Cause Industrial Development? Grid Expansion and Firm Turnover in Indonesia

Job Market Paper

Click here for recent version

Dana Kassem*

May 25, 2018

Abstract

I ask whether electrification causes industrial development. I combine newly digitized data from the Indonesian state electricity company with rich manufacturing census data. To understand when and how electrification can cause industrial development, I shed light on an important economic mechanism - firm turnover. In particular, I study the effect of the extensive margin of electrification (grid expansion) on the extensive margin of industrial development (firm entry and exit). To deal with endogenous grid placement, I build a hypothetical electric transmission grid based on colonial incumbent infrastructure and geographic cost factors. I find that electrification causes industrial development, represented by an increase in the number of manufacturing firms, manufacturing workers, and manufacturing output. Electrification increases firm entry rates, but also exit rates. Higher turnover rates lead to higher average productivity and induce reallocation towards more productive firms in electrified areas. This is consistent with electrification lowering entry costs, increasing competition and forcing unproductive firms to exit more often. Without the possibility of entry or competitive effects of entry, the effects of electrification are likely to be smaller.
(*JEL* D24, L60, O13, O14, Q41)

* Department of Economics and STICERD, London School of Economics, d.kassem@lse.ac.uk. I thank Robin Burgess and Oriana Bandiera for helpful advice and continuous support. I also thank Gharad Bryan, Michel Azulai, Clare Balboni, Matteo Benetton, Jan De Loecker, Thomas Drechsel, Greg Fischer, Alessandro Gavazza, Maitreesh Ghatak, Hanwei Huang, Rachael Meager, Panos Mavrokonstantis, Marco Gonzalez-Navarro, Kieu-Trang Nguyen, Michael Peters, Alex Rothenberg, Mark Schankerman, Pasquale Schiraldi, Arthur Seibold, John Sutton, Junichi Yamasaki, and various seminar and conference participants. I am grateful for the support of Suroso Isnandar, Muhammad Ikbal Nur, Musa Partahi Marbun, Ahmed Yusuf Salile and various others at the Indonesian State Electricity Company Perusahaan Listrik Negara (PLN). Financial support from the IGC and STICERD is gratefully acknowledged.

1 Introduction

The idea that electrification causes industrial development dates back as far as Lenin[1]. Even today, many governments and aid agencies[2] invest in energy infrastructure projects, especially in developing countries. In 2017, the Indonesian government invested around $1.8 billion in electricity, 7% out of its total budget for infrastructure. The Kenyan government is currently investing $2.1 billion in the grid expansion to rural areas. The Kenyan policymakers expect this investment ”to enhance industrialization and emergence of [...] industries”. There is consensus among policymakers that access to electricity is an essential ingredient for industrial development, which is considered a fundamental driver of growth.

However, recent economic evidence shows that the benefits of electrification are not as large as previously thought (Lee, Miguel, and Wolfram (2016), Burlig and Preonas (2016)). In fact, electrification in various African countries has increased substantially over the last decades, but these countries have not witnessed industrial development. So I ask, does electrification cause industrial development? Or do these investments have little impact on the pace of industrial development?

To answer this question, I use a rapid, government-led grid expansion during a period of rapid industrialization in Indonesia. I travelled multiple times to Indonesia and put together a comprehensive data-set over a period of 11 years from 1990 to 2000 from various current and historical sources. I first map the expansion of the electric transmission grid over time and space in Java, the main island in Indonesia. I then map manufacturing activity in 25,000 administrative areas for more than 29,000 unique firm observations in Java, where 80% of Indonesian manufacturing firms are located. These data allow me to understand *when* and *how* electrification affects industrial development.

This is the first paper to examine the effect of the extensive margin of electrification (grid expansion) on the extensive margin of industrial development (firm entry and exit). The effect of the extensive margin of electrification, i.e. extending the electric grid to new

[1] Lenin (1920)”*Communism is Soviet power plus the electrification of the whole country”.* Lenin believed that electrification would transform Russia from a ”*small-peasant basis into a large-scale industrial basis*”

[2] The World Bank has committed to lending $6.3 billion to the Energy and Mining sector worldwide. From *The World Bank Annual Report 2017*, `http://www.worldbank.org/en/about/annual-report`.

locations, has been studied on employment (Dinkelman (2011)) and general development-level indices (Lipscomb, Mobarak, and Barham (2013)). Other papers have estimated the demand and cost of rural electrification for households in a controlled environment (Lee, Miguel, and Wolfram (2016)). The link between electrification and firms has been studied on the intensive margin and is mostly focused on the effect of shortages on firm outcomes (e.g. Allcott, Collard-Wexler, and O'Connell (2016)). Variation in shortages creates short-run firm responses by affecting the input price of electricity which in turn affects the firm's production decision on the intensive margin. The evidence on the intensive margin of electrification and industrial development is important, but the effect of the extensive margin of electrification on industrialization is potentially different, and of greater relevance to those interested in long run development. Changes on the extensive margin of electrification, meaning whether the firm can be connected to the electric grid or not, can create long-run firm responses by affecting the extensive margin of firm decisions, namely, entry and exit.

An economic mechanism through which electrification potentially affects industrial development is therefore firm turnover, driven by the entry and exit of firms. Electrifying a new location can influence firms' entry and exit decisions in that particular location. This affects the composition of firms in the market, and hence, average productivity. Whether or not electrification enhances or decreases manufacturing productivity is therefore a question that requires empirical verification.

Indonesia is an ideal setting to answer this research question. For historical reasons, the Indonesian power sector remained underdeveloped compared to countries with a similar GDP. In 1990, Java, the most developed and densely populated island in Indonesia, was only around 40% electrified. The island has since witnessed a massive and successful government-led effort to expand access to electricity up until the year 2000. During that period, transmission capacity in Java quadrupled and electrification ratios increased to more than 90%. At the same time, Indonesia experienced fast growth in the manufacturing sector. This allows me to match modern type firm-level micro data with sufficient recent variation in access to the grid to detailed data on the electrification infrastructure.

Establishing a causal link between electrification and industrial development is empirically challenging. In any emerging economy, infrastructure and industrialization occur simultaneously, and separating demand-side from supply-side factors is difficult. This poses an empirical challenge in identifying the effect of electrification on industrial outcomes. My empirical strategy tries to make progress on this issue by using an instrumental variable strategy inspired by the transportation infrastructure literature[3]. I exploit a supply-side natural experiment based on the need of the state electricity monopoly to have a single interconnected electricity grid in Java. I construct a hypothetical interconnected electric transmission grid that is a function of incumbent disconnected electrification infrastructure built by Dutch colonial electric utilities and geographic cost factors. The hypothetical grid abstracts from endogenous demand factors that could be driving the expansion of the grid and focuses on cost factors only. The use of the colonial infrastructure also means that the incumbent infrastructure is unlikely to be correlated with economic forces in 1990. Distance to the hypothetical grid is used to instrument for endogenous access to electricity, conditional on various controls, including other types of infrastructure.

The data-sets used in this paper come from various sources. I collected and digitized spatial data on the electrification infrastructure from the Indonesian state electricity monopoly Perusahaan Listrik Negara (PLN) in Jakarta. This includes data on the location, operation year, and capacity of power plants and transmission substations. To build a time-series, I use administrative documents from PLN. Gaps are then filled from World Bank loan reports from 1969 to 1992. I then construct measures of access to the grid based on the distance from the centroid of a desa to the nearest transmission substation. A desa is the lowest administrative division in Indonesia. To study firm turnover, I construct yearly maps of manufacturing activity in Java, which includes the number of firms, manufacturing output, number of manufacturing workers, and entry and exit rates in any desa in Java. The information on manufacturing activity at the desa level comes form the Indonesian annual manufacturing census 1990-2000. This is a census of Indonesian manufacturing firms with 20 or more employees. The firm-level data is also used to get information on firm output, inputs, exit and entry decisions, as well as to get estimates of revenue productivity. I complement the firm-level data with product-level data where I observe product prices. This data allows me to estimate physical productivity. Together with revenue productivity, these variables will allow me to look at the

[3]For example, see Banerjee, Duflo, and Qian (2012), Chandra and Thompson (2000), Redding and Turner (2014) and Faber (2014)

effect of electrification on different measures of productivity. I then combine productivity estimates with firm market share data to study the effect of electrification on reallocation at an aggregate industry level.

This paper contributes to the literature on infrastructure and development. A strand of literature examines the effect of different types of infrastructure on economic outcomes. These include the effect of dams on agricultural productivity and poverty (Duflo and Pande (2007)), and the effect of transportation (roads, railways, highways) infrastructure on regional economic outcomes (examples include Donaldson (2010), Banerjee, Duflo, and Qian (2012), Faber (2014), Donaldson and Hornbeck (2016), and Gertler, Gonzalez-Navarro, Gracner, and Rothenberg (2014)). In terms of electrification infrastructure, a growing literature evaluates the effects of grid expansion as in Dinkelman (2011) who estimates the effect of electrification on employment in South Africa and Lipscomb, Mobarak, and Barham (2013) where they look at the effect of electrification in Brazil. Rud (2012) looks at the effect of electrification on industrialization in India at the state level. He shows that industrial output in a state increases with electrification.

While these papers focus on the extensive margin of electricity supply, many papers study the relationship between electricity supply and firms on the intensive margin, i.e. shortages. Reinikka and Svensson (1999) show that unreliable power supply in Uganda reduces private investment productivity by forcing firms to invest in generators and other low-productivity substitutes for reliable public provision of power. Fisher-Vanden, Mansur, and Wang (2015) use Chinese firm-level panel data to examine the response of firms to power shortages. They find that firms respond by re-optimizing among inputs, which increases their unit cost of production but allows them to avoid substantial productivity losses. Allcott, Collard-Wexler, and O'Connell (2016) find that electricity shortages in India reduce revenue but have no effect on revenue productivity.

Another strand of literature this paper is related to is the one on productivity and firm dynamics. Many papers study the determinants of firm turnover and its role in reallocating resources from less productive to more productive firms (examples include Syverson (2004), Syverson (2007), Foster, Haltiwanger, and Syverson (2008), Bartelsman, Haltiwanger, and Scarpetta (2013), Nguyen (2014)). An extensive literature as in Tybout (2000), Hsieh and Klenow (2009), Bloom, Mahajan, McKenzie, and Roberts (2010), and

Faber (2014) aims at understanding the productivity gap between firms in developing countries and firms in developed countries. These differences in productivity across countries imply substantial differences in aggregate performance of countries. Infrastructure is one suggested explanation to the lower productivity level of firms in developing countries, in particular, electricity access. I contribute to this literature in this paper by linking infrastructure to reallocation and turnover in explaining the low productivity of firms in developing countries.

My results show that electrification causes an increase in manufacturing activity. This is manifested by an increase in the number of firms, number of manufacturing workers, and manufacturing output at the desa level. Interestingly, electrification increases firm turnover by increasing entry rates but also exit rates.

At the firm level, I find that electrification causes average firm size to increase, both in terms of how much output the firm produces and how much inputs it demands. The results on firm turnover are confirmed in the firm-level analysis. Electrification increases the probability of exit, making it harder for inefficient firms to survive. In addition, electrification shifts the firm age distribution towards younger firms. This is a sign of churning in the industry, created by increased entry (more young firms) and increased exit (firms die more often).

At both the desa-level and the firm-level, I find that electrification creates new industrial activity, as opposed to only relocating economic activity from non-electrified areas to electrified areas.

Finally, I find that electrification increases average productivity, but using revenue productivity underestimates the productivity gains relative to physical productivity because of demand side biases. Finally, I use a decomposition of an aggregate revenue-weighted average productivity following Olley and Pakes (1996). I find that electrification increases allocative efficiency where the covariance between firm productivity and market shares is higher in electrified areas. These results are theoretically consistent with a decrease in the entry costs, suggesting that electrification increases aggregate productivity by allowing more productive firms in the market, increasing firm turnover, and enhancing allocative efficiency.

Section 2 below presents the institutional background of electrification in Indonesia, summarizing the history of the Indonesian power sector and the objective of the Indonesian government during the period of the study. Section 3 presents the data on electrification infrastructure and manufacturing activity, and describes the empirical strategy. Results on the effect of electrification and industrial outcomes are presented in section 4. In Sectrion 5, I explore whether electrification creates new economic activity or it relocates economic activity across space. Section 6 presents the results on electrification and manufacturing productivity. Section 6 concludes.

2 Institutional Background

2.1 History of the Indonesian Power Sector

Knowing the historical context of the power sector in Indonesia is crucial to understand why the Indonesian electricity supply was underdeveloped, including in Java. During the period of Dutch colonization of Indonesia, access to electricity was unequal and mainly reserved to colonial establishments. Between 1953 and 1957 the three Dutch owned electric utilities in Indonesia were nationalized by the Government. The transfer was not friendly, and was without a transition period where the new Indonesian management could have been trained by its colonial predecessors and many documents were destroyed in the process. Political unrest, lack of funds, hyperinflation and the lack of qualified management and engineers lead to a period of decline in efficiency, poor operating conditions, and inadequate expansion[4]. This in turn lead to a large electric supply deficit, which meant low household electrification ratios and that businesses and industries had to rely on self-generation. Over the next decades, with the help of various international aid agencies, PLN was expanding steadily both in terms of physical and human capital.

2.2 Objective of the Government of Indonesia 1990-2000

The main sources of electricity supply in Indonesia in the late 1980s and early 1990s comprised of PLN, the state electricity monopoly, and self-generation (around 40% of

[4]McCawley (1971).

generating capacity), mainly by the manufacturing sector. As Indonesia was witnessing an expansion of the PLN generation capacity, the manufacturing sector was shifting from relying exclusively on self-generation towards the use of captive generation for solely on a stand-by basis. Trends in PLN sales and captive power suggested that manufacturing firms, even after incurring the sunk cost of acquiring a generator, prefer grid electricity. This suggests that the marginal price of electricity from the grid is lower that the marginal price of electricity from self-generation. In 1989, the level of electricity consumption remained low in Indonesia relative to other countries at the same development level. This low level of electricity consumption was due to the lack of supply facilities. PLN's investment program in the late eighties was designed to meet the goals set by the Government's Five-Year Development Program (REPELITA V) by 1994. These included a 75% electrification ratio in urban areas, 29% electrification ratio, and finally, the substitution of 80% of captive generation by the industrial sector. The apparent objective of the Government at that time was to replace self-generation, i.e. providing grid electricity to non-connected incumbents, as opposed to expanding the grid to industrialize new locations. The subsequent Five-Year Development Program (REPELITA VI 1994-1999) by the Indonesian government had the following objectives for the power sector: (i) provide adequate, reliable, and reasonably priced supply of energy to rapidly growing economy, (ii) conserve and diversify the sources of energy, and (iii) minimize social and environmental adverse impacts. Goal (i) illustrates the simultaneity problem of growing adequate infrastructure provision and economic growth. The government of Indonesia was investing heavily in electricity supply to keep up with a rapidly growing economy, which poses the empirical challenge of identifying the causal effect of the expansion of electricity supply on industrial development. In 1997, the Asian financial crisis hit, followed by the end of the Suharto dictatorship and political unrest, which all lead to a lack of funds. Investment in the power sector continued during that period, albeit at a slower pace. By 2000, more than 90% of firms Java had access to electricity.

3 Data and Empirical Strategy

3.1 Data Description

I first construct a time-series of the electricity transmission network in Java between 1990 and 2000 using data from various sources. Java is the most dense island in Indonesia with 60% of the population and 80% of manufacturing firms. The main source are current administrative records from Perusahaan Listrik Negara (PLN), the Indonesian state electricity monopolist. I digitized information on the location, capacity and operation date of equipment within power firms and transmission substations in Java today from the PLN Head Office in Jakarta. To build the time-series from 1990 to 2000, gaps in administrative data were filled using World Bank power project reports, which evaluate electricity infrastructure loans given by the World Bank to Indonesian government between 1969 and 1996. In addition, because location data from PLN is not always accurate, I manually cross-checked power plant and substation coordinates using data downloaded from OSM (Open Street Maps). The resulting data-set is a panel of all transmission substations in Java. Figure 1 shows the expansion of the grid during the sample period. The expansion of the transmission grid in Java during that period was rapid and substantial. In 1990, the number of substations was 115. By 2000, there was a total of 279 transmission substations in Java. Total electricity transmission capacity increased from 6620 MVA to 25061 MVA, almost 4 times.

There are multiple units of analysis. I start my empirical analysis by looking at the effect of access on desa-level manufacturing outcomes. A desa is the lowest administrative division in Indonesia[5]. Data on desa level boundaries were acquired from BIG, the Indonesian National Mapping Agency. To get information on manufacturing activity in these desas, I use the Indonesian annual census of all manufacturing firms in Indonesia with 20 or more employees, where I observe in which desa each firm is located. I restrict the analysis to firms located in Java, which constitute around 80% of all Medium and Large firms in Indonesia. This allows me to create variables such as the number of manufacturing firms, number of manufacturing workers and total manufacturing output in each desa. The resulting data-set is a yearly balanced panel of all desas in Java from 1990 to 2000. I use

[5]There are 4 administrative divisions in Indonesia: province, regency, district and desa.

the the Desa Potential Statistics (PODES) survey for 1990, 1993, 1996 and 2000. The PODES data-set contains on all Indonesian desas, which I use to get data on desa level characteristics such as population, political status, development classification and most importantly, various infrastructure variables. These include information on the type of infrastructure available in the desa such as railway, motor station, river pier, road quality, and so on. This data is used to construct a digital map of desas in Java with various desa-level characteristics over time.

I then take advantage of the richness of information in the firm-level data from the census of manufacturing and analyze the effect of access to electricity on firm-level outcomes. The final level of analysis is at the product level. I supplement the firm-level data with product-level data at the 9 digit level where I observe the sales and physical output of each product produced by the firm. I can therefore calculate product price and using structural techniques of estimating production functions, I estimate physical productivity. This product data is however only available from 1994 onward. Finally, I use GIS data on cities, waterways, coastline and roads in Java. I measure the distance from each desa (centroid) to each of these geographic features in addition to the nearest electric substation and the hypothetical least cost grid. I also use data on elevation to measure land gradient at each location.

3.2 Empirical Strategy

The expansion of the grid is demand driven. In fact, PLN follows a demand forecast methodology where they forecast demand in a certain area and compare it to existing supply infrastructure. PLN then decides to expand it if they believe there will be a gap between supply and demand in the future. I explain this methodology in detail in Appendix E. Importantly, this methodology implies that the bias in ordinary least square estimates can go either way. On the one hand, more productive regions have higher demand forecasts, which means that OLS will be upward bias. On the other hand, areas with generally poor infrastructure, where firms are less productive, will have a higher gap between demand forecasts and existing supply, meaning that OLS will be downward bias. Another element in the decision of expanding the grid is cost of construction, which is potentially exogenous.

Using the data described above, I estimate the effect of access to the grid $Access_{vpt}$ on outcome Y_{vpt} of desa v, province p and year t using the following specification:

$$Y_{vpt} = \alpha + \beta Access_{vpt} + \eta \mathbf{V}_{vpt} + \gamma_p + \delta_t + \epsilon_{vpt} \tag{1}$$

and the firm-level equivalent where I estimate the effect access $Access_{vpst}$ on outcome y_{ivpst} of firm i in desa v, province p, industry s and year t.

$$y_{ivpst} = \alpha + \beta Access_{vpst} + \nu \mathbf{X}_{ivpst} + \eta \mathbf{V}_{vpt} + \gamma_p + \delta_{st} + \epsilon_{ivpst} \tag{2}$$

where $\mathbf{X}_{ivpst}$ is a vector of firm controls, $\mathbf{V}_{vpst}$ is a vector if desa level controls, γ_p are province fixed effects, δ_t are year fixed effects and δ_{st} are industry-by-year fixed effects.

Electricity grids are placed endogenously to industrial outcomes. Even conditional on all the listed controls, estimating the above model by OLS will give biased results. In order to deal with the endogeneity problem, I propose an instrumental variable approach exploiting a supply-side natural experiment. Up until the late 1980's, the electricity grid in Java was not interconnected. My empirical strategy exploits the fact that PLN needed to build an interconnection of the grid, which occurred by the start of my sample period. This interconnection created a change in the probability of receiving electricity in the future in certain desas that lie between two grids. The section below describes how this strategy in detail.

Hypothetical Least Cost Grid

In 1969, electricity grid in Java consisted of 5 different disconnected grids across the island (Figure 2). Having disconnected grids is inefficient, prevents load-sharing across regions, and increases the price of supplying electricity. Therefore, the 1970's and the 1980's witnessed a huge and successful effort by PLN with the help of agencies such as the World Bank and the Asian Development Bank to connect the various grids on the island (Figure 3). Various transmission lines were built for the main purpose of interconnecting the grid. As a result, desas nearby the lines connecting the grids faced an exogenous shock to the probability of receiving electricity access in the future as it is cheaper to connect desas that are closer to the existing network. To deal with the concern that transmission lines could be targeted at areas that are different than others, for example, non-farming land, I create a hypothetical grid to connect the main power firms in the separate grids. In total, I consider 15 power firms which I identify from historical maps as the main power-plants in the 5 separate grids. I calculate the least cost path for each pair of power-plants based on elevation and waterway data. I then use Kruskal's algorithm[6] to calculate the hypothetical least cost grid that connects all power plants on a single grid. Figure 4 shows the resulting hypothetical least cost grid. The distance to the hypothetical least cost grid is then used as the instrumental variable.

In addition, I control for various desa-level characteristics. One concern is that the location of the power plants is endogenous. In Java, many of these power plants are hydroelectric power plants, meaning their location is tied to the natural source. In addition, these power plants have been built by the Dutch electric utilities decades before the start of the sample period[7]. It is likely then that the factors determining the location of these power-plants do not directly affect outcomes in 1990 (conditional on controls). Nonetheless, I exclude desas within a certain radius of power-plants to deal with the concern that power plants are endogenously located. Power-plants are built close to the load centers that they are meant to supply electricity to in order to minimize transmission losses. Because load centers are typically cities and urban areas, one concern is that the instrument is correlated to distance to closest city. To alleviate this concern, I include distance to nearest city as a control variable. Because most economic activity is located along the coast of the island, many of the power plants are located there as well. One reason is that the coast is flatter and therefore it is cheaper to build there. In addition, proximity to coal sources for thermal power plants is crucial. Coal in Indonesia is mostly in the islands of Sumatera and Kalimantan, which are easily reachable from the north coast because of proximity and good wave conditions in the Java sea. Furthermore, because the coast is flatter, Kruskal's algorithm will favor lines along the coast. It is then important to control for distance to coast in any empirical specification to avoid any threats to exclusion. Controlling for desa elevation is also necessary because land gradient is correlated with distance to hypothetical least coast grid. Another potential confounder is the possible correlation between distance to the hypothetical grid and the road network on Java. For that reason, controlling for distance to road and road characteristics is important to guar-

[6]Kruskal's algorithm is a minimum spanning tree algorithm. The minimum spanning tree is the spanning tree that has the lowest cost among all the possible spanning trees. The cost of the spanning tree is defined as the sum of the weights of all the edges in the tree.

[7]http://maps.library.leiden.edu/apps/search?code=04693focus

antee the exclusion of the instrument. In all my specification, I control for the distance to the nearest regional road, as a well as various characteristics of the road at the desa level such as whether the road is wide enough for a four-wheel vehicle. I also control for the availability of non-energy infrastructure facilities. These include railway station, motor station, river pier, sea port, and airport. In addition to geographic controls, I also control for the desa political status and development classification. Political status is an indicator for whether the desa is the district capital. Desas are classified by the Government in three development categories; developed, transitional, and traditional. Bappenas, the planning ministry, gives priority to desas classified as developed. At the firm level, I control for whether the firm is public or private to deal with any favoritism in access towards government owned firms. I also control for firm age, legal status, and export status. The identification assumption is that, conditional on controls, the potential outcomes of desas or firms are independent of their distance to the hypothetical least cost grid.

To summarize, geographic desa controls include distance to coast, elevation, and distance to nearest city, road quality and distance to nearest road. Other desa level controls include various infrastructure availability dummies, political status, legal status and development classification. Firm level controls include firm age, export status, legal status and ownership type. Figure 5 illustrates the empirical strategy in a simplified manner. Consider two disconnected grids Grid 1 and Grid 2. These represent the incumbent infrastructure built by the Dutch electricity company and were existent by 1969. During the 1970s and the 1980s, the two grids became interconnected by the green line. Consider two firms (or desas) A and B that only differ in their distance to the green line. Because Firm A is closer to the green line, it is then more likely to get connected to the electricity grid in the 1990s compared to Firm B. The blue lines therefore represent the instrument. Because of potential concerns regarding the placement of the green line, I create a hypothetical green line that is based solely on cost factors. The hypothetical least cost grid is essentially an instrument for the actual interconnection transmission network.

4 Effect of Electrification on Industrial Outcomes

In this section, I use the data described above to measure the impact of the electricity grid expansion on the manufacturing sector. The first level of analysis is at the desa level, the lowest administrative division in Indonesia. I look at how electrification affects desa manufacturing activity. I choose to study desa-level outcomes because access to electricity varies at this same level. Second, I zoom in to the firm-level, where I compare the performance of firms in desas with and without access. In the final section, I am interested in studying the effect of access on manufacturing productivity. I start by looking at firm revenue productivity. To deal with biases related to revenue-productivity estimates, I take advantage of the supplementary product-level data on physical output and prices to analyze the effect of access on firm physical productivity. Finally, I study how reallocation at the industry-regency level is affected by access. An industry is defined as a two-digit industry code, and a regency is that second highest administrative division in Indonesia. This allows me to decompose industry-level market share weighted productivity into an average term and a covariance term, where the covariance term captures how efficient the market is in reallocating resources from less productive to more productive firms.

4.1 Access to Electricity at the Desa Level

I examine whether the expansion of the grid affected the number of firms, manufacturing employment and manufacturing revenue at the desa level. To this end, I construct a data-set of all desas in Java between 1990 and 2000. I superimpose the firm-level data on a digital map of desas in Java to get the number of firms per desa per year, as well as number of workers in manufacturing, and industrial output. I use a desa-level survey, PODES, to get information on local characteristics and the availability of different types of infrastructure[8] and political, legal and development status. Distance variables and elevation were measured in GIS and include distance from the desa centroid to the nearest transmission substation, the least cost hypthetical grid, the nearest incumbent power-plant, the nearest city, the nearest road, and to the coast. The resulting data-set is a balanced panel of $24,824$ desas over 11 years.

[8]railway station, motor station, seaport, river pier, airport, road width

4.1.1 First Stage: Desa-Level

The first Column of table 1 shows the first-stage regression using distance to the hypothetical least cost grid $Z(KM)$ as an instrumental variable and using all the controls discussed above. The dependent variable, $Access_{vt}$, is an indicator variable equal to one if the desa is within 15 KM[9] of the nearest transmission substation in year t. The coefficient in Column (1) is negative and significant, indicating that the further away a desa is from the hypothetical least cost network, the less likely it is to have access to electricity. The first stage F-statistic is high enough to guarantee relevance of the instrument, avoiding weak instrument bias. The coefficient in Column (1) then shows that even conditional on various controls, this difference in means is still significant and distance to the hypothetical grid is a good predictor of access to electricity at the desa level.

Figure 6 plots the unconditional probability of a desa having access to the grid as a function of the distance to the hypothetical least cost grid. The closer a desa is to the hypothetical grid, the more likely it is to have access to the actual grid. The relationship between the probability of access to the actual grid and the instrument is negative. I also plot the median and 90^{th} percentile of the instrument. At large values of the instrument, i.e. for desas very far from the hypothetical, the instrument doesn't predict the probability of access very well. However, this is not much of a concern as there are few observations in that region (beyond the 90^{th} percentile). Figure 7 plots the probability of a desa having access to the grid for the years 1990, 1995 and 2000, against the distance to the hypothetical grid. The graph shows that the negative relationship between access and the instrument persists over time. Holding distance to the hypothetical grid fixed, the probability of having access to the grid is increasing over time. This captures the fact that the electricity grid was expanded substantially between 1990 and 2000, increasing access from around 43% of Java's desas to 71%[10].

[9] This threshold was chosen based on conversations with electrical engineers at the Indonesian state electricity monopoly. The results are not sensitive to this particular choice.

[10] PLN reports an electrification ratio of 50% in 1990.

4.1.2 Access, Number of Firms, and Industry Outcomes

The first three Columns of table 2 shows the OLS, IV and reduced-form regression results for three desa-level outcomes as in specification (1) : number of firms, total number of workers in the manufacturing sector, and total manufacturing output. Because I have many desas that don't have any medium or large manufacturing firm, hence many zero values, I use the level of these variables instead of the log (See table C1 in the appendix for results with zero-preserving log transformations). Across all outcome variables, the OLS estimates in Panel A are positive and significant, suggesting that there is a positive correlation between access to electricity and industrial outcomes. Compared to the IV estimates in Panel B, OLS is consistently of smaller magnitude. This result is in line with the infrastructure literature both on electrification (e.g. Dinkelman (2011), and Lipscomb, Mobarak, and Barham (2013)) and transport (Baum-Snow (2007), Duranton and Turner (2012), and Duranton, Morrow, and Turner (2014)) indicating that infrastructure is allocated to less productive areas. This means that the OLS estimates will underestimate the effect of electrification on manufacturing, as the results show. The IV estimates in Panel B are positive and significant. The coefficient in Column (1) in panel Bsays that the causal effect of grid connection on the number of firms in a desa is an increase of 0.9 firm. Considering that the average number of firms per desa in the sample is 0.84, this effect is large and around 100% increase over the average. Theoretically, a larger number of firms is associated with a tougher competition. Therefore, electrification potentially intensifies competition by increasing the number of active producers. Similarly for the number of workers and manufacturing output, the IV estimates in Columns (2) and (3) are positive, large and strongly significant. A caveat is that I don't observe the universe of manufacturing firms, but instead I observe the universe of medium and large manufacturing firms with 20 or more employees. To mitigate this issue, for the number of firms, I use the reported start year of production in the survey as opposed to the first year I observe the firm in the data. I take that into account when calculating the total number of firms in a desa which greatly alleviates this issue.[11] As for the total number of workers and output, I don't observe any information for these firms before

[11] Of course, I still don't observe those firms that exited before they reached the threshold to be included in the survey. This is however not a major concern as these firm are naturally small both in number of workers and probably in production relative to the total manufacturing sector.

they are in the survey. Therefore coefficients in panel B Columns (2) and (3) should be interpreted as the causal difference in the number of workers and manufacturing output between connected and unconnected desas with Medium and Large manufacturing firms. The IV estimates are significantly larger in magnitude that the OLS estimates. Given that I am estimating a local average treatment effect of access on industrial outcomes; this difference in magnitudes is potentially driven by a complier sub-population of desas that would benefit *more* from electrification. A simple exercise of characterizing compliers as in Angrist and Pischke (2008) and more generally following Abadie (2003) shows that complier desas are almost three times more likely to be classified as a transitional, or emerging, desa according the Indonesian goverment's development classification[12]. It is likely that transitional villages have a higher marginal benefit of electrification in terms on industrializing relative to more developed desas.

Figures 8, 9 and 10 show the kernel regression of the of number of manufacturing firms, number of workers and manufacturing output as a function of the distance to the hypothetical least cost grid. The relationship between each of these desa-level outcome variables and the distance to the hypthetical gridis negative, illustrating the reduced-form effect of the instrument on the outcome variables. Panel C of table 2 presents the reduced-form regressions from regressing desa outcomes on the instrument, distance to the hypothetical grid. Coefficients in Columns (1), (2) and (3) all show the closer a desa is to the least cost network, the larger the number of firms, number of manufacturing workers and manufacturing output.

Columns (4) and (5) of table 2 looks at the effect of access on firm turnover. The first outcome is entry rate, defined as the ratio of entrants to the total number of firms. The second outcome variable is the exit rate, defined as the ratio of exiting firms to the total number of firms. These outcomes are only defined for desas with a positive number of firms. Focusing on panel B, the IV estimate in Column (4) show that access to the grid increases firm entry rate by around 10%. In Column (5), the coefficient on access shows that the exit rate also increases due to electrification, although by a smaller amount than the entry rate. This is consistent with the an increase in the total number of

[12]There are three desa-level development classifications. Swadaya (traditional), swakarya (transitional), and swasembada (developed). The proportion of each category in 1990 in Java are 5%, 20%, and 75% respectively.

manufacturing firms from Column (1). Electrification therefore increases firm turnover, leading to more churning in a given desa. Higher churning is a sign of efficiency where firm selection into and out of the desa is at work. These findings suggest that the extensive margin of electrification induces long-run firm responses; entry and exit.

4.2 Access to Electricity and Firm Performance

So far, results show that the expansion of the electricity grid caused an increase in manufacturing activity and increased firm turnover in Java. In the section, I test how average firm performance measures respond to access to electricity. I investigate whether firms in electrified desas are different. I begin by looking at firm output and inputs. I then look at how electrification is affecting firm turnover by looking at firm-level exit probabilities and the age distribution of firms. If firm turnover is a mechanism through which electrification affects manufacturing activity, we expect the probability of exit to be higher and firms to be younger in electrified desas.

4.2.1 First-Stage: Firm-Level

Before turning to the firm-level results, it is necessary to check whether my empirical strategy is still valid. I now check if distance to the hypothetical least cost grid still explains access to electricity at the firm-level. In the current section, I use the same definition of access, $Access_{vt}$. This is an indicator is equal to one if an firm is located in a desa within $15km$ of the nearest transmission substation. Based on the results from the previous section, firms are located in desas that are on average closer to the hypothetical least cost grid. One concern is therefore whether the instrument is still strong enough. The second Column of table 1 show the first-stage regressions of access on z_{vt}, the distance to the hypothetical least cost grid. In addition to the above controls defined at the desa-level, I include firm-level controls and year-by-industry fixed effects. The coefficient in Column (2) is negative and significant and the first stage F-statistic is high. The instrument is therefore still relevant. Figures 11 shows again a negative relationship between the unconditional probability of having access and distance to the least cost network.

4.2.2 Results: Firm Performance Measures

I first present the estimation results of specification (2) for different firm-level outcome variables. Table 3 shows the OLS, IV and reduced-form versions of specification (2) for the log values of firm-level deflated sales, deflated capital, wage bill, number of workers, energy bill and quantity of electricity consumed in kWh. The treatment variable here again is $Access_{vt}$, instrumented with z_v, the distance to the hypothetical least cost grid in kilometers. Table 3 Panel B shows that consuming grid electricity increases output and production inputs. The IV coefficients are all positive and significant at the 1% level. Looking at the first Column of Panel B, the causal effect of access on average firm sales is large and positive. Columns (2) to (4) show that connection also causes firm input demand for capital and labor (wage bill and number of workers) to increase substantially, with a larger effect on capital relative to labor. Perhaps not surprisingly, the effect on the energy bill in Columns (5), which include both spending on electricity and fuels, is the largest. Column (6) shows that firms with access to the grid do indeed consume a substantially greater quantity of electricity in kWh. The fact that electricity consumed increases by more than the increase in the energy bill reassuringly means that the unit price of electricity is lower in electrified areas. Panel C presents the results from the reduced-form regressions. Across all Columns, being closer to the hypothetical grid causes all firm-level outcomes to be significantly larger.

Relative to the existing literature, the most readily comparable results to what I find are from Allcott, Collard-Wexler, and O'Connell (2016). In their paper, the authors look at the effect of shortages on firm-level outcomes. They find that a 1 percentage point increase in shortages causes a 1.1% decrease in firm sales. Access to electricity can be thought of as a 100 percentage points decrease in shortages, which would then translate into a 110% increase in sales revenue. The effect, although comparable to the Allcott, Collard-Wexler, and O'Connell (2016) result, is four times larger. The size of the effect confirms the fact that the extensive margin of electricity supply has a more substantial effect on the industrial sector relative to the intensive margin. One explanation is that electrification is likely to reduce entry costs by more relative to improvements in the reliability of electricity supply. If sunk costs of entry are significantly affected by electrification, the effect on average firm outcomes will be larger. Allcott, Collard-Wexler, and O'Connell

(2016) also find that shortages do not affect labor input. In contrast, I find a large effect of access on number of workers, confirming that the extensive margin of electricity has a more considerable effect on the industrial sector.

I now examine whether electrification affects turnover in the economy. In other words, does the expanded access to electricity help with the weeding out of inefficient firms from the desa? I start by investigating the effect of electrification on the probability of exit. I estimate a linear probability model where I regress an exit dummy on access, instrumented with distance to the hypothetical and controlling for desa-level and firm-level characteristics as above. Before presenting the results, a discussion about how exit is defined is necessary. I define exit in period t as a dummy variable equal to one if the firm drops out of the census in period $t+1$. Because this is a census of firms with 20 or more employees, it could be that the firm did not actually exit the market, but instead shrank below the size threshold. For that reason, I restrict the definition of exiting firms to those who are not in the survey in year $t+1$ and have at least 25 employees, which is the 25^{th} percentile of size in the data. Table 4 Column (1) shows results from the OLS, IV and reduced-form regressions. Column (1) panel A presents the OLS estimate which is positive and significant indicating that the probability of exit and being a connected desa are positively correlated. The corresponding IV regression is in Column (1) panel B, and as before, the magnitude of the OLS estimate is smaller that the IV estimate. The coefficient shows that the causal effect of electrification on selection is an increase of around 5% in the probability of exit. Column (1) panel C show the reduced-form regressions of exit on the distance to the hypothetical least cost grid, showing that the closer the firm is to the least cost network, the more likely it is to exit. This suggests that survival in the industry is less likely in electrified desas.

Table 4 Column (2) shows the effect of electrification on the age distribution of firms. It presents the OLS, IV and reduced form regressions of a dummy variable $young_{it}$ equal to 1 if a firm is below the median age. Results show that firms in electrified desas are on average younger. This finding is consistent with electrification shifting the age distribution of firms towards younger firms by (i) increasing entry, therefore having more younger firms, and (ii) increasing exit, therefore shortening the average firm age in the desa.

5 Electrification, Industrial Development and Relocation of Economic Activity

The results in the previous section indicate that electrification increases industrial activity at the desa-level by attracting more firms. One important question is thus whether these firms are new firms or whether they are firms that have relocated from other non-connected desas. In particular, it is interesting to understand if these firms would have existed anyway, regardless of electrification. Results from the firm-level analysis show that connected firms sell more. Another interesting question is therefore whether these firms are stealing business from unconnected firms. In both this case of business stealing and the case where firms would relocate, the effect of electrification would be a reorganization of economic activity across the island as opposed to creation of *new* economic activity.

Put differently, a potential concern is that the stable unit treatment value assumption (SUTVA) is violated in the identification strategy of this paper. SUTVA requires that the treatment applied to one unit does not affect the outcome for another unit. If connecting one desa (or firm) to electricity will create firm relocation or business stealing for competitors (because of lower prices), then SUTVA is violated. The presence of spillovers across different desas complicates the interpretation of my results. Electrifying one desa can have an effect on firms in other desas, and these effects are likely to be negative. What I estimate as the average difference between electrified and non-electrified desas could be therefore a combination of creation of new economic activity and displacement of economic activity from those who don't get electrified (or are already electrified) to desas that get newly electrified.

In the following subsections, I attempt to address the question of whether electrification creates new economic activity or whether it is relocating economic activity. I start by looking at the possibility of firm relocation. I then look at demand side effects where connected firms steal business from non-connected firms.

5.1 Relocation of Firms

5.1.1 Relocation of Incumbent Firms

Can electrifying a new desa induce firms in non-electrified desas to close their factory and move it to the newly electrified desa? This could happen if a firm finds in profitable to do so, i.e. when the cost of relocation is smaller than the benefit of relocating. Firms choose to locate in certain desas presumably because the benefits from being in that location are the highest for that particular firm, so moving would be costly, in addition to the physical relocation costs.

Unlike a network of highways or subways, access to the electrification infrastructure is not restricted to particular locations such as a train station or a highway entrance. There is no technological limit on where the grid can go. In the context the island of Java, even if a desa is faraway from the grid at a certain point in time, it will eventually be connected to the grid. Given that this is a period of rapid expansion of the grid in Java, eventually all desas became connected to the grid. So unless the firm is really impatient, the benefit of moving to an electrified desa today versus waiting to get a connection in the future is unlikely to be a profitable action. Confirming this insight, I observe no firm movements across desas in the dataset[13],[14].

Finally, the evidence from desa-level regressions in table 2 Column (5) and firm-level regressions table 4 Column (1) shows that there is more exit in electrified desas. This result on exit rates is evidence against exit of firms from non-electrified desas to electrified desas.

5.1.2 Empirical Tests

To test whether relocation of firms is important in this context, I perform three main empirical tests. Given the technology argument made above and the rapid grid expansion,

[13]Less that 5% of the firms change desas between 1990-2000. I exclude these firms from the analysis.

[14]Another possibility is that entrepreneurs could be closing their factories in non-electrified desas and opening new factories producing *different* products in electrified desas. In this case, the firm will show up with a new firm identifier in the data, and it will be counted it as an exiting firm from the non-electrified desa and a new entry in the electrified desa. However, since I don't observe the identity of the owners, it is not possible for me to track this firm. Given that it is producing a different product, it wouldn't be unreasonable to consider this firm as a new firm.

relocation is likely to happen at a local geographic level where the benefits from being in different desas are comparable within a certain proximity. This argument applies both to incumbent firms as well as entrants. In fact, it is expected for these local spillover effects to be larger for entrants since these do not need to incur a physical cost of relocation.

First, I estimate equation (1) at the district[15]-level, a higher administrative division than a desa[16]. If spillovers are prominent, then the estimates should be smaller at the district-level. Table 7 presents the OLS and IV results. For comparability with the desa-level results in table 2, I use the average number of firms, average number of manufacturing workers and average manufacturing output in a district as opposed to the total[17] in Columns (1), (2) and (3) as the dependent variables. In Columns (4) and (5), I present the results for the entry and exit rates, defined as the total number of entrants and exiting firms divided by the total number of firms at the district-level, respectively. Comparing to the desa-level results, the effect of access on these industrial outcomes at the district level is very close to the effect at the desa-level. The estimated coefficients are if anything somewhat larger that the estimated coefficients from table 2, meaning that relocation of economic activity within district is unlikely. The IV results in Panel B therefore confirm that spillovers or relocation of economic activity are not prominent in this context.

Second, I test if an increase in the number of neighboring desas that switch for being non-electrified to electrified in a certain year negatively affects the number of firms and the number of entrants in desas that are not electrified and that do no switch. If there are any relocation effects, I would be expect them to be largest for this sub-sample.

I run the following specification where I test the effect of N^S_{vpt}, the number of switching neighboring desas on desa outcome Y_{vpt}, conditional on the total number of neighboring desas N_{vp} defined as the number of desas within a 7 km radius of the desa.

$$Y_{vpt} = \alpha + \beta N^S_{vpt} + \theta N_{vp} + \mu Z_{vp} + \eta \mathbf{V}_{vpt} + \gamma_p + \delta_t + \epsilon_{vpt} \quad (3)$$

Of course, N^S_{vpt} is endogenous. I instrument N^S_{vpt} with the average distance of neighboring desas to the hypothetical grid[18], conditional on the desa's distance to the least cost

[15]Kecamatan in Bahasa Indonesia

[16]The average number of desas per district is 16.

[17]Results are similar when using the total then dividing by average number of desas in a district.

[18]Variation in the shape of the grid across space means that there average neighbors distance to the grid and the desas own distance to the grid are not perfectly collinear. Interacting the IV with time dummies also helps with power.

hypothetical grid Z_{vp}.

Table 5 shows the OLS and IV results for this first test. Panel B Column (1) shows the IV estimate for the effect of an increase in the number of switching neighbors on the number of firms in the desa. The coefficient is statistically indistinguishable from zero and is small in magnitude. Give the mean number of switching neighbors in a given year for a given desa, this says that when one neighbor gets electricity in a certain year, the number of firms decreases by 0.007 firms; approximately zero. The coefficient in Panel B Column (2) shows the same IV regression for the number of entrants. The estimated effect is small and insignificant, but also positive. This shows that if a neighboring desas gets electrified, that does not decrease the number of entrants in the non electrified desa. Columns (3) and (4) panel B show the IV estimates for entry and exit rates. Results indicate that there is no effect of switching neighbors on firm turnover. In appendix C I show the same test in table **??** for the sub-sample restricting the sample to positive number of switching neighbors, where the effects should be larger. The results are similar and do not show any evidence for local spillovers.

Finally, I repeat the desa-level analysis from equation (1) but jointly estimating the main effect of access $Access_{vpt}$ and the spillover effect N^C_{vpt}. N^C_{vpt} is defined as the number of connected neighboring desas. I also condition on the total number of neighboring desas N_{vp}.

$$Y_{vpt} = \alpha + \beta Access_{vpt} + \mu N^C_{vpt} + \theta N_{vp} + \eta \mathbf{V}_{vpt} + \gamma_p + \delta_t + \epsilon_{vpt} \tag{4}$$

The coefficient on N^C_{vpt} will therefore measure the effect of having an additional electrified neighboring desa on desa outcome Y_{vpt}. If $\hat{\beta}$ and $\hat{\mu} * \bar{N^C_{vpt}}$ sum up to zero, where $\bar{N^C_{vpt}}$ is the average number of connected neighboring desas, then the effect of electrification evaluated at the average number of connected neighbors is only a relocation one. Otherwise, if the sum of $\hat{\beta}$ and $\hat{\mu} * \bar{N^C_{vpt}}$ is larger than zero, then electrification creates *new* economic activity.

As before, I instrument access with the desa's own distance to the hypothetical grid, and the number of connected neighbors by the average distance of neighbors to the hypothetical grid, both interacted with time dummies to aid with power.

Table 6 presents the OLS and IV results of equation (4). Focusing on the IV results in panel B, the estimated coefficients across all industrial outcomes are comparable to the IV results in table 2. The effect of access on industrial outcomes is positive and significant. On the other hand, the IV estimate for the effect of the number of connected neighbors N^C_{vpt} is small and negative, but not always significant. It is significant only in Columns (3), (4) and (5). This indicates that spillovers are stronger in the output market, consistent with high relocation costs of firms and workers. The last row of table 6 presents the p-value of the joint test where the null is $H0 : \hat{\beta} + \hat{\mu} * \bar{N^C_{vpt}} = 0$. The null is rejected in Columns (1) to (4). This indicates that indeed electrification does create new economic activity, and the effects are not restricted to relocation of economic activity.

5.2 Business Stealing Effects

In reality, firms in certain desa can potentially sell in different desas. An important question is whether there is any creation of new output in response to electrification. To check if spillovers or business stealing effects are present in my context, I run three tests. The extent to which these spillovers exists depends on various factors. First, it depends on the type of goods produced and their tradability. For example, we except these spillovers to minimal in the context of non-tradable goods. To test this, I estimate the effect of access on firm sales in the non-tradable sectors[19]. I consider certain products to be non-tradables because of their heavy weight which involves really large transportation costs. Table 8 presents the IV results for this exercise. I find a coefficient of 2.3, which is very close to the estimate found using the whole sample in table 3 panel B Column (1). This shows that in a setting where business stealing effects should be minimal because of large transportation costs, electrification still increases average firm sales. This indicates that there is some new economic activity being generated from electrification. Given the large number of desas (23,000 per year), and the large number of firms (16,104 per year on average), such spillover effects are unlikely in my settings because each unit is too small to affect its competitors if you consider Java as one single market. To confirm the absence of spillovers, I test for general equilibrium effects by regressing firm sales on the number of switching neighboring districts:

$$y_{ivpst} = \alpha + \beta M_{vpst} + \eta \mathbf{X}_{ivpst} + \theta Z_{vpt} + \eta \mathbf{V}_{vpst} + \gamma_p + \delta_{st} + \epsilon_{ivpst}$$

[19]These are two three-digit industries (263 and 264 ISIC Rev3). They include the following categories: Refactory bricks, clay products, clay bricks, clay tiles, structural clay, cement, lime plaster, gips

The idea is that if spacial spillovers exist, then the number of switching districts around the firm desa should affect firm revenue negatively. Here the assumption is that trade costs are infinite for further away districts. It is a strong assumption but it is supposed to capture that trade costs increase with distance. If there are spillovers, they will be strongest between neighboring districts. Because the number of switching neighbors is endogenous, I instrument for it with the average distance to hypothetical least cost grid in the district, conditional on the firm's own distance to the least cost network[20]. Table 9 shows the corresponding OLS and IV regressions. Column (2) presents the IV regression. The coefficient on number of switching firms is negative statistically insignificant. This rejects the presence of spacial spillovers. Even if the coefficient were to be significant, the implied effect is very small[21] (0.3) relative to the effect of access I find in table 3 Column (2) Panel B and cannot explain more than 13% of the difference in average sales between connected and unconnected firms.

Finally, I look for spillovers within narrowly defined industries across the whole island. Results are presented in table 10. The estimated effect of an increase in the number of switching competitors on the sales of non-switchers is a precisely estimated zero. I estimate this relationship again by industry, and I find no evidence for spillovers. For some industries, I find a precisely estimated zero. For other industries, the coefficients are statistically indistinguishable from zero, and the magnitudes are small: spillovers can explain at most around 20% of the estimated effect.

6 Electrification and Manufacturing Productivity

Results in the previous section show that electrification induces long-run responses in firm decisions by affecting firm entry and exit. I this section, I try to understand how these long-run responses translate into productivity effects. I also examine if the increase in churning resulting from electrification implies reallocation of activity towards more productive firms. I first need to measure productivity. A large literature exists on productivity estimation methods and the multiple difficulties involved. Typically, measuring

[20] These two distances are not collinear given the variation in the shape of the hypothetical least cost grid.

[21] This is equal to the estimated coefficient -0.108 times the average number of switching neighbors, which is around 3

productivity requires estimating a production function at the industry level, and productivity is calculated as the residual between firm output and firm predicted output using production function estimates and observed inputs. The first challenge in to estimate a production function consistently. This is because of the simultaneity bias stemming from the fact that productivity is unobservable to the econometrician but is observed by the firm when it chooses its flexible inputs. The second challenge arises from the data: we typically observe firm revenue and not physical output. Using revenue instead of quantity can confound productivity estimates by demand shocks and markups. In what follows, I will present two measures of productivity. The first is a firm-level productivity measure using revenue data estimated following Olley and Pakes (1996), a method that deals with simultaneity bias. I then move to a product-level analysis where I observe quantity produced to avoid demand side biases.

6.1 Measuring Productivity

6.1.1 Firm-Level Revenue Productivity

In this section, I investigate the effect of electrification on average firm-level productivity. Productivity is defined as the efficiency with which a firm transforms inputs into output. Let $F(.)$ be an industry level production technology. Output quantity Q_{it} of firm i in year t if produced according to $Q_{it} = exp(\phi_{it})F(\mathbf{X}_{it}, \beta)$. Firm productivity is ϕ_{it}, $\mathbf{X}_{it}$ is a vector of production inputs; capital, labor, and electricity. Typically, physical output Q is not observed. Instead we observe firms sales revenue $R_{it} = P_{it} * Q_{it}$. Consider the revenue based production function (in logs):

$$r_{it} = p_{it} + q_{it} = f(\mathbf{x}_{it}, \beta) + \phi_{it} + p_{it} + \epsilon_{it} \tag{5}$$

where ϵ_{it} is an error term. Since also prices are unobersvable, the literature typically estimates revenue productivity, or profitability, TFPR, defined as:

$$TFPR_{it} = \phi_{it} + p_{it} \tag{6}$$

Since TFPR is unobservable, and it is correlated with inputs, estimating the production function with OLS will give biased estimates of the production function coefficients. Following Olley and Pakes (1996), I use investment as a proxy for productivity, and assume a Cobb-Douglas production function. This methodology accounts for the simultaneity bias by proxying for the omitted variable, productivity ϕ_{it}. Under the assumption of monotonicity, more productive firms will invest more. Therefore, using a first order condition of the firm optimization problem, investment can be inverted to infer productivity. This method however fails to account for demand side biases caused by the presence of price in $TFPR$. The goal is to check if connected firms have on average higher physical productivity, ϕ_{it}. Testing this channel with regressions of $TFPR_{it}$ on access is not ideal. To see why, consider equation (6). Suppose that access increases the average productivity ϕ_{it}. But price and productivity ϕ_{it} are negatively correlated: more productive firms have lower marginal costs and therefore lower prices. This means that if access increases the average ϕ_{it} in the market and decreases the average price, the two effects can potentially cancel out.

6.1.2 Product-Level Physical Productivity

This calls for the estimation of a quantity-based product-level production function to avoid these biases. I therefore take advantage of price and physical quantity data which I observe (and are most likely set) at the product level. Two additional biases arise in this case. The first is an input price bias since input quality is not observed. The second is the input allocation bias as input allocation across products within multi-product firms[22] is unobserved. I closely follow De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) in dealing with these biases with two differences. The first is the choice of inputs in the production function. I use a Translog production function in capital, labor and electricity. The choice of functional form allows for a richer substitution pattern (relative to a Cobb-Douglas) between inputs to understand the role of access to energy in affecting marginal cost. Second, I allow unobservable input prices to depend on access. I describe briefly the procedure below[23].

Production Function Estimation

[22] The median number of products per firm per year is 2.

[23] I refer the reader to De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) for a more detailed discussion.

First consider the production function of product j produced by firm i in year t in logs:

$$q_{ijt} = f_j(\mathbf{x}_{ijt}, \beta) + \phi_{it} + \epsilon_{ijt} \tag{7}$$

where the vector $\mathbf{x}_{ijt}$ contains $k_{ijt}, l_{ijt}, e_{ijt}$, the product specific physical capital, labor, and energy and β is a vector of production function parameters. In practice, for input x, we observe a deflated version of x_{ijt} at the firm level $\tilde{x}_{it}$ where the following relationship holds in logs:

$$x_{ijt} = \rho_{ijt} + \tilde{x}_{it} - w^x_{ijt} \tag{8}$$

In equation 8, ρ_{ijt} is the log share of firm input expenditure dedicated to product j and w^x_{ijt} if the log deviation of firm-product specific price of input from the industry average. Substituting 8 in 7 yields:

$$q_{ijt} = f_j(\tilde{\mathbf{x}}_{ijt}, \beta) + A(\rho_{ijt}, \tilde{\mathbf{x}}_{ijt}, \beta) + B(w^x_{ijt}, \rho_{ijt}, \tilde{\mathbf{x}}_{ijt}, \beta) + \phi_{it} + \epsilon_{ijt} \tag{9}$$

The $A(.)$ function represents the bias stemming from unobserved input allocation across products within firm. I deal with this bias first by estimating the production function for single product firms only[24] while correcting for selection into being a single product firm[25]. The $B(.)$ term represents the input price bias. De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) show that input prices are a function of output prices p_{it}[26] and other variables proxying for product quality such as market share ms_{it}, location dummies G_i and product dummies K_i. In addition to these variable, I allow input prices to depend on access C_{it}. This gives rise to the following input price control function[27]:

$$w^x_{it} = w_t(p_{it}, ms_{it}, K_i, G_i, C_{it}) \tag{10}$$

This leaves one bias remaining, which is the classical bias from unobserved productivity ϕ_{it}. I follow the literature as in Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg, Caves, and Frazer (2015) by using the first order condition of a variable input, in my case electricity spending, as a proxy for productivity[28]. Given the estimated

[24] This is a sub-sample of all firms that are producing a single product at any point in time, including firms that become multiproduct firms in later periods (and vice versa) and those who remain single product.

[25] a procedure similar to controlling survival as in **?**.

[26] Vertical differentiation model

[27] Coefficients of the input price control function are not separately identified by input, so they have to be firm specific instead of product specific.

[28] I implement the one step estimator as suggested by Wooldridge (2009)

production function coefficients and the input price control, ϕ_{it} and the ρ_{ijt}'s can be solved for using the residual from 9, as the only unknown is the ρ_{ijt}'s from $A(.)$ function[29] and ϕ_{it} is the constant. The production function estimates and average output elasticities can be found in tables C3 and C4 in Appendix C.

6.2 Results

I first estimate the effect of electrifying a desa on average revenue productivity estimated following Olley and Pakes (1996) by running the following regression:

$$TFPR_{ivpst} = \alpha + \beta Access_{vpst} + \nu \mathbf{X}_{ivpst} + \eta \mathbf{V}_{vpt} + \gamma_p + \delta_{st} + \epsilon_{ivpst} \quad (11)$$

Table 11 shows the OLS, IV, and reduced form results. The OLS estimates are again smaller in magnitude than the IV estimate. Focusing on Column (1) panel B, I find that on average electrifying a desa increases revenue productivity in the desa. To explore heterogeneity in the effect of access average revenue productivity across entrants and incumbents, proxied by firm age, I estimate the same equation for young and old firms separately. A young firm is a firm whose age is below the median age. IV regressions in panel B show that this increase in average revenue productivity is driven by an increase in the revenue productivity of younger firms. This evidence is not necessarily consistent with a turnover channel where electrification induces the inefficient incumbents to exit. We would expect that in that case the average productivity of older firms is also higher. However, given that TFPR estimates are a combination of productivity and prices, this could be driven by a differential effect of access on prices for younger and older firms. To separate the effects, I use the product-level price data and physical productivity estimates. I estimate the following equation for product j (which is a subset of industry s) produced by firm i in desa v, province p, industry s and year t is:

$$y_{jivpst} = \alpha + \beta Access_{vpst} + \nu \mathbf{X}_{ivpst} + \eta \mathbf{V}_{vpt} + \gamma_p + \delta_t + \delta_j + \epsilon_{jivpst} \quad (12)$$

where δ_j are product-level fixed effects.

Table 12 shows the results from regressing log price and ϕ_{it} on access for all, young and old firms. The OLS estimates in panel A are smaller in magnitude than the IV estimates

[29] We know the functional form

as before. The IV estimates of the effect of access on ϕ_{it} in panel B Columns (2), (4) and (6) are all positive, significant and of the same magnitude, indicating that the difference in the average physical productivity of connected and unconnected firms is the same across firm cohorts. The coefficient in Column (3) panel B shows that the difference in price between products produced by young connected firms and young unconnected firms is not statistically different from zero. However, there is an negative effect of access on the average price of products produced by older connected firms. This explains the results on TFPR from table 11. These findings indicate that access to electricity increases average productivity by bringing in more productive firms to the market, and exerting competitive pressure on incumbents, leading to an increase in the average productivity of incumbents.

6.2.1 Reallocation at the Regency-by-Industry Level

The evidence so far indicates that electrification increases firm turnover in a desa by allowing more firms in and increasing the probability of exit. This leads to an increase in the average firm productivity in the manufacturing sector. Does electrification improve the reallocation of resources towards more productive firms? To answer this question, I aggregate revenue productivity at the regency-by-industry level. A regency is the second highest administrative division in Indonesia. There are around 100 regencies in Java. On average a regency has 250 desas and around 250 firms per regency. An industry is a two-digit industry classification. I call each regency-by-industry pair a sector. I decompose the sector TFPR index Ω_{st}, defined as the revenue-weighted average of log firm revenue productivity TFPR in an industry s in year t, into an unweighted average and a covariance term (Olley and Pakes (1996)):

$$\begin{aligned}
\Omega_{st} &= \sum_{i=1}^{N} S_{it} TFPR_{it} \\
&= \frac{1}{N}\sum_{i=1}^{N} TFPR_{it} \sum_{i=1}^{N}(S_{it} - \frac{1}{N})(TFPR_{it} - \frac{1}{N}\sum_{i=1}^{N} TFPR_{it}) \\
&= \overline{TFPR_{st}} + Ncov(S_{it}, TFPR_{it})
\end{aligned} \tag{13}$$

where S_{it} is firm i revenue share in sector s. $\overline{TFPR_{st}}$ is the unweighted average of log revenue productivity across all firms in industry s in year t. The Olley-Pakes covariance term measures allocative efficiency. It is higher when more productive firms have larger market shares. I test how electrifying more desas within a regency affects the industry. I define $Access_{st}$ as a dummy = 1 if at least 0.5 of firms are within $15KM$ of the nearest substations. I use a similar identification strategy as at the desa level where I instrument access with the average distance in the industry to the hypothetical grid. The estimating equation is:

$$Y_{st} = \alpha + \beta Access_{st} + \gamma_{pt} + \delta_s + \epsilon_{st} \tag{14}$$

where with province-by-year fixed effect and sector fixed effect. Table 13 presents the results. The IV estimates in panel B show that access increases both weighted and unweighted productivity at the sector level. In addition, the Olley Pakes covariance term increases with access. This means that electrification increases the covariance between market share and revenue productivity. Reallocation is more efficient in regions-by-industry groups with larger electrified proportions. This is evidence for a firm turnover mechanism where electrification helps reallocating resources towards more productive firms.

7 Conclusion

In this paper, I show that electrification has a substantial causal impact on the industrial sector. I highlight a new mechanism through which this effect can occur. This mechanism, firm turnover, is unlikely to operate in response to short-run improvements in electricity supply. The extensive margin of electrification induces extensive margin responses in firm decisions, which affects the composition of firms in the industry. Electrification attracts more firms into a market. This creates more competition and makes it more difficult for unproductive firms to survive. By increasing firm turnover, electrification increases average productivity in the market. This mechanism is similar to selection induced by trade liberalization where exposing domestic firms to international competition forces the least productive firms to exit as in Pavcnik (2002) and Melitz (2003). Electrification therefore promotes industrial development by increasing the efficiency with which markets allocate resources from unproductive firms towards more productive firms.

While the infrastructure literature has made substantial progress in understanding the effect of transportation (roads, railways) on development, we are at the very beginning of understanding how access to energy affects economic development. This paper has taken a small step towards a better understanding of the relationship between energy infrastructure and development. However, there is still a lot to be learned. Electrification projects are typically large-scale costly investments and it is important to quantify their benefits. In some instances, like in Lee, Miguel, and Wolfram (2016) and Burlig and Preonas (2016), benefits from electrification do not necessarily justify the investment and are not as large as we expect them to be. Large investments in electrification have been made in various African countries over the last decades, but Africa is yet to industrialize.

It is therefore important to understand how electrification and other institutional features might interact. For instance, other large institutional barriers to entry or to market access might prevent electrification from triggering entry and allowing for productivity gains. In the presence of credit constraints, the effect of electrification could be even larger, because it can lower the cost of entry for constrained entrepreneurs and reduce the extent of misallocation. These are a few of the open questions that remain to be answered in future work on electrification and development.

Once we have a better understanding of how and when access to energy leads to growth, it is then important to think about how we can provide energy and use it to grow the economy without harming the environment. Energy is potentially essential to bring people out of poverty, but it is also important to provide it in a cheap and sustainable way. This provides us with a new set of challenges and research opportunities that we have not thought about previously in the experience of electrification and industrialization in the developed world.

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of econometrics 113*(2), 231–263.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

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

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

Banerjee, A., E. Duflo, and N. Qian (2012). On the road: Access to transportation infrastructure and economic growth in china. Working Paper 17897, National Bureau of Economic Research.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013). Cross-country differences in productivity: The role of allocation and selection. *The American Economic Review 103*(1), 305–334.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics 122*(2), 775–805.

Bloom, N., A. Mahajan, D. McKenzie, and J. Roberts (2010, May). Why do firms in developing countries have low productivity? *American Economic Review 100*(2), 619–23.

Burlig, F. and L. Preonas (2016). Out of the darkness and into the light? development effects of rural electrification. *Working Paper*.

Chandra, A. and E. Thompson (2000). Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional Science and Urban Economics 30*(4), 457–490.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica 80*(6), 2543–2594.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016). Prices, markups, and trade reform. *Econometrica 84*(2), 445–510.

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

Donaldson, D. (2010). Railroads of the raj: Estimating the impact of transportation infrastructure. Technical report, National Bureau of Economic Research.

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A "market access" approach *. *The Quarterly Journal of Economics 131*(2), 799–858.

Duflo, E. and R. Pande (2007). Dams. *122*(2), 601–646.

Duranton, G., P. M. Morrow, and M. A. Turner (2014). Roads and trade: Evidence from the us. *The Review of Economic Studies 81*(2), 681–724.

Duranton, G. and M. A. Turner (2012). Urban growth and transportation. *Review of Economic Studies 79*(4), 1407–1440.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from china's national trunk highway system. *The Review of Economic Studies*.

Fisher-Vanden, K., E. T. Mansur, and Q. J. Wang (2015). Electricity shortages and firm productivity: Evidence from China's industrial firms. *Journal of Development Economics 114*(C), 172–188.

Foster, L., J. Haltiwanger, and C. Syverson (2008, March). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review 98*(1), 394–425.

Gertler, P. J., M. Gonzalez-Navarro, T. Gracner, and A. D. Rothenberg (2014). The role of road quality investments on economic activity and welfare: Evidence from indonesia's highways. *Unpublished manuscript*.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Lee, K., E. Miguel, and C. Wolfram (2016). Experimental evidence on the demand for and costs of rural electrification. Technical report, National Bureau of Economic Research.

Lenin, V. I. (1920). Collected works. *Moscow: Progress Publishers XXXI*, 513–518.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies 70*(2), 317–341.

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

Liu, Q. and L. D. Qiu (2016). Intermediate input imports and innovations: Evidence from chinese firms' patent filings. *Journal of International Economics 103*, 166–183.

McCawley, P. a. (1971). *The Indonesian electric supply industry.* Ph. D. thesis, The Australian National University.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Nguyen, K. T. (2014). Privatization and productivity of upstream industries: Firm and industry level evidence from vietnam. *Working Paper*.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from chilean plants. *The Review of Economic Studies 69*(1), 245–276.

Redding, S. J. and M. A. Turner (2014). Transportation costs and the spatial organization of economic activity. Working Paper 20235, National Bureau of Economic Research.

Reinikka, R. and J. Svensson (1999). How inadequate provision of public infrastructure and services affects private investment. Policy research working paper series, The World Bank.

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

Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy 112*(6), 1181–1222.

Syverson, C. (2007). Prices, spatial competition and heterogeneous producers: an empirical test. *The Journal of Industrial Economics 55*(2), 197–222.

Tybout, J. R. (2000, March). Manufacturing firms in developing countries: How well do they do, and why? *Journal of Economic Literature 38*(1), 11–44.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters 104*(3), 112 – 114.

A Figures

Figure 1: Expansion of the Grid 1990-2000

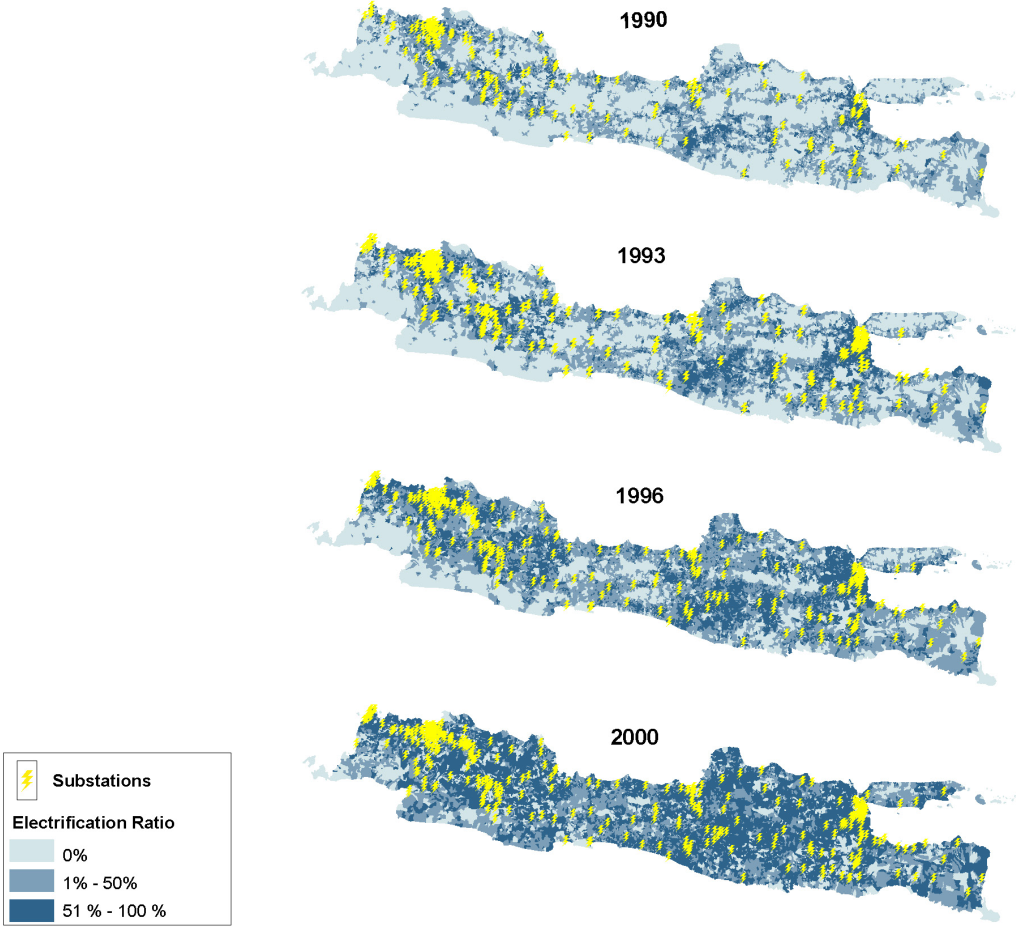


Figure 2: Java Network 1969

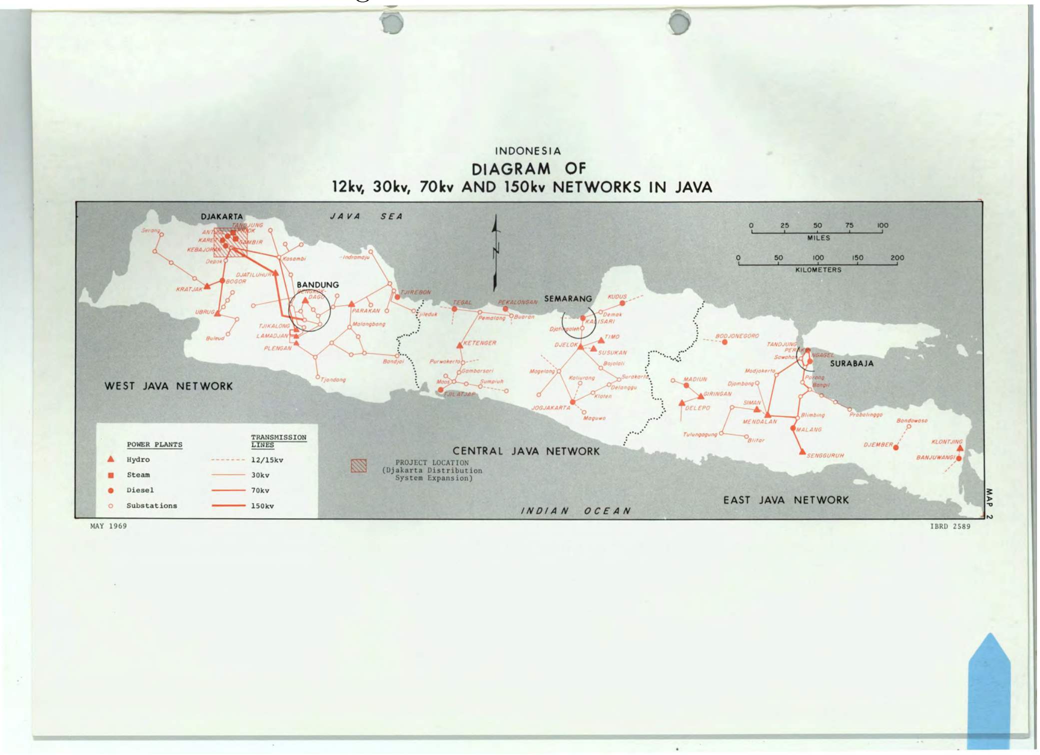


Figure 3: Java Network 1989

Figure 4: Least Cost Network

Figure 5: Empirical Strategy

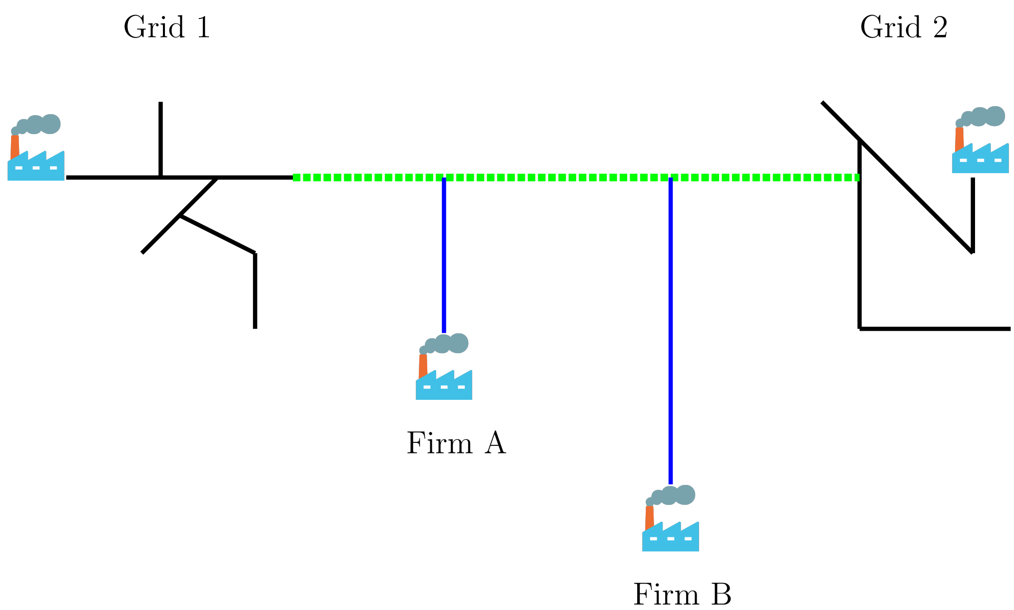
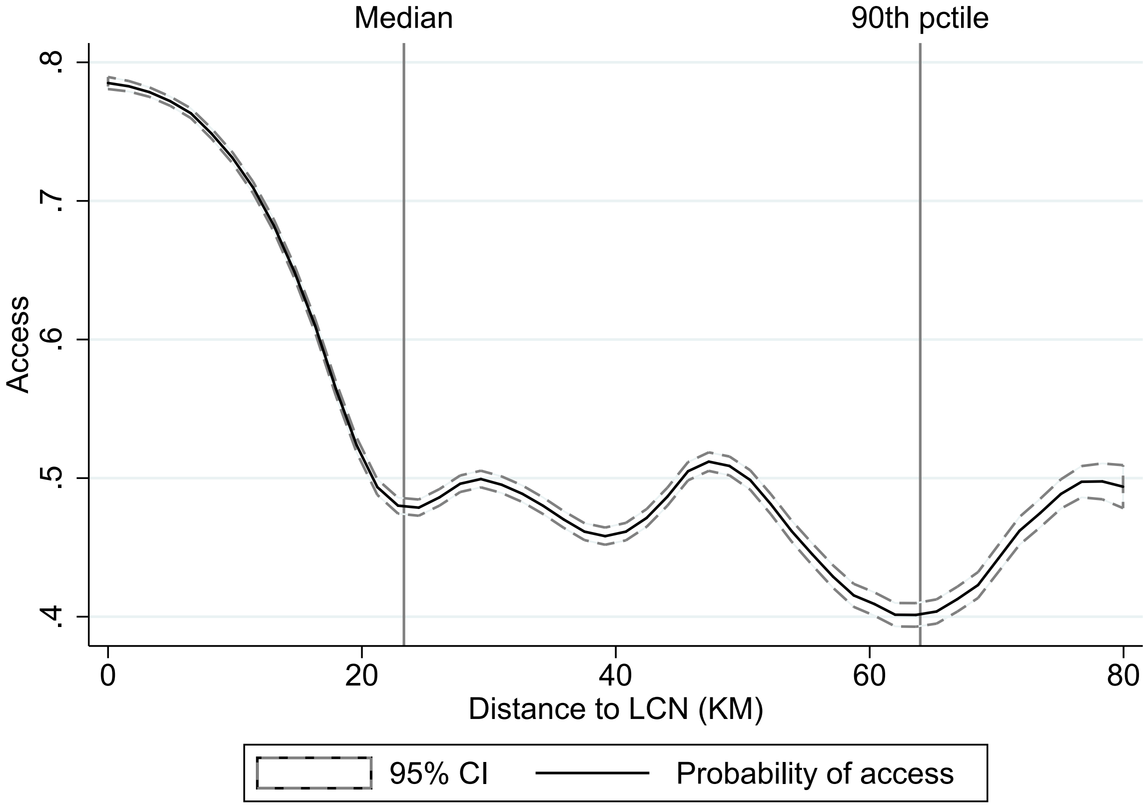
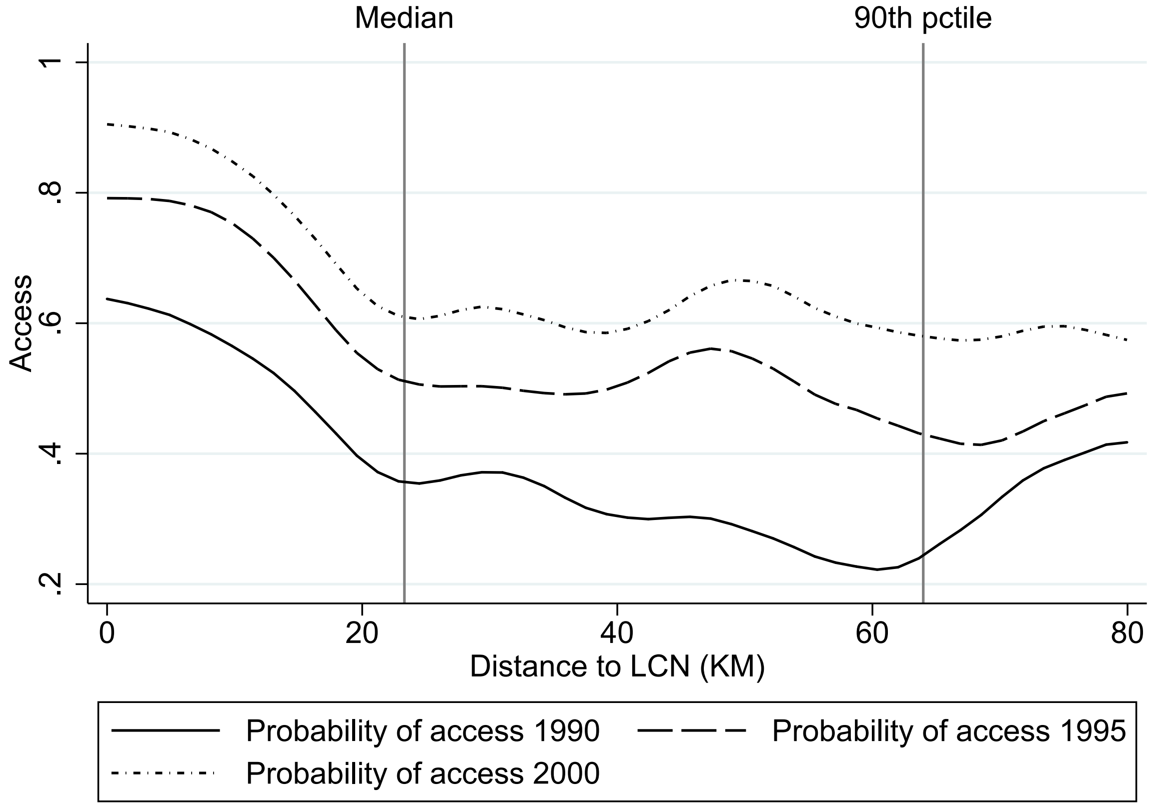


Figure 6: Distance to Hypothetical Grid and Probability of Being Connected



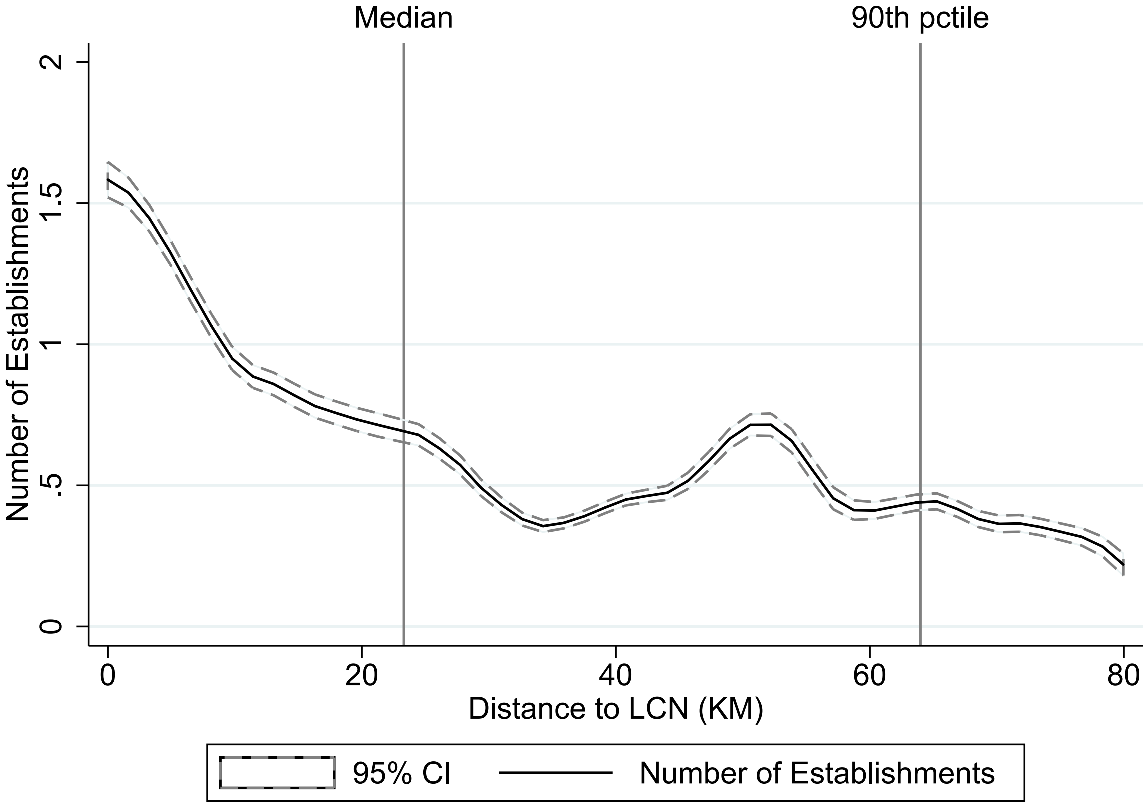
The y-axis presents probability of a desa being connected to the grid, where $Access_{vt}$ is a dummy variable equal to 1 if a desa is within 15 KM of the nearest transmission substation. The probability is estimated using an Epanechnikov kernel function with a bandwidth of 2.16. The x-axis shows the distance from the desa to the hypothetical least cost grid. The median and the 90^{th} percentile of the distance to the hypothetical gridare shown for reference.

Figure 7: Distance to hypothetical gridand Probability of Being Connected, by Year



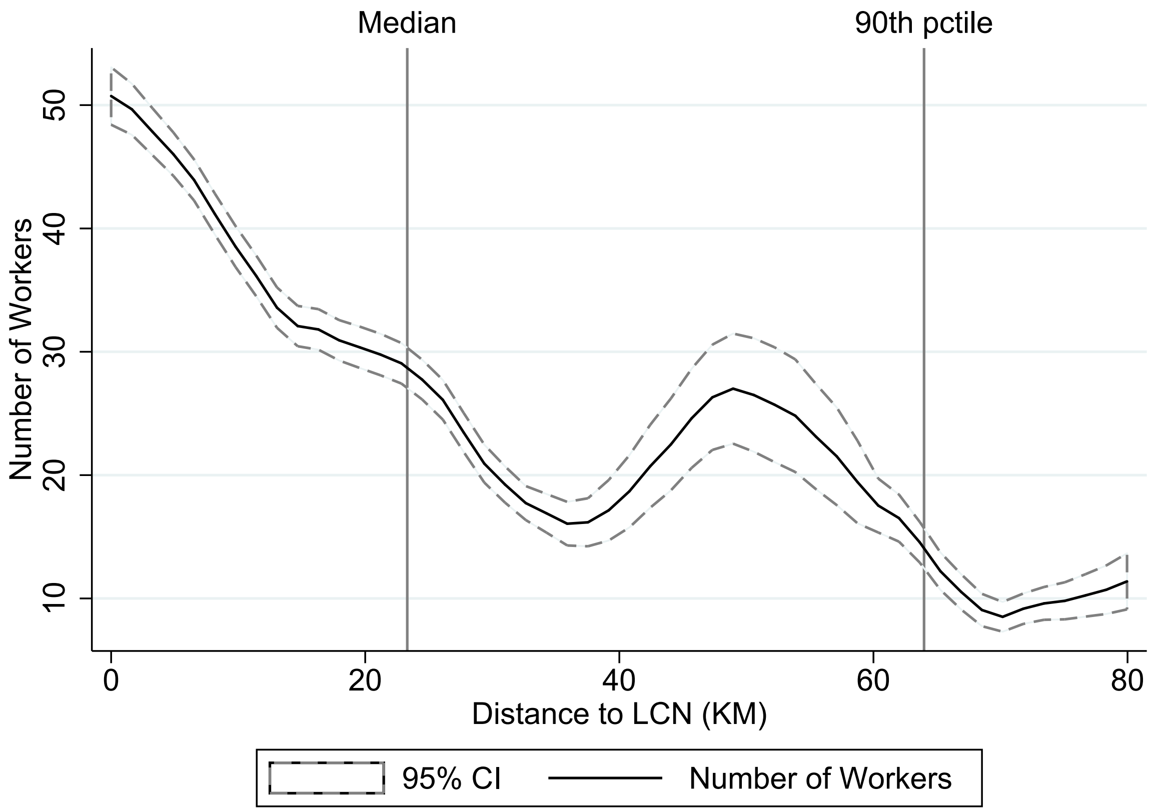
The y-axis presents probability of a desa being connected to the grid for years 1990, 1995 and 2000, where $Access_{vt}$ is a dummy variable equal to 1 if a desa is within 15 KM of the nearest transmission substation. The probability is estimated using an Epanechnikov kernel function with a bandwidth of 2.16. The x-axis shows the distance from the desa to the hypothetical least cost grid. The median and the 90^{th} percentile of the distance to the hypothetical gridare shown for reference.

Figure 8: Distance to hypothetical gridand Number of Manufacturing firms



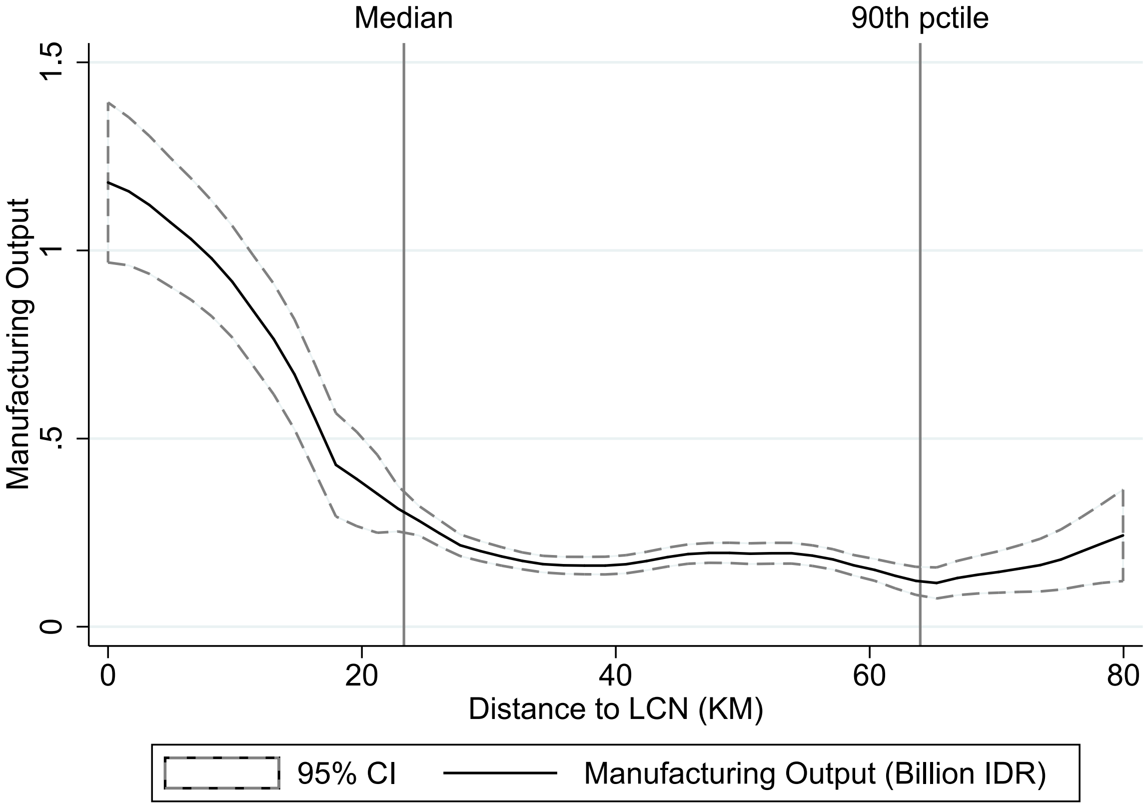
The y-axis presents the number of manufacturing firms at the desa level as a function of the distance of that desa to the hypothetical least cost grid. This is estimated using an Epanechnikov kernel function with a bandwidth of 2.42. The x-axis shows the distance from the desa to the hypothetical least cost grid. The median and the 90^{th} percentile of the distance to the hypothetical gridare shown for reference.

Figure 9: Distance to hypothetical gridand Number of Manufacturing Workers



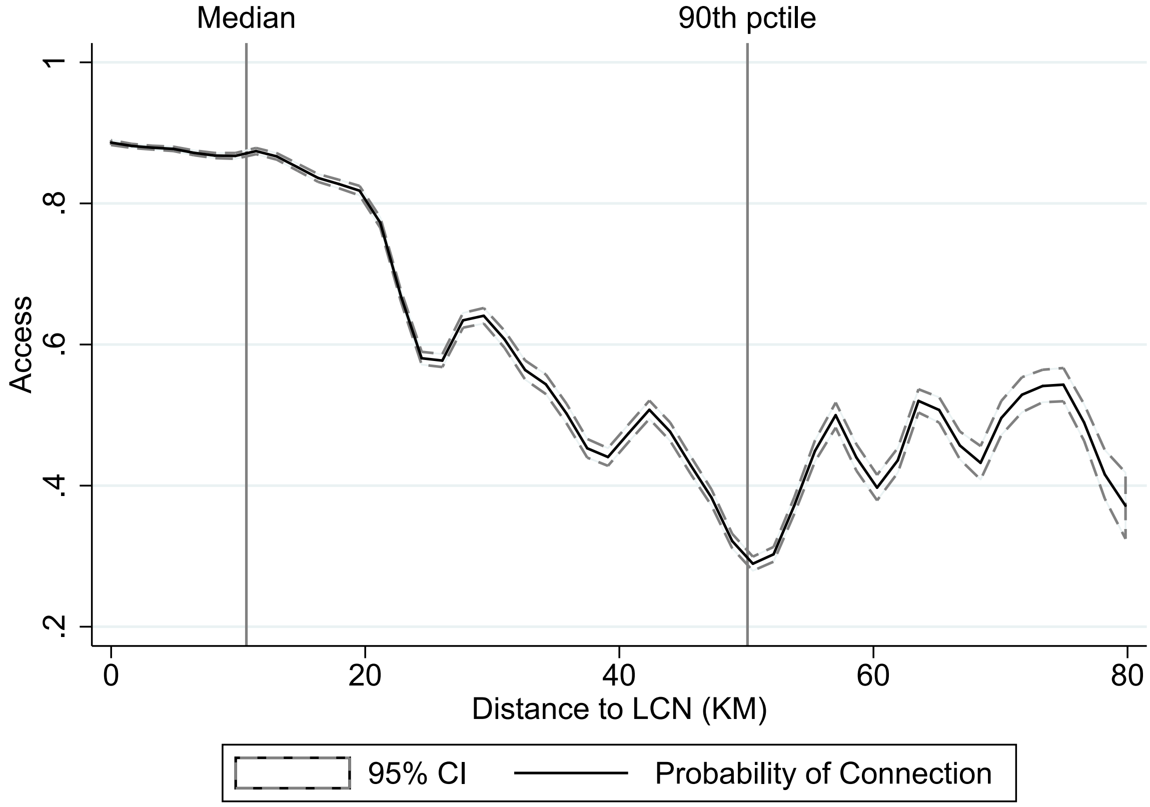
The y-axis presents the number of manufacturing workers at the desa level as a function of the distance of that desa to the hypothetical least cost grid. This is estimated using an Epanechnikov kernel function with a bandwidth of 3.35. The x-axis shows the distance from the desa to the hypothetical least cost grid. The median and the 90^{th} percentile of the distance to the hypothetical gridare shown for reference.

Figure 10: Distance to hypothetical gridand Manufacturing Output



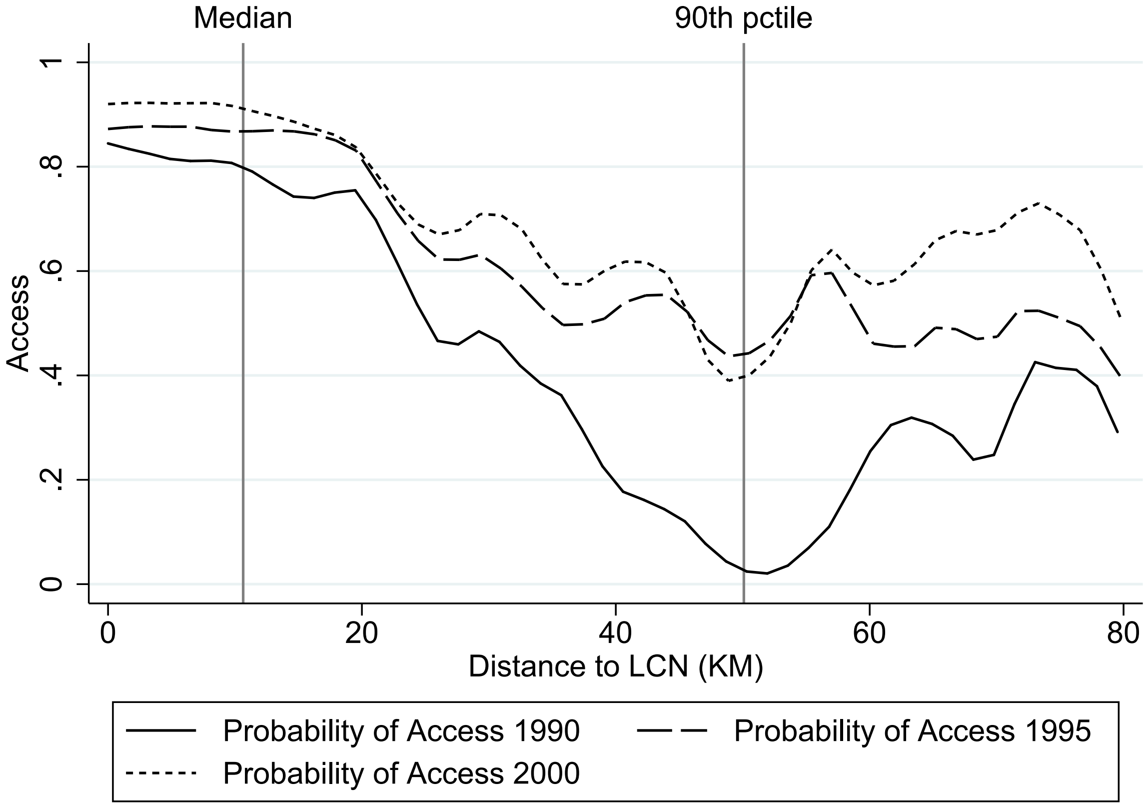
The y-axis presents the manufacturing output (Billion IDR) at the desa level as a function of the distance of that desa to the hypothetical least cost grid. This is estimated using an Epanechnikov kernel function with a bandwidth of 5.02. The x-axis shows the distance from the desa to the hypothetical least cost grid. The median and the 90^{th} percentile of the distance to the hypothetical gridare shown for reference.

Figure 11: Distance to hypothetical grid and Firm Access



The y-axis presents probability of a firm being in a desa with access to the grid, where $Access_{vt}$ is a dummy variable equal to 1 if a desa is within 15 KM of the nearest transmission substation. The probability is estimated using an Epanechnikov kernel function with a bandwidth of 2.49. The x-axis shows the distance from the desa to the hypothetical least cost grid. The median and the 90^{th} percentile of the distance to the hypothetical gridare shown for reference.

Figure 12: Distance to hypothetical grid and Firm-Level Access, by Year



The y-axis presents probability of a firm being in a desa with access to the grid for years 1990, 1995 and 2000, where $Access_{vt}$ is a dummy variable equal to 1 if a desa is within 15 KM of the nearest transmission substation. The probability is estimated using an Epanechnikov kernel function with a bandwidth of 2.49. The x-axis shows the distance from the desa to the hypothetical least cost grid. The median and the 90^{th} percentile of the distance to the hypothetical grid are shown for reference.

B Tables

Table 1: First-Stage Regressions

Sample	**Desas**	**Firms**
	(1)	(2)
	Access_{vt}	Access_{vt}
Z (KM)	**-0.00165*** **	**-0.00296*** **
	(0.000152)	**(0.000460)**
Distance to city	-0.00263***	-0.00320***
	(0.000131)	(0.000304)
Distance to coast	5.56e-05	0.00163***
	(0.000149)	(0.000455)
Elevation	-0.191***	-0.0858**
	(0.00940)	(0.0401)
Distance to road dist	-0.00410***	-0.000329
	(0.000664)	(0.000524)
Motorstation	-0.0281**	-0.00699
	(0.0136)	(0.0142)
Railway	0.0419**	0.00927
	(0.0191)	(0.0220)
Seaport	-0.0545	-0.174***
	(0.0537)	(0.0646)
Airport	0.167***	0.0203
	(0.0423)	(0.0174)
First Stage F	118.7	41.55
Observations	261,470	141,615
Year FE	✓	
Desa Controls	✓	✓
Province FE	✓	✓
YearxIndustry FE		✓
Firm Controls		✓

Notes: First stage regressions of access instrument with distance to hypothetical least cost grid. Access is defined at the desa level. A desa has $access_{vt} = 1$ if it is within 15 Km of the nearest substation. Desa controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, road type and width dummies, desa political status, and development classification, and infrastructure controls. Robust standard errors in parentheses clustered at the desa level.

Table 2: Impact of access on desa level outcomes.

Sample: Desa-Level					
	(1)	(2)	(3)	(4)	(5)
Dependent Variable	Nb of Firms	Nb of Workers in Manufacturing	Output Billion IDR	Entry Rate	Exit Rate
Panel A: OLS					
Access_{vt}	0.378***	74.64***	3.973***	0.00719***	0.00171***
	(0.0288)	(6.196)	(0.491)	(0.00263)	(0.000581)
Panel B: IV					
Access_{vt}	0.887*	513.9***	39.74***	0.106***	0.0157**
	(0.480)	(113.8)	(8.175)	(0.0284)	(0.00658)
First Stage F	118.7	118.7	118.7	58.39	58.39
Panel C: Reduced Form IV					
Z (KM)	-0.00148*	-0.856***	-0.0662***	-0.000249***	-3.68e-05**
	(0.000793)	(0.176)	(0.0125)	(6.00e-05)	(1.50e-05)
Observations	261,470	261,470	261,470	54,210	54,210
Year FE	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓
Mean Dep Var	0.84	110	6.7	0.07	0.01

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

Notes: Results from OLS, IV and reduced-form regressions of equation (1). Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Robust standard errors in parentheses clustered at the desa level.

Table 3: Impact of connection on the sales and inputs at the firm level.

Sample: Firm-Level						
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable (Log)	Sales	Capital	Wage Bill	Nb Workers	Energy Bill	Electricity (kWh)
Panel A: OLS						
Access	0.466***	0.416***	0.348***	0.197***	0.447***	0.499***
	(0.0592)	(0.0592)	(0.0422)	(0.0275)	(0.0888)	(0.0933)
Panel B: IV						
Access	2.511***	3.417***	1.788***	1.169***	4.015***	5.125***
	(0.615)	(0.648)	(0.403)	(0.266)	(0.781)	(1.256)
First Stage F	41.55	41.55	41.51	41.55	40.48	30.89
Panel C: Reduced Form IV						
Z (KM)	-0.00665***	-0.00933***	-0.00505***	-0.00336***	-0.0114***	-0.0110***
	(0.00134)	(0.00139)	(0.00102)	(0.000661)	(0.00180)	(0.00204)
Observations	141,659	141,659	141,642	141,659	139,481	120,453
IndustryxYear FE	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓

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

Notes: Results from OLS, IV, and reduced form regressions. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table 4: Electrification, exit, and the age distribution.

Sample: Firm-Level		
	(1)	(2)
Dependent Variable	Exit	Young
Panel A: OLS		
Access_{vt}	0.0077***	0.0371***
	(0.002)	(0.0148)
Panel B: IV		
Access_{vt}	0.049**	0.242**
	(0.016)	(0.099)
First Stage F	41.55	41.55
Panel C: Reduced Form IV		
Z (KM)	-0.000144***	-0.000718**
	(3.82e-05)	(0.000281)
Observations	141,615	141,615
IndustryxYear FE	✓	✓
Province FE	✓	✓
Geo Controls	✓	✓
Firm Controls	✓	✓

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

Notes: Results from OLS, IV and reduced-form regressions for a young dummy access to electricity defined at the desa level. Young is a dummy equal to 1 if the firm's age is below the median age. Robust standard errors in parentheses clustered at the desa level. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table 5: Relocation of Economic Activity Desa-Level

Sample: Desa-Level				
	(1)	(2)	(3)	(4)
Dependent Variable	No of Firms	No of Entrants	Entry Rate	Exit Rate
Panel A: OLS				
N^S_{vpt}	0.0082	-0.000199	-0.00035	-9.5e-05
	(0.0051)	(0.00033)	(0.00049)	(8.45e-05)
N_{vp}	0.0097***	0.00066***	-0.000160	-3e-05
	(0.00132)	(0.000104)	(0.000104)	(2.00e-05)
Z(KM)	-0.00019	-3.2e-06	-0.0002*	2.9e-06
	(0.000814)	(6.82e-05)	(0.000106)	(2.10e-05)
Panel B: IV				
N^S_{vpt}	-0.0177	0.00349	0.000424	-0.00114
	(0.0135)	(0.00349)	(0.00363)	(0.0008)
N_{vp}	0.0101***	0.00061***	-0.0002	-8e-06
	(0.00135)	(9.60e-05)	(0.000129)	(2.66e-05)
Z(KM)	-0.00019	-4.31e-06	-0.000203*	5e-06
	(0.00081)	(6.8e-05)	(0.0001)	(2e-05)
First Stage F	40.60	40.60	12.44	12.44
Observations	113,312	113,312	15,446	15,446
Year FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓
Mean Dep Var	0.39	0.03	0.08	0.006
Mean N^S_{vpt}	0.38	0.38	0.53	0.53
Mean N_{vt}	35	35	42	42

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

Notes: Results from OLS, and IV regressions of equation (3). Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Robust standard errors in parentheses clustered at the desa level.

Table 6: Access and spillover effects at the desa-level.

Sample: Desa-Level					
	(1)	(2)	(3)	(4)	(5)
Dependent Variable	Nb of Firms	Nb of Workers in Manufacturing	Output Billion IDR	Entry Rate	Exit Rate
Panel A: OLS					
Access_{vt}	0.234***	21.03	2.816**	0.00959**	-0.00033
	(0.0577)	(13.46)	(1.170)	(0.00395)	(0.0009)
N^C_{vpt}	0.0014	1.607***	0.049	-3.8e-05	4.9e-05***
	(0.00161)	(0.389)	(0.0357)	(7.6e-05)	(1.8e-05)
N_{vp}	0.00804***	-0.621***	-0.0620***	-8.2e-05	-8.4e-06
	(0.00134)	(0.190)	(0.0125)	(7.6e-05)	(1.7e-05)
Panel B: IV					
Access_{vt}	2.001**	545.3**	100.7***	0.152***	0.031**
	(0.886)	(222.5)	(19.68)	(0.053)	(0.0148)
N^C_{vpt}	-0.0318	-6.407	-2.916***	-0.002**	-0.0007***
	(0.0249)	(5.775)	(0.555)	(0.001)	(0.0003)
N_{vp}	0.0277	3.617	1.998***	0.00127*	0.0006***
	(0.0172)	(3.982)	(0.389)	(0.00076)	(0.0002)
First Stage F	39.63	39.63	39.63	5.078	5.078
Observations	261,470	261,470	261,470	54,210	54,210
Year FE	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓
Mean Dep Var	0.84	110	6.7	0.07	0.01
Mean N^C_{vpt}	27.8	27.8	27.8	39.6	39.6
Mean N_{vt}	43	43	43	52	52
P-value of joint effect	0.0016	0.00	0.011	0.0013	0.17

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

Notes: Results from OLS and IV regressions of equation (4). Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Robust standard errors in parentheses clustered at the desa level. The p-value in the last row corresponds to the null of H0: $\hat{\beta} + 27.8 * \hat{\mu} = 0$.

Table 7: Impact of access on district level outcomes.

Sample: District-Level					
	(1)	(2)	(3)	(4)	(5)
Dependent Variable	Nb of Firms	Nb of Workers in Manufacturing	Output Billion IDR	Entry Rate	Exit Rate
Panel A: OLS					
$Access_{dt}$	0.447***	3.312***	85.95***	0.00738*	0.00254***
	(0.0716)	(0.818)	(13.51)	(0.00395)	(0.000756)
Panel B: IV					
$Access_{dt}$	1.616*	39.05***	617.5***	0.101***	0.0143*
	(0.846)	(13.43)	(229.6)	(0.0367)	(0.00737)
First Stage F	20.12	20.12	20.12	19.73	19.73
Observations	17,941	17,941	17,941	13,407	13,407
Year FE	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓
Mean Dep Var	1.08	153	8.9	0.072	0.009

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

Notes: Results from OLS, and IV regressions of equation (1) at the district level. Geographic controls are defined at the district level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Robust standard errors in parentheses clustered at the district level. Access is defined as a dummy equal to 1 if at least 50% of desas in the district are within 15Km of the closest substation.

Table 8: Effect of electrification on sales of nontradables

	(1)
Dependent Variable	Sales
Access	2.277**
	(0.907)
First Stage F	12.80
Observations	11,462
IndustryxYear FE	✓
Province FE	✓
Geo Controls	✓
Firm Controls	✓

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

Notes: Results from OLS and IV regressions. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table 9: Testing For Spillovers

	(1)	(2)
	OLS	IV
Dependent Variable (Log)	Deflated Sales	Deflated Sales
Nb switching neighbors	0.149***	-0.108
	(0.0234)	(0.153)
Z (KM)	-0.00776***	-0.00668***
	(0.00129)	(0.00156)
Observations	141,615	141,420
First Stage F		45.02
IndustryxYear FE	✓	✓
Province FE	✓	✓
Geo Controls	✓	✓
Firm Controls	✓	✓

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

Notes: Results from OLS and IV regressions. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table 10: Testing For Spillovers within a 5-digit industry

Dependent Variable	Log Sales			
	(1)	(2)	(3)	(4)
Industry	ALL	non-tradables	Food and Bev	Textiles
Number of Switching Competitors	0.00181	-0.000866	-0.0234	-0.000858
	(0.00542)	(0.00346)	(0.0253)	(0.0509)
First Stage F	86.47	124.5	91.83	50.64
Observations	113,115	10,861	24,329	15,317
Mean RHS	10.1	20.5	6.2	5.8
	(5)	(6)	(7)	(8)
	Apparel and Footwear	Furniture	Rubber and plastic	All
	-0.0164	-0.0305	-0.0270	
	(0.0102)	(0.0443)	(0.0412)	
Access				2.057***
				(0.497)
Observations	16,058	10,836	6,887	113,115
First Stage F	340.5	45.80	477.2	35.11
Mean RHS	11.8	11.9	2.5	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Results from IV regressions. The dependent variable is log sales. The first Column shows the regression of the whole sample of firms. The RHS variable is the number of switching competitors. A switching competitor is a firm in the same 5-digit industry that switches from being without access to having access to the grid. Columns (2) - (7) shows the same regression for each of the top 6 largest industries separately. Column (8) presents the effect of access on sales of all firms in the 6 largest industries. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table 11: Effect of electrification on TFPR by Age Group.

Sample: Firm-Level			
	All	**Young**	**Old**
	(1)	(2)	(3)
Dependent Variable	log(TFPR)	log(TFPR)	log(TFPR)
	Panel A: OLS		
Access_{vt}	0.0184*	0.0179	0.0169
	(0.0100)	(0.0148)	(0.0105)
	Panel B: IV		
Access_{vt}	0.177**	0.369***	0.060
	(0.089)	(0.003)	(0.096)
First Stage F	43.76	36.81	33.08
	Panel C: Reduced Form IV		
Z (KM)	-0.000486**	-0.0010***	-0.00016
	(0.000236)	(0.00032)	(0.00025)
Observations	134,391	47,921	86,439
IndustryxYear FE	✓	✓	✓
Province FE	✓	✓	✓
Geo Controls	✓	✓	✓
Firm Controls	✓	✓	✓

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

Notes: Results from OLS and IV and Reduced-Form regressions of TFPR on access defined at the desa level. TFPR is measured following Olley and Pakes (1996). Robust standard errors in parentheses clustered at the desa level. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table 12: Impact of connection on Price and ϕ_{it} by Age Group.

	All		Young		Old	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	log(Price)	ϕ_{it}	log(Price)	ϕ_{it}	log(Price)	ϕ_{it}
			Panel A: OLS			
$Access_{vt}$	-0.0125	0.108***	-0.0191	0.208***	-0.0129	0.0633
	(0.0261)	(0.0340)	(0.0414)	(0.0532)	(0.0291)	(0.0388)
Observations	127,427	127,427	40,406	40,406	86,226	86,226
			Panel B: IV			
$Access_{vt}$	-0.375	0.932***	0.0845	0.931*	-0.576*	0.804*
	(0.245)	(0.355)	(0.397)	(0.532)	(0.319)	(0.427)
Observations	127,427	127,427	40,406	40,406	86,226	86,226
First Stage F	25.23	25.23	17	17	16.27	16.27
			Panel C: Reduced Form IV			
Z (KM)	0.000803	-0.00199***	-0.000193	-0.00213*	0.00109**	-0.00152**
	(0.000500)	(0.000678)	(0.000901)	(0.00120)	(0.000521)	(0.000713)
Observations	127,427	127,427	40,406	40,406	86,226	86,226
IndustryxYear FE	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓

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

Notes: Results from OLS and IV and Reduced-Form regressions of two different measures of TFPR on access defined at the desa level. Robust standard errors in parentheses clustered at the desa level. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table 13: Olley-Pakes Revenue Weighted Productivity Decomposition

	Sample: Sector-Level		
	(1)	(2)	(3)
Dependent Variable	Weighted Average	Unweighted Average	Covariance
	log(TFPR$_{OP}$)	log(TFPR$_{OP}$)	(log(TFPR$_{OP}$), *share*)
	Panel A: OLS		
Access$_{vt}$	0.140***	0.0114	0.121***
	(0.0249)	(0.0123)	(0.0197)
	Panel B: IV		
Access$_{vt}$	0.550***	0.261***	0.278**
	(0.163)	(0.0945)	(0.109)
First Stage F	36	36	36
	Panel C: Reduced Form IV		
Z (KM)	-0.00213***	-0.000998***	-0.00106***
	(0.000549)	(0.000292)	(0.000410)
Observations	9,899	9,899	9,899
Industry FE	✓	✓	✓
ProvincexYear FE	✓	✓	✓

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

Robust standard errors in parentheses clustered at the sector level

C Additional Results

Table C1: Impact of electrification on desa level industrial outcomes - Log transformations.

	Sample: Desa-Level					
	(1)	(2)	(3)	(4)	(5)	(6)
	Log(1+Nb Firms)	Log(h(Nb Firms)) in Manufacturing	Log(1+Nb Workers) in Manufacturing	Log(h(Nb Workers)) Billion IDR	Log(1+Output) Billion IDR	Log(h(Output))
			Panel A: OLS			
Access$_{vt}$	0.103***	0.131***	0.329***	0.368***	0.126***	0.149***
	(0.00553)	(0.00699)	(0.0171)	(0.0193)	(0.00611)	(0.00725)
			Panel B: IV			
Access$_{vt}$	0.210**	0.266**	0.918***	0.961***	0.774***	0.906***
	(0.0983)	(0.125)	(0.307)	(0.346)	(0.125)	(0.147)
First Stage F	120.4	120.4	120.4	120.4	120.4	120.4
			Panel C: Reduced Form IV			
Z (KM)	-0.000350**	-0.000443**	-0.00153***	-0.00160***	-0.00129***	-0.00151***
	(0.000162)	(0.000206)	(0.000500)	(0.000566)	(0.000181)	(0.000214)
Observations	261,470	261,470	261,470	261,470	261,470	261,470
Year FE	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓

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

Notes: Results from OLS and IV and Reduced-Form regressions of two different measures of log transformations that preserve zeros for the number of firms, number of workers in manufacturing and total manufacturing output. The first transformation is a $log(1+X)$. The second transformation is $log(h(X))$ where $h(X) = X + (X^2+1)^{\frac{1}{2}}$ following Liu and Qiu (2016). Robust standard errors in parentheses clustered at the desa level. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table C2 shows how connection to the grid affects firm-level input ratios. As in Table 3, the OLS estimates in Panel A are positive but smaller in magnitude relative to the IV estimates in Panel B. Column (1) Panel B shows connection causes the capital-labor ratio of the firm to increase. From Columns (2) and (3), both the energy-capital and energy-labor ratios increase, but the second increases by twice as much. This explains the increase in the capital-labor ratio. In other words, conditional on capital, there is no effect of connection on labor. All these results depict a particular input substitution pattern where capital and energy are complimentary and labor and energy are more substitutable (or at least, there is less substitution between capital and energy than labor and energy).

Table C2: Electrification and the firm's input ratios.

	(1)	(2)	(3)
Dependent Variable	log(K/L)	log(E/K)	log(E/L)
	Panel A: OLS		
Connected	0.328***	0.0512	0.316***
	(0.0361)	(0.0635)	(0.0587)
Observations	141,642	139,481	139,468
	Panel B: IV		
Connected	2.135***	1.621**	3.767***
	(0.562)	(0.710)	(0.832)
Observations	141,642	139,481	139,468
First Stage F	26.80	20.93	20.97
IndustryxYear FE	✓	✓	✓
Province FE	✓	✓	✓
Geo Controls	✓	✓	✓
Firm Controls	✓	✓	✓

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

Notes: Results from OLS and IV regressions. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

C.1 Production Function

Table C3: Average Output Elasticities

Sector	N	nrobs	Capital	Labor	Energy
15 Food and Beverages	29555	12520	0.03	0.40	0.28
			(0.04)	(0.07)	(0.09)
16 Tobacco Products	4197	3435	0.00	0.69	0.26
			(0.02)	(0.19)	(0.13)
17 Textiles	15517	3796	0.06	0.37	0.29
			(0.09)	(0.14)	(0.15)
18 Wearing Apparel , Fur	14614	3581	0.02	0.41	0.10
			(0.04)	(0.19)	(0.04)
19 Leather, leather products and footwear	5036	1691	0.02	0.67	0.27
			(0.06)	(0.22)	(0.13)
20 Wood Products (excl. furniture)	7128	2312	0.07	0.15	0.11
			(0.06)	(0.18)	(0.06)
21 Paper and paper products	2584	1013	0.17	0.20	0.32
			(0.14)	(0.49)	(0.22)
22 Printing and Publishing	3846	740	-0.02	0.52	0.41
			(0.12)	(0.13)	(0.24)
23 Coke, refine petroleum products, nuclear fuel	260	140	0.07	0.44	0.69
			(0.53)	(1.03)	(0.30)
24 Chemicals and chemical products	9386	1761	0.04	0.53	0.27
			(0.06)	(0.34)	(0.11)
25 Rubber and plastic products	9312	3226	0.04	0.20	0.29
			(0.07)	(0.13)	(0.12)
26 Non-metallic mineral products	14797	3290	0.02	0.34	0.31
			(0.06)	(0.12)	(0.14)
27 Basic metals	1065	503	0.11	0.20	0.31
			(0.20)	(0.27)	(0.19)
28 Fabricated metal products	5829	2198	0.02	0.24	0.20
			(0.05)	(0.06)	(0.06)
29 Machinery and equipment n.e.c.	3410	1158	-0.04	0.51	0.24
			(0.18)	(0.21)	(0.09)
31 Electrical Machinery and apparatus	1095	633	0.04	-0.00	0.32
			(0.34)	(0.55)	(0.31)
32 Radio, television and communication equipment	498	336	0.13	0.47	0.33
			(0.13)	(0.54)	(0.21)
33 Medical, precision and optical instruments	457	310	0.10	0.03	0.32
			(0.23)	(0.39)	(0.19)
34 Motor vehicles, trailers, semi-trailers	934	691	0.04	-0.02	0.47
			(0.15)	(0.38)	(0.28)
35 Other Transport Equipment	1693	790	0.03	0.35	0.40
			(0.19)	(0.30)	(0.19)
36 Furniture, manufacturing n.e.c	11543	3393	-0.07	0.55	0.06
			(0.07)	(0.09)	(0.02)
Average	14142.96	4762.09	0.03	0.39	0.24

Table C4: Production Function Coefficients

Sector	β_k	β_k	β_e	β_{kk}	β_{ll}	β_{ee}	β_{lk}	β_{ke}	β_{le}	β_{lek}
15	0.280	0.632	0.399	-0.014	0.016	0.035	-0.004	-0.009	-0.088	0.002
16	-0.082	0.012	0.291	-0.005	0.062	0.035	0.023	0.017	-0.062	-0.002
17	0.750	0.465	0.213	-0.004	0.067	0.039	-0.071	0.006	-0.088	0.001
18	-0.737	-1.564	-0.770	0.003	0.068	0.016	0.071	0.084	0.067	-0.009
19	0.658	1.002	1.290	0.008	0.118	0.049	-0.118	-0.006	-0.228	0.005
20	-1.086	-2.216	-1.821	-0.001	0.063	-0.013	0.093	0.177	0.176	-0.014
21	2.401	3.239	2.779	-0.019	0.105	0.087	-0.218	-0.108	-0.458	0.015
22	1.993	2.931	3.994	0.004	-0.008	-0.049	-0.179	-0.223	-0.228	0.018
23	6.391	13.630	10.023	0.146	0.188	-0.010	-1.074	-0.461	-1.005	0.055
24	-0.048	2.323	0.904	0.011	-0.019	0.045	-0.022	-0.017	-0.140	0.002
25	-1.660	-1.374	-2.248	-0.013	-0.034	0.050	0.202	0.144	0.163	-0.015
26	-0.193	0.583	0.078	0.005	0.028	0.048	-0.005	0.016	-0.084	0.000
27	-2.752	-4.106	-6.890	-0.013	0.076	0.064	0.202	0.413	0.368	-0.028
28	0.121	0.290	0.497	0.008	0.014	0.022	-0.015	-0.040	-0.048	0.002
29	-0.538	1.568	0.870	-0.017	-0.071	-0.017	0.061	-0.001	-0.064	0.003
31	2.556	2.693	2.254	0.019	0.188	-0.097	-0.463	0.039	-0.294	0.016
32	-7.002	-3.840	-11.889	-0.007	-0.145	0.115	0.568	0.664	0.735	-0.051
33	-4.390	-5.584	-5.044	0.075	0.207	0.068	0.156	0.403	0.269	-0.029
34	1.090	-0.862	4.668	0.049	0.310	0.118	-0.166	-0.160	-0.528	0.010
35	2.002	2.190	4.505	0.050	0.147	0.094	-0.270	-0.306	-0.497	0.024
36	-0.893	-0.373	-0.523	0.005	0.001	-0.009	0.084	0.060	0.081	-0.007
37	-0.926	-1.149	-0.915	0.000	-0.025	-0.002	0.117	0.054	0.133	-0.008
Average	-0.082	0.228	0.076	-0.001	0.033	0.029	0.005	0.023	-0.048	-0.001

Table C5: Impact of connection on the marginal cost at the product level.

Dependent Variable	Log Marginal Cost				
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) RF
Connected_{it}	0.00321	-1.310***			
	(0.0336)	(0.495)			
Access_{vt}			-0.0578*	-0.872***	
			(0.0335)	(0.332)	
Z (KM)					0.00187***
					(0.000631)
Observations	127,427	127,427	127,427	127,427	127,427
First Stage F		27.60		25.23	
Product FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Results from OLS, IV and RF (reduced form) regressions. Connected is an firm-level dummy variable equal to one if the firm consumes a positive amount of grid electricity. Access is a desa-level dummy equal to one if the desa is within 15km of the closest transmission substation. Z is the distance to the hypothetical least cost grid in km. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, desa political and legal status, and infrastructure controls. Firm Controls include export, cohort, and ownership dummies. Robust standard errors in parentheses clustered at the desa level.

Table C6: Impact of connection on the product-level outcomes.

	(1)	(2)	(3)	(4)
Dependent Variable (Log)	Sales	Volume	Price	Markup
Panel A: OLS				
Connected	0.210***	0.155***	0.0653***	0.0621**
	(0.0381)	(0.0431)	(0.0199)	(0.0255)
Observations	127,427	127,427	127,427	127,427
Panel B: IV				
Connected	1.551**	2.150**	-0.564	0.747*
	(0.774)	(0.961)	(0.377)	(0.399)
Observations	127,427	127,427	127,427	127,427
First Stage F	27.60	27.60	27.60	27.60
Panel C: Reduced Form IV				
Z (KM)	0.00221**	-0.00306**	0.000803	-0.00106*
	(0.00101)	(0.00122)	(0.000500)	(0.000564)
Observations	127,427	127,427	127,427	127,427
Product FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓

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

Notes: Results from OLS and IV regressions. Robust standard errors in parentheses clustered at the desa level. Geographic controls are defined at the desa level and include distance to coast, distance to nearest city, elevation, distance to road, road type and width dummies, desa political status, development classification and an urban dummy. Firm controls include export and ownership dummies, age and a generator dummy.

D Model

In this section, I present a model of a monopolistically competitive industry to illustrate the effects of the grid expansion on the manufacturing sector. The goal is to analyze selection, allowing for competition effects. As the grid reaches more areas, the entry decision of firms in these areas will be affected through a reduction in the fixed cost of entry. In addition, as more firms in the market are getting connected, and thus becoming more efficient, this will affect the survival of incumbents (and expected value of entry) as a higher proportion of more efficient firms in the market means more intense competition.

D.1 Demand

Consider an industry with a continuum of firms of measure N, each indexed by i. Firm i produces a differentiated variety in the market. Consumers have utility U defined over these differentiated varieties indexed by i in set I and a Hicksian composite commodity:

$$U = H + \int_{i \in I} \alpha q_i di - \frac{1}{2}\eta\Big(\int_{i \in I} q_i di\Big)^2 - \frac{1}{2}\gamma \int_{i \in I} {q_i}^2 di \tag{15}$$

where H is the consumption of the Hicksian composite good and q_i is the consumption of variety i. The demand parameter $\eta \in (0,1)$ represents the degree of substitutability between different varieties. Utility maximization implies the following demand function:

$$q_i = \frac{\alpha}{\eta N + \gamma} + \frac{\eta N}{\gamma(\eta N + \gamma)}\bar{p} - \frac{1}{\gamma} p_i \tag{16}$$

where $\bar{p} \equiv \frac{1}{N}\int_{i \in I} p_i$ is the average price in the market conditional on survival. Define p^{max} as the highest price consumers are willing to pay which can calculated from setting demand in equation (16) to zero:

$$p^{max} = \frac{\gamma\alpha}{\eta N + \gamma} + \frac{\eta N}{\eta N + \gamma}\bar{p} \tag{17}$$

The residual demand for product i from (16) can therefore be written as:

$$q_i = \frac{1}{\gamma}(p^{max} - p_i) \tag{18}$$

D.2 Production

On the production side, consider a single input technology[30] where firm i produces according the the following production function:

$$q_i = \phi_i x_i \tag{19}$$

[30] The assumption of a single input production process is without loss of generality when considering a multiple input production function with constant returns to scale.

where ϕ_i is the firm's physical productivity and x_i is the input of production which is supplied inelastically at a constant[31] price w. Therefore, firm i's marginal cost is $c_i = \frac{w}{\phi_i}$. Combined with the demand form, the profit maximizing price is:

$$p(c_i) = \frac{1}{2}(p^{max} + c_i) \tag{20}$$

The equilibrium profit is:

$$\pi(c_i) = \frac{1}{4\gamma}(p^{max} - c_i)^2 \tag{21}$$

Firm i will stay in the market as long as $\pi(w, \phi_i) \geq 0$. This gives the cut-off level of marginal cost c^* such that the firm will not want to stay in the market if its marginal cost exceeds it:

$$c^* = p^{max} = \frac{\gamma\alpha}{\eta N + \gamma} + \frac{\eta N}{\eta N + \gamma}\bar{p} \tag{22}$$

Firm price, mark-up and quantity can therefore be written as:

$$p(c_i) = \frac{1}{2}(c^* + c_i) \tag{23}$$

$$\mu(c_i) = \frac{1}{2}(c^* - c_i) \tag{24}$$

$$q(c_i) = \frac{1}{2\gamma}(c^* - c_i) \tag{25}$$

Firm price is increasing in its own marginal cost, but more efficient firms charge relatively higher markups and produced relatively more. The more efficient the marginal firm is (lower c^*), the tougher competition is, reducing firm prices, markups and quantity demanded, conditional of the firm's own marginal cost. The cutoff c^* then implies implies a cutoff level for firm productivity:

$$\phi^* = \frac{w}{c^*} \tag{26}$$

Firms with productivity $\phi_i < \phi^*$ will not be profitable and will exit the market. Therefore, $p^{max} = \frac{w}{\phi^*}$.

D.3 Long Run Equilibrium

In the long run, a large number of ex-ante identical potential firms decide whether to enter the market. Before observing their productivity, potential entrants have to pay a sunk cost of entry s. They then receive a productivity draw from a distribution $G(\phi)$ with support $[\underline{\phi}, \infty]$. In equilibrium, the expected value of entry should be equal to zero for positive entry to occur:

$$V^e = \frac{w^2}{4\gamma}\int_{\phi^*}^{\infty}\left(\frac{1}{\phi^*} - \frac{1}{\phi}\right)^2 dG(\phi) - s = 0 \tag{27}$$

Equation (27) pins down ϕ^* which summarizes the equilibrium. The equilibrium mass of firms N is determined using equations (20) and (22).

[31] This simple representation is meant to capture that although firms are heterogeneous in their productivity they face the same price of electricity which is set by the state, either directly (price per kWh or price of fuel). This is true in the case of Indonesia where the energy sector is heavily regulated and the price is the same everywhere in the country.

D.4 Predictions

The goal of this exercise is to see how the equilibrium cut-off changes with access to electricity. This can be studied through comparative statics with respect to two parameters. The first is the input price w. The conjecture is that access to the grid reduces the per-unit price of electricity. The second is the sunk cost of entry s. Entry to a location where the grid hasn't arrived is potentially more expensive as the firm will need to purchase its own generator. Starting with comparative statics with respect to w, and using the implicit function theorem:

$$\frac{d\phi^*}{dw} = -\frac{\partial V^e/\partial w}{\partial V^e/\partial \phi^*} > 0 \tag{28}$$

since $\partial V^e/\partial \phi^* < 0$ and $\partial V^e/\partial w > 0$. Therefore, a decrease in w will lead to a lower productivity cut-off. Intuitively, as the input price is lower, a firm that wasn't able to survive before will be able to do so now. As for the sunk cost of entry, the cutoff ϕ^* is decreasing in s since the derivative of the value function with respect to s is -1:

$$\frac{d\phi^*}{ds} = -\frac{\partial V^e/\partial s}{\partial V^e/\partial \phi^*} < 0 \tag{29}$$

This says that if access to electricity reduces the sunk cost of entry, then this will increase the productivity of firms in the industry. The intuition is as follows. If access to electricity lowers barriers to entry, more firms will enter the market. This intensifies competitive pressure and makes it more difficult for relatively unproductive firms to survive in equilibrium.

In order to understand how average industry and firm outcomes could be affected by electrification, it is useful to focus the analysis on changes in the marginal cost cutoff c^*. This is because although the effect of access on ϕ^* is interesting, what ultimately determines the equilibrium outcomes is a combination of input prices and firm productivity, i.e. the marginal cost of the firm. Revisiting the comparative statics with respect to input price w and sunk cost of entry s gives the following predictions. The effect of a decrease in w on c^* is ambiguous. Although ϕ^* increases with a decrease in w, this doesn't necessarily mean that the marginal cost of the marginal firm c^* is lower. The overall effect depends on the relative effects of the decrease in w and increase in ϕ^*. As for the sunk cost of entry, conditional on w, a decrease in s unambiguously leads to a decrease c^*.

Define the average marginal cost of surviving firms $\bar{c} = \frac{1}{1-G(\phi^*)} \int_{\phi^*}^{\infty} \frac{w}{\phi} dG(\phi)$. Given a distribution of productivity $G(.)$, the averages of firm outcomes in equations (23)-(25) conditional of survival are:

$$\bar{p} = \frac{1}{2}(c^* + \bar{c}) \tag{30}$$

$$\bar{\mu} = \frac{1}{2}(c^* - \bar{c}) \tag{31}$$

$$\bar{q} = \frac{1}{2\gamma}(c^* - \bar{c}) \tag{32}$$

where $\bar{z} = \frac{1}{1-G(\phi^*)} \int_{\phi^*}^{\infty} z(\phi) dG(\phi)$. Intuitively, $\bar{c}$ is increasing in c^*. If the marginal firm is more efficient (lower c^*), then the average firm efficiency in the industry is higher (lower $\bar{c}$). Equation (30) predicts that the average observed prices conditional of firm survival is lower when c^* is lower. Equations (31) and (32) however give an ambiguous prediction on a change in c^* on average markups and quantities. On the one hand, a lower c^* means tougher competition in the market, reducing firm markups and quantities produced. However, tougher selection also means that the set of surviving firms are more efficient (lower $\bar{c}$), and as seen from equations (24) and (25), more efficient firms charge relatively higher markups and produced more. Which effects dominates depends on the distribution of productivity $G(.)$ and its support.

Recall that in equilibrium, the zero profit condition states that the profit of the marginal firm should be equal to zero. This condition requires that $c^* = p^{max}$:

$$c^* = \bar{p} + \frac{\gamma(\alpha - \bar{p})}{\eta N + \gamma} \tag{33}$$

The equilibrium mass of active firms as a function of c^* is therefore:

$$N = \frac{2\gamma(\alpha - c^*)}{\eta(c^* - \bar{c})} \tag{34}$$

These equations state that tougher competition (lower c^*) is associated with a higher mass of active firms N and a lower average price[32] $\bar{p}$. To see this[33], suppose N increases, and that surviving firms don't change their prices following entry, keeping $\bar{p}$ constant. From equation (33), c^* will decrease. From equation (30), $\bar{p}$ will decrease as result, which further decreases c^*. In addition, the model predicts that firm exit rates unambiguously increase when the marginal cost cutoff c^* is lower. The probability of survival, which is equal to $\tilde{G}(c^*) = 1 - G(\frac{w}{\phi^*})$, is decreasing in c^*. Intuitively, tougher competition is associated with tougher selection where conditional on its own efficiency, a firm's probability of survival is lower.

The relationship between access to electricity and firm-level and industry-level outcomes can be interpreted through the lens of the model. The averages of firm outcomes in (30)-(32) correspond to the respective observed firm outcomes in the data. Also, if access to the grid reduces fixed cost of entry, the model predicts that access will lead to tougher selection in the market induced by entry of a larger number of firms. In addition, the model predicts that higher exit rates are associated with tougher selection and a higher efficiency cutoff. Finally, equations (28) and (29) state that average physical productivity ϕ increases if barriers to entry are lower, but decreases in response to an increase in the input price. This sharp prediction is informative regarding the channels through which access to electricity is affecting the manufacturing sector.

D.5 Limitations of the Model

This model is very simple and abstracts from many features that could be important in determining the effect of electrification on industry productivity.

- <u>Trade</u>: I assume that each location is a separate market and that firms don't sell in other locations. This is obviously an unrealistic assumption as these firms are

[32] An implicit assumption here is that $\alpha > c^*$ which implies that α is greater than $\bar{p}$ and $\bar{c}$.
[33] The intuition is the same as in Combes, Duranton, Gobillon, Puga, and Roux (2012).

medium and large manufacturing firms and the desas are too small to constitute their whole market. The model can be extended to allow for trade across location as in Melitz Ottaviano (2008) and the comparative statics with respect to sunk cost of entry and input price in the location's own cutoff all go through. Therefore, we can still learn something from the simple closed economy model about the effect of electrification on productivity at the location level.

- Spillovers: Given that the true model involves trade across different locations and since most firms in my data produce tradable goods, the presence of spillovers across different locations complicates the interpretation of my results. Electrifying one location can have an effect on firms in other locations, and these effects are likely to be negative. What I estimate as the average difference between electrified and non-electrified locations could be therefore a combination of creation of new economic activity and relocation of economic activity from those who don't get electrified (or are already electrified) to locations that get newly electrified. An important question is whether there is any creation of new economic activity in response to electrification, or does electrification only displace economic activity? I address this question in the empirical section where I test for the presence spillovers. Theoretically, the size of the spillovers depend the substitutability of the products being traded, transportation costs, and the number of trading partners. If transportation costs are very large, then spillovers will be minimal. Spillovers can also be minimal if there is a very large number of markets: the general equilibrium effects will be small because each market is too small to affect other markets. This can be shown in a model such as Melitz Ottaviano (2008) where the number of markets (countries in their context) is very large.

E Demand Forecasts

E.1 Methodology Overview: DKL

The model combines multiple methods; mainly trend projections and what PLN calls econometrics (calculating elasticities). PLN conducts its forecast at the sectoral level before aggregating at the regional level. In the case of Java, the forecast is aggregated at the system level. PLN considers four sectors: Residential, Commercial, Public and Industrial. For each of these sectors, energy consumption is forecasted as a function of historical PLN data, macroeconomic variables, and elasticities of energy sales in that sector with respect to economic growth.

E.2 Residential Sector

- Energy Consumed: $E_t^R = E_{t-1}^R * (1 + \epsilon_t^R * g_t) + \Delta Nb_t^R * UK_t^R$ where:
  - ϵ_t^R is the elasticity of residential energy sales (kWh) with respect to regional GDP growth. Elasticities are optained using the "econometrics" model where they calculate the elasticity either by using actual yearly data or by regressing log sales on log gdp.
  - g_t is the regional GDP growth rate. This is either taken from BPS the Indonesian Statistics Bureau or projected linearly.
  - ΔNb_t^R is the change in the number of residential customers between year t and year $t-1$. For future years, it is the change in the FORECASTED number of customers between two years. The number of customers is projected linearly using customer factor (the equivalent of elasticity) and population growth rates where CF_t^R is calculated as the elasticity of the number of customers with respect to economic growth[34].
  - $Nb_t^R = Nb_{t-1}^R * (1 + CF_t^R * g_t)$
  - UK_t^R is energy consumption per customer (kWh/hh). The customer is one household.
- In order to forecast electrification ratios, the future number of households in the economy is forecasted using population forecasts and average number of individuals per household and then used with the forecasted number of customers to calculated the implied electrification ratio.

E.3 Commercial, Industrial and Public Sectors

Similarly, for each sector i, the goal is to get an estimate of energy consumption. This is done as follows: the number of customers is calculated/projected:

- Energy Consumed: $E_t^i = E_{t-1}^i * (1 + \epsilon_t^i * g_t)$ where:

[34] PLN seems to assume elasticities if the calculated ones are unreasonable.

- ϵ_t^i is the elasticity of energy sales (kWh) in sector i with respect to regional GDP growth.
- g_t is the regional GDP growth rate.

- In order to forecast power contracted, average power (VA) per customer is multiplied by the number of new customers in sector, then it is added to the previous year's power contracted:
- $PC_t^i = PC_{t-1}^i + \Delta Nb_t^i * UK$
- ΔNb_t^i the change in the number of customers between year t and year $t-1$ in sector i
- $Nb_t^R = Nb_{t-1}^R * (1 + CF_t^R * g_t)$
- UK_t^i is energy consumption per customer (kWh/hh) which is the average from historical data.

E.4 Forecasted TOTAL Demand and Load Factor

- Total Energy Sales (GWh): $ES_t = E_t^R + E_t^C + E_t^P + E_t^I$
- Forecasted energy sales represent the energy needs of PLN customers
- Required energy production (GWh) needs to take account of inefficiencies such as transmission and distribution losses (L%) and station use(SU%):
  $P_t = \frac{ES_t}{(1-L-SU)}$
- The Final form of demand forecast is called **Peak Load**(MW). To calculate that from required production, the load factor is needed:
  $LF_t = 0.605 * \frac{E_t^R}{ES_t} + 0.7 * \frac{E_t^C + E_t^P}{ES_t} + 0.9 * \frac{E_t^I}{ES_t} < 1$
- Finally, the peak load of the system, which is the goal of this procedure, is:

$$PL_t = \frac{P_t}{365 * 24 * LF_t * 1000}$$

E.5 Disaggregation

Because the Java-Bali system is interconnected, forecast is done at the system level. The next step is to disaggregate this forecast at the substation level. The way this is done is by looking at the proportion of the load borne by each substation out of the whole system load, and assuming that in the future these proportions will be the same. Then divide the forecasted load according to each substation's proportion. Once the load forecast is calculated for each substation, it is then compared to the capacity of each substation. If the load is greater than 80% of the capacity, then the substation should be extended or a new substation is commissioned.

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Subscribe to our newsletter
and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE

International
Growth Centre

DIRECTED BY

FUNDED BY

Designed by soapbox.co.uk