American Economic Journal: Applied Economics 2013, 5(2): 200–231
http://dx.doi.org/10.1257/app.5.2.200

Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil†

By Molly Lipscomb, A. Mushfiq Mobarak, and Tania Barham*

We estimate the development effects of electrification across Brazil over the period 1960–2000. We simulate a time series of hypothetical electricity grids for Brazil for the period 1960–2000 that show how the grid would have evolved had infrastructure investments been made based solely on geography-based cost considerations. Using the model as an instrument, we document large positive effects of electrification on development that are underestimated when one fails to account for endogenous targeting. Broad-based improvement in labor productivity across sectors and regions rather than general equilibrium re-sorting appears to be the likely mechanism by which these development gains are realized. (*JEL* H54, L94, O11, O13, Q41, Q43)

Construction of large-scale infrastructure projects was a popular use of development funds until the 1970s, but this was replaced by a trend toward smaller programs in health and education in the 1980s and 1990s. There is now renewed support for large infrastructure projects as a means of poverty reduction (The World Bank 2003; Ali and Pernia 2003). Despite the renewed investment, there is relatively little causal evidence of the effects of large infrastructure investment in general,[1,2]

* Lipscomb: University of Virginia, PO Box 400893, Charlottesville, VA 22904 (e-mail: molly.lipscomb@virginia.edu); Mobarak: Yale University, School of Management, 135 Prospect Street, P.O. Box 208200, New Haven, CT 06520-8200 (e-mail: ahmed.mobarak@yale.edu); Barham: University of Colorado, Boulder, 483 UCB, Boulder CO 80309 (e-mail: tania.barham@colorado.edu). We thank the University of Colorado NICHD Population Center, Corporación Andina de Fomento, Center for Advancement in Research and Teaching in the Social Sciences at the University of Colorado, the National Science Foundation, and the Macmillan Center at Yale University for the financial support that made this data collection possible. We thank Jorge Nunes at CHESF and many other electricity management professionals at CHESF, Eletrobras, and ANEEL for their help in data collection. We also thank Daniel Ortega, Taryn Dinkelman, Steven Puller, Rohini Pande, Arik Levinson, Erin Mansur, Sheila Olmstead, Judy Chevalier, Bill Evans, Hilary Hoynes, three anonymous referees, and seminar participants at NBER Environmental and Energy Economics Summer Institute, Corporación Andina de Fomento, UC-Berkeley Energy Institute, International Growth Centre at LSE, IDEI-World Bank International Conference on Infrastructure Economics, UCL/LSE, University of Virginia, Harvard University, Cornell University, Inter-American Development Bank, Center for Global Development, Fudan University, Yale University (Economics), and Yale School of Management for comments, and Vanessa Empinotti and Steven Li for excellent research assistance.

† To comment on this article in the online discussion forum, or to view additional materials including the author disclosure statement(s), visit the article page at http://dx.doi.org/10.1257/app.5.2.200.

[1] Some recent studies investigate the effects of irrigation dams (Duflo and Pande 2007), highways (Chandra and Thompson 2000; Michaels 2008), and railroads (Atack et al. 2009; Donaldson 2008; Banerjee, Duflo, and Qian 2012). See Estache (2010) for a review of the literature on infrastructure impact evaluations.

[2] Aschauer (1989); Canning and Bennathan (2002); Esfahani (2003); Estache, Speciale, and Verdas (2005); Canning and Pedroni (2004); Hulten, Bennathan, and Srinivasan (2005); and Yeaple and Golub (2007) estimate macro growth effects of infrastructure expansion.

and electrification in particular.[3] This is because electricity networks and other infrastructure are expanded in a planned manner, leading to reverse causality and program placement bias. Unlike health and education programs, large infrastructure projects do not lend themselves easily to researcher manipulation and randomization. Understanding the effects of investment in energy is important. A quarter of the world's population, and the majority in the poorest nations, still do not have access to electricity (Legros et al. 2009), and unreliable energy access can have large effects on firm productivity (Straub 2008).

This paper examines the effects of electricity grid expansions in Brazil between the years 1960 and 2000 on local economic development using a county and time fixed effects instrumental variables (IV) approach. To address endogeneity, we develop a model to forecast hydropower dam placement and grid expansion for Brazil that produces hypothetical maps that show how the electrical grid would have evolved over these 40 years had infrastructure investments been based solely on geographic cost considerations, ignoring demand-side concerns. This allows us to isolate the portion of the variation in grid expansion in Brazil that is attributable to exogenous cost considerations, and use it as an instrument to estimate the development effects of the impressive growth in electrification in Brazil over this period. This empirical strategy takes advantage of the fact that Brazil relies almost exclusively on hydropower to meet its electricity needs, and the cost of hydropower dam construction depends on topographic factors such as water flow and river gradient, since hydropower generation requires intercepting large amounts of water at high velocity. Hydropower is the fastest growing source of electricity worldwide (US Energy Information Administration 2010), and Brazil's experience over this period is therefore extremely relevant to understand the development effects of electrification.

We simulate the evolution of electrification as follows. First, the national budget for generation plants in each decade determines the number of hydropower dams to be constructed by the model in each decade. Second, each location in Brazil is ranked on its suitability for hydropower dam placement based on its geographic characteristics. The model places the new dams in the highest-ranking (i.e., most suitable) locations that do not yet have electricity according to the model until the national budget is exhausted. Third, we use a cost-minimizing algorithm to build transmission lines attached to each hydropower generation point. In summary, the evolution of the hypothetical electricity network is a function of geographic characteristics that affect suitability—water flow and river gradient, and the construction budget in each decade. Since the budget limits the number of dams constructed each decade, this creates discontinuities in the model, where (say) the 20 most suitable points may receive dams earlier, whereas the twenty-first most suitable location waits to get electrified.

The validity of our identification strategy depends on whether cost-side concerns in hydropower dam placement can be fully separated from demand-side concerns at

[3] Dinkelman (2011); Rud (2012); Asaduzzaman, Barnes, and Khandker (2010); Ketlogetswe, Mothudi, and Mothibi (2007); Grogan and Sadanand (2009); Khandker et al. (2009), and Kirubi et al. (2009); Fan, Zhang, and Zhang (2002) have examined the effects of electrification, and a subset of these studies have used instrumental variables strategies.

the level of variation in the data within a county. Cross-sectional variation would not provide the necessary exogenous variation, since people and firms are more likely to be located in water-rich areas. The estimator would also be biased in panel data settings if people and firms move over time along the same spatial lines as the forecasted placement of the electricity network within a county; from the lowest cost locations (robust water flow with a steep river gradient) in the early decades to slightly more expensive (flatter and less water-rich) locations in later years.

A number of characteristics of our data suggest that such concerns are likely to be minimal. We show that most areas of Brazil that are settled today were already settled by the beginning of our sample period, so growth of new settlements was not driven by people seeking untapped water resources. And, that the contemporaneous expansion of economic and demographic settlements follows a different spatial pattern than modeled grid expansion. This is probably because population and economic activity would not move in the same spatial order as our electricity grid forecasts that prioritize water and steep gradients. In fact, the rank order correlation between points where grid placement is predicted by our model, and where high density or GDP is predicted by the same geographic factors, is very low. Lastly, we show that lagged development does not predict the model's placement of new dams and transmission lines.

We run a battery of robustness tests to further address concerns about identification. For example, results are robust to the inclusion of other public services, such as roads, household sanitation, and water infrastructure. Even when we control for time-varying flexible trends in water resources, river gradient, and Amazon and Pantanal location, and rely only on discontinuities in the ranking of locations that receive a hydropower dam in a given decade (due to the decade-specific budgetary allocations), we retain enough first-stage predictive power to produce robust results.

We find large effects of electrification on two summary measures of development: the UN Human Development Index (HDI) computed for each county, and average housing values (under the assumption that improvements in living and working conditions in the county get capitalized into the value of housing). OLS regressions substantially underestimate the gains from electrification, which is consistent with either targeting of infrastructure to poorer areas, or that compliers in the IV approach (the hydropower dams identified by the cost-minimizing model) are the most cost-effective projects not built on the basis of political or other motivations.

The large development gains we observe are consistent with productivity improvements such as that modeled in Morrison and Schwartz (1996). However, the positive effects could also reflect selective in-migration of the most productive workers and firms into electrified areas, which creates larger regional disparities. To determine which of the two mechanisms is at play, we examine effects on a broader set of variables including in-migration, urbanization, salaries, employment, and education. Our analysis suggests that migration is unlikely to account for the large magnitude of development gains observed. We estimate large, positive effects of electrification on employment, salaries, and investments in education, but not health. The effects are of similar magnitude across sectors and across urban and rural economies. The pattern of results suggests that electricity led to some broad-based improvements as workers gained both post-secondary education and work experience in the decade following electrification. We compute the economic returns to electrification implied by our estimates and find that they exceed returns that have traditionally been calculated through demand estimation. This suggests that there are external market effects or agglomeration benefits to electrification not captured in a private demand function.

Our estimation strategy is related to Duflo and Pande (2007) which uses slope interacted with a time-varying state budget variable to predict irrigation dam placement in India (and Strobl and Strobl 2011 for Africa). They find that the poverty rate increases by 0.77 percent and agricultural wages decrease by 4.5 percent in the districts where dams are placed, but poverty decreases by 0.15 percent and wages increase by 6.9 percent in downstream districts. For hydropower dams in Brazil, we observe a temporary contemporaneous decrease in agricultural production and area harvested in counties where dams are built, but production rebounds quickly, and increases in comparison to areas without dams (see online Appendix Table A1). We also see a large reduction in poverty from electrification, and an increase in formal employment in rural areas. We attribute the difference in results to our focus on hydropower rather than irrigation dams, and to the fact that we estimate the effects over a longer time period.

Our results are related to Dinkelman's (2011) study of the employment effects of household access to electricity in a rural province of South Africa, KwaZulu-Natal, from 1995–2001. While that paper delves into specific household mechanisms, a distinctive contribution of this study is to report the long-run effects of electrification on a broad range of development indicators over a forty-year period—a more appropriate length of time to investigate macroeconomic changes. Dinkelman (2011) finds that female employment increases by 30–35 percent following electrification, while there is no statistically significant impact on male employment.

The rest of the paper proceeds as follows. Section I provides contextual information about the electricity sector in Brazil. Section II describes the data, and Section III describes the construction of the instrument. The estimate strategy is in Section IV, and our estimates of the development effects of electrification and possible mechanisms are in Section V. Section VI concludes.

I. Background on the Electricity Sector

Brazil provides an ideal setting to implement our estimation strategy because 85 percent of its electricity is generated from hydropower plants (US Energy Information Administration 2010). This dependence on hydropower enables us to forecast the expansion of electricity access well based on topographic features in Brazil. In addition, we benefit from the substantial variation in electrification over the period of interest. The transmission network in Brazil grew at an average rate of 8.9 percent per year, increasing in size from 2,359 kilometers in 1950 to 167,443 kilometers in 2000 (Operador Nacional do Sistema Elétrico (ONS) 2008). Generation capacity has also increased; 775 major electricity plants have been constructed in Brazil since 1910 (Agência Nacional de Energia Elétrica (ANEEL) 2008).

Electricity expansion was planned independently in each of the five major regions of Brazil until a 1961 law mandated coordination at the national level. Examining expansion plans from Brazil's south and southeast during this period (Canambra

Engineering Consultants 1969) indicates that both the government's development goals and load factors linked to local GDP and projected GDP growth were used to determine the expansion of the grid. Development goals and anticipation of growth are important opposing sources of selection bias in estimating the effects of electrification.

The 1961 legislation created a new national electricity company, Eletrobrás, to coordinate the financing of electricity projects and ensure that projects met the government's overall development goals for the country. Even though Eletrobrás took control of the four existing regional (North, Northeast, South, and South Central) electricity companies, the majority of the planning for grid expansion was devolved back to the four regions. This left the system fragmented, and given the high cost of transmission between regions, local infrastructure continues to matter for local electricity access.

During the 1960s and 1970s, electricity access was expanded primarily by increasing the number of isolated power generators. The generators provided power to local areas, but the electricity was not transmitted further than the region (Canambra Engineering Consultants 1968). The impressive expansion in generation capacity during this period was made possible by high electricity rates and the easy availability of financing. Investment in the electricity network slowed in the mid-1970s due to reduced financial means and remained low through the 1990s, leading to deterioration in network reliability (Gall 2002). Consequently, the largest variation in grid expansion in the data is for the earlier decades within our sample period.

Brazil reformed the electricity market in 1995, privatizing some of the state-owned electricity and distribution companies, but political conflict and economic crisis weakened the reforms. As a result, the government still owns 80 percent of the generation capacity in Brazil (Gall 2002). Despite an effort to integrate transmission across the four regions in the 1990s, much local electricity continues to be sourced from local or relatively nearby plants (Gall 2002).

In Brazil, the benefits of improved access to electricity accrue to many sectors of the economy. The industrial sector is the largest energy consumer, accounting for about half of Brazil's power usage since the 1970s. The agriculture sector uses a relatively small share of power, but usage increased from less than 1 percent in 1970 to almost 4 percent by 2009 due to more intensive use of water irrigation. The public and commercial sectors' shares of power usage have remained constant since 1970, at about 10 percent and 15 percent, respectively. The share of power used in the residential sector oscillates between approximately 20 percent and 25 percent (Instituto de Pesquisa Econômica Aplicada (IPEA) 2010).

Despite the impressive expansion of electricity, 27 percent of rural Brazilians still lack access to electricity (Energy Sector Management Assistance Program (ESMAP) 2005). Infrastructure investment has recently garnered renewed support as a development priority of the government of Brazil, and efforts are being made to provide electricity to all rural areas of Brazil.

II. Data and Variable Construction

We construct a county-level panel dataset for Brazil from 1960 to 2000 by combining data from a variety of sources. Data on historical electricity infrastructure (E) had to be constructed from feasibility studies, inventories, and maps provided by regional electricity companies. The procedures and data sources are described in detail in Section IIA and in online Appendix 1. Predicted electricity infrastructure (the instrument, Z) is constructed using GIS data, and the procedure and data sources are described in Section III, with further details in online Appendix 2. The dependent variables and other controls are drawn from the decennial census conducted by IBGE (2010), and aggregations from the census and other surveys constructed by IPEA (Instituto de Pesquisa Econômica Aplicada (IPEA) 2010). Exact definitions of the variables used in the analyses are in online Appendix 3.[4]

County (or *municipio*) borders in Brazil change over time predominantly due to a splitting of large counties into smaller counties. To create a panel dataset that compares the same geographic areas over time, we aggregate the smaller split counties using a crosswalk provided by IBGE and IPEA. We will refer to these 2,184 "standardized" counties as "counties" for brevity.

A. *Data on Actual Electricity Infrastructure*

We construct a measure of actual electrification (E) in each county for each decade by digitizing the locations of all major plants and electricity transmission substations. The raw data are derived from historical sources such as the feasibility studies and inventories conducted by Brazilian electricity companies, and come in the following two forms:

- Tables with inventories of all transmission lines that typically specify the county of origin, the destination county, length and voltage, and similar tables of power plants specifying location, type, and wattage.
- Large paper maps of generation plants and transmission lines by region of Brazil. Online Appendix Figures A1 and A2 provide examples of the maps and tables used to construct the dataset. We digitize and combine this information into GIS maps of the Brazilian electricity network for the 1960s, 1970s, 1980s, 1990s, and 2000s.[5] Power plants were placed on the digital map according to their reported latitude and longitude, while transmission substations were assumed to be located at the centroid of their county of record.

We collected data on generation plants and transmission lines, but not on the third component of the electricity grid—distribution networks. We restrict our attention to hydropower plants. Transmission lines transfer electricity from the generation plants to the region, while distribution networks transport electricity from the major local transmission substation to the household, industrial and agricultural consumers

[4] Most variables are available from 1960–2000, except county in-migration (only available in the 1990 and 2000 census) and car ownership rates (only available for 2000). Housing values are imputed from census data on rents.

[5] The 1960s network is based on the comprehensive inventory taken by Canambra (1968) and Canambra Engineering Consultants (1969) for 1965 and 1967. The 1970s network is pieced together from maps and tables for the different regions of Brazil from Eletrobràs. The 1980s network is based on another comprehensive inventory by Sistema de Informações Empresariais do Sector de Energia Elétrica (SIESE) (1987). The 1990s is again pieced together from various sources (e.g., Eletrobras Furnas 1993), and the 2000 network is based on data from 2000–2008 from SIESE (2008).

of electricity. We are unable to map the exact distribution networks over the period,[6] but, based on conversations with electricity sector professionals in Brazil, assume that, on average, distribution networks stretch one-hundred kilometers across. We divide Brazil into 33,342 evenly spaced grid points, and all grid points within a 50 kilometer radius of the centroid of a county containing a transmission substation are assumed to have access to electricity.[7] For regression analysis, these data are aggregated to the county level, so that actual electricity provision is defined as the proportion of electrified grid points within a county.

Figure 1 maps the evolution of the electricity network in Brazil from the 1960s through 2000. The early development of the electricity network was focused in the relatively affluent and industrial south, and from the 1970s onward the grid was expanded to the populous (but poorer) Southeast and Northeast. The network has expanded westward every decade since the 1970s, and by 2000 the coastal areas of the Southeast and Northeast had almost full coverage. The Amazon and Pantanal areas have remained largely unconnected and continue to have substantially less access to electricity than the rest of Brazil. We refer to these regions together as the Amazon in the remainder of the paper and in the tables.

III. The Instrument: A Model of Construction Cost Minimization

Our instrument (Z) is a prediction on electricity availability at each grid point in each decade, based on a model that simulates the evolution of generation plants and transmission lines in a way that minimizes construction cost. The model takes as inputs data on the geographic characteristics of each location and the national budget for each decade, and produces predictions for whether each of the 33,342 evenly spaced grid points has electricity access in each of the 5 time periods of data between 1960 and 2000. The geographic data are matched to existing hydropower dam data by 12 km buffer zones around each grid point.

We first outline the three steps needed to construct the instrument. Subsections IIIB–D provide details on specific assumptions and data sources for each of the three steps in the model, and online Appendix 2 provides additional technical details.

In step 1, the national budget for the number of hydropower plants built in each decade is set. The budget equals the number of dams that were actually built in all of Brazil during the decade. The data are from electricity company inventories in each period.

In step 2, we use data on the topographic characteristics of each grid point to rank-order all grid points in Brazil in terms of suitability for a low-cost hydropower dam placement based on topographic factors. The highest ranked grid points in terms of topographic suitability receive dams in the first decade until the budget (calculated in step 1) is exhausted. In the next decade, the next highest ranked grid points not already forecasted to have electricity receive dams, until the national budget for that decade is reached.

[6] Data on electricity distribution in Brazil is difficult to assemble because the data are decentralized across 64 privatized electricity companies, and there is no central clearinghouse.

[7] Figure A3 in the online Appendix illustrates this assumption on distribution coverage for Southern Brazil. The dark blue polygons are counties which have transmission substations, and the light blue circles surrounding them are assumed to be the distribution networks associated with those substations.

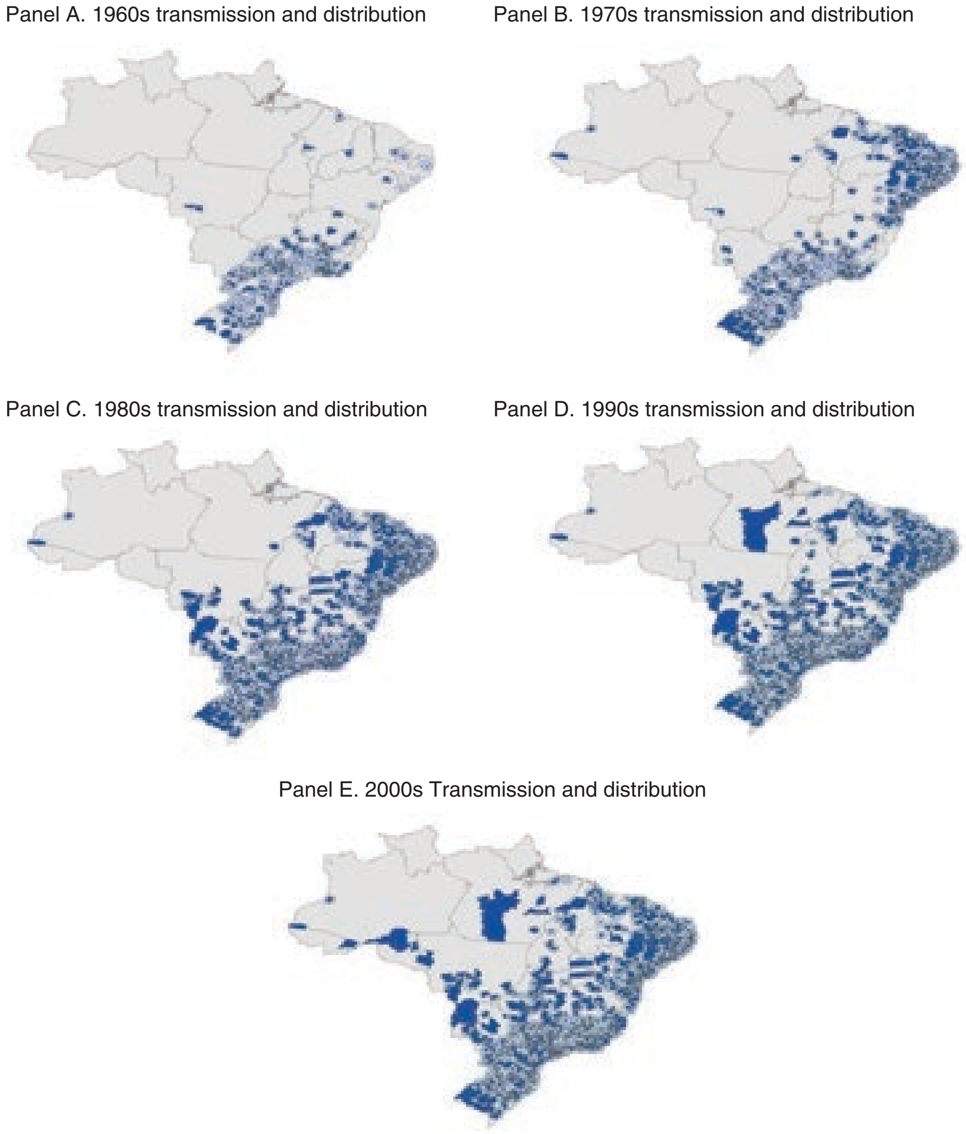

FIGURE 1

Source: Authors' calculations

In step 3, a cost-minimizing algorithm is used to construct two transmission lines to carry electricity from each hydropower dam that has been built. Electricity access is expanded to the area around the endpoints of the transmission lines.

A. *Step 1: Budget and Time Periods*

Our objective is to predict electricity availability for the five time periods for which actual electricity grid data is available. Our forecasting model matches the scale of expansion between two periods to the scale of investment in hydropower plants observed in the data for all of Brazil. A beginning balance of 240 power plants is allocated in the 1960s because that is the number of hydropower plants in existence in Brazil at the time of the inventory in 1967. The budget for the 1970s was 53 additional power plants, for the 1980s it was 36 additional plants, for the 1990s it was 25 additional plants, and for the 2000s it was 24 additional plants. The model takes these country-wide budgets as given and chooses the optimal location of hydropower plants and transmission lines within Brazil based on geographic factors.

B. *Step 2: Ranking the Suitability of Locations for Hydropower Dam Construction*

While electricity companies expand networks primarily in response to expected demand, access to electricity in a country that relies heavily on hydropower has an exogenous topographic component because the cost of access depends on the suitability of the local environment for power generation. In evaluating a new location for a hydropower plant, engineers consider available head, flow duration, and daily peaking operation to determine generation cost (Gulliver and Arndt 1991). "Available head" is determined by the amount of water flow and the change in elevation between the top and bottom of the dam. The head determines the amount of power that will be produced. Generating a given amount of power is cheaper when the available head is larger. The "flow duration" is determined by the amount of time in a given month (day, or year) the water flow required by the turbines is adequate for operation. The "daily peaking operation" is the flow duration that occurs during peak demand hours (Gulliver and Arndt 1991). When choosing a hydropower plant site, engineers also consider distance to the existing transmission network, as developing new transmission lines is expensive and can comprise a large component of the overall budget for the network (Canambra Engineering Consultants 1968).

Our model requires a measure of the cost of building a hydropower at each grid point. To do so, we gather data on topographic factors that affect the cost of generating electricity, and assign a weight or cost parameter to each of those factors. We collect information on topographic factors within ten kilometers of each grid point including whether there is a river, the average and maximum gradient of the river, maximum water flow accumulation,[8] and an indicator for whether the the grid point is in the Amazon.[9]

To estimate the cost parameters (the relative importance of each geographic factor in the dam location decision), we run a probit regression of an indicator for whether a location has a dam in the inventory data on the topography measures. Table 1 reports the results and shows that water resources (flow accumulation) and steep gradients are positive factors in dam placement, while the Amazon is a negative factor. The coefficients from this regression are used to assign cost parameters to each factor, which

[8] The geographic inputs in the model are from GIS maps from the US Geological Survey Hydro1k program (US Geological Survey (USGS) 2011). Using a GIS raster map of water bodies, we create two kilometer buffers on either side of each river, and compute the gradient along the river using elevation maps.

[9] To create a separate control variable for average land slope for the cross-sectional (comparison) specifications, we compute average slope within ten kilometers of each grid point using the elevation map. Figure A4 in the online Appendix explains the construction of the river gradient variable.

Table 1—Probit Regression for Hydropower Geographic Cost Parameters

Log of maximum flow accumulation	0.029**
	(0.014)
Average slope in the river	0.044
	(0.030)
Maximum slope in the river	0.062***
	(0.012)
Amazon indicator	−0.753***
	(0.066)
Indicator for location has a river	−0.030
	(0.063)
Observations	33,342

Notes: The dependent variable is an indicator for location has a hydropower plant. Standard errors clustered by county in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

allows us to rank all grid points in terms of suitability for hydropower dam construction. The highest ranking points are forecasted to receive hydropower dams first.

As a robustness check, we estimate the same probit regression of power plant placement on topographic factors using data on hydropower plant locations and topographic factors in the United States.[10] If our probit regression using Brazil data was mistakenly capturing some nonengineering cost determinants of dam placement, then the US model would likely find different coefficients on the same topographic variables. We find that the allocation of electricity is quite similar across the two models using either Brazilian or American data, which suggests that the model is responding to the relative geographic attractiveness for hydropower construction. The IV results are also qualitatively similar, but the results are more precise using the Brazilian data.[11]

Online Appendix Figure A5 maps the locations of the first 240 power plants that were predicted by the model for the 1960s. The red dots represent the predicted plants, the yellow dots represent the actual plant locations, and the color of the background reflects the elevation (darker colors are closer to sea level), and the blue lines show river locations. Our model predicts a large number of power plants along the Southeast to Northeast corridor (São Francisco river basin), where elevation changes quickly from the low-lying coastal areas, implying a steep increase in slope.

[10] The use of Brazilian data to parameterize the cost function may introduce concerns about endogenous placement bias, but, in our panel specifications, only if endogenous demand-side parameters move over time in a similar sequence from the lowest cost locations for building new power plants would those factors introduce endogeneity bias.

[11] This is not surprising since hydropower plants and water resource conditions in the United States are different from those in Brazil, and a larger percentage of generated power in Brazil is from hydropower. Geography therefore matters in different ways.

C. *Step 3: Placement of Transmission Lines*

The model next predicts the locations of substations (end-points of transmission lines) that deliver the electricity generated at each plant predicted from the previous step. We make the simplifying assumption that each power plant has the same generation capacity and is connected to two transmission substations. This assumption was made based on the average number of transmission substations per hydropower dam in the inventory data (Operador Nacional do Sistema Elétrico (ONS) 2008). The electricity network is assumed to be fully durable, and new substations and power plants cannot be placed on grid points that have already received electricity in prior decades, or from a generation plant constructed in a previous decade.

The model arrives at the lowest cost electricity network in each decade by computing costs for all possible arrangements of transmission lines. The model assumes that cost increases with distance and is prohibitively high when building substations in the Amazon (due to high material transport costs). We do not use any data to influence the direction of expansion of the transmission lines. There are a large number of possible permutations of transmission lines, and we use a numeric method to arrive at the lowest cost grid in equilibrium. Details are in online Appendix 2. Cost minimization in the transmission model results in short transmission lines, so the model predicts substations located close to generation points. To verify that the assumption that building in the Amazon is expensive does not drive the results, we run the first-stage regressions with and without controls of an Amazon indicator and an Amazon-specific time varying trend.

Once the equilibrium set of transmission lines is determined, we assume that all grid points within a 50 kilometer radius[12] of any substation will receive access to electricity, to account for the distribution network surrounding that substation. We assume a radius of 50 kilometers because that is average size of distribution networks and mirrors our treatment of distribution networks in the creation of the actual provision of electricity in Section IIA. We remain agnostic about the direction in which the distribution networks expand. In subsequent decades, new power plants are placed in the highest probability circles among those that have not yet received plants or substations, and locations for transmission substations are similarly proposed from among the grid points that remain without electricity infrastructure in the previous decades.

D. *Summary*

To summarize, our model consists of three simple steps. First, the national budget determines the number of dams (generation plants) to be built. Second, a probit model provides a priority ranking of grid point, which determines the spatial order in which dams will be allocated until the budget is exhausted. Third, a cost-minimizing algorithm determines the endpoints of the transmission lines, and a fixed area around those endpoints is assumed to receive electricity.

[12] The assumed access was slightly expanded in the model in comparison to the assumed radius for the distribution networks in order to simplify the Matlab algorithm (grid points on the diagonal are included). This has a minimal impact on the number of areas forecast to receive electricity in each decade.

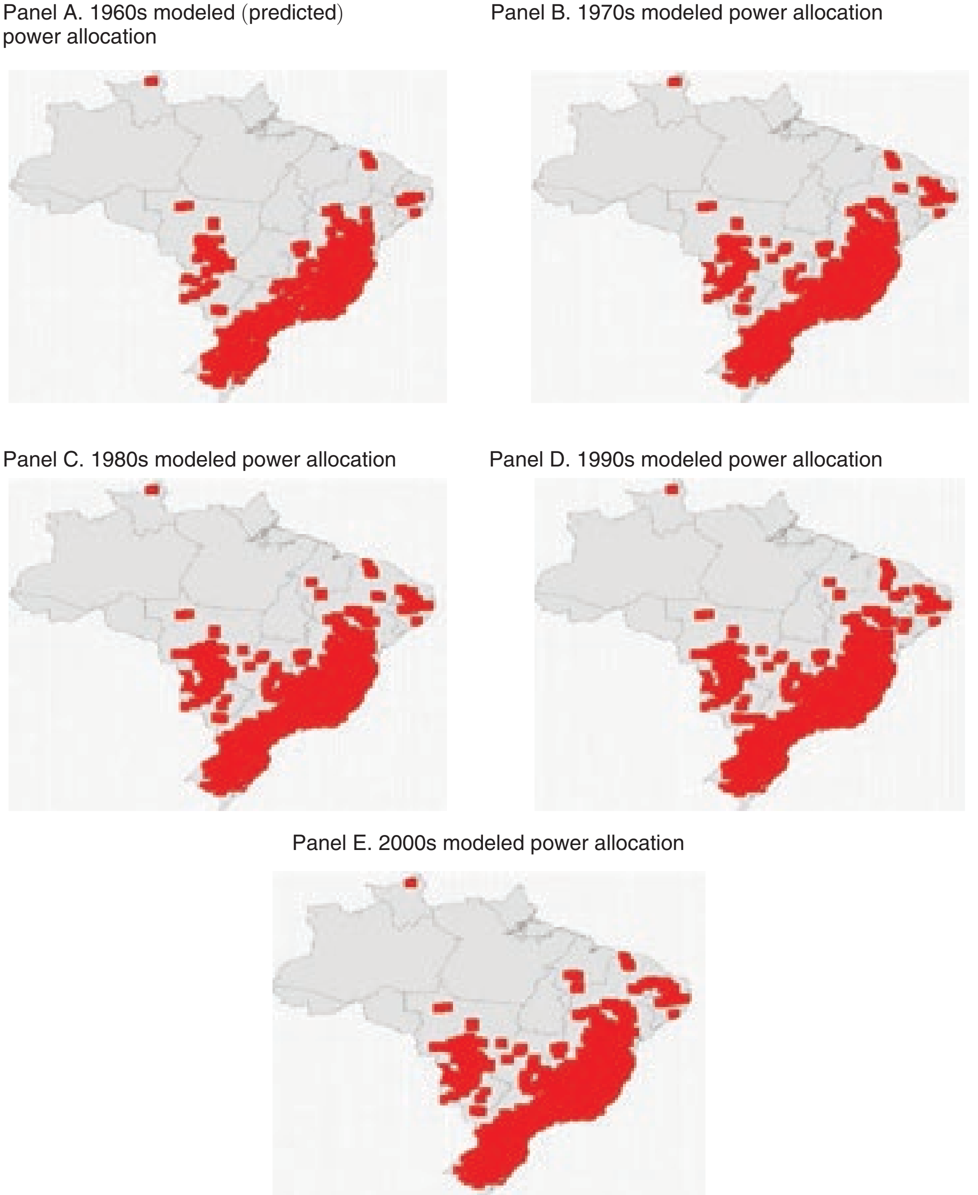

FIGURE 2

Source: Authors' calculations

Figure 2 plots the areas predicted to receive electricity by this model by decade. There is a reasonably good cross-sectional spatial correlation with the actual electricity network for Brazil (Figure 1), and the direction of expansion is similar.[13]

[13] Consistent with the vast expansion of the electricity network in 1960s through the mid-1980s, there is substantial variation across decades in the 1960s through 1980s, while the underinvestment in electricity infrastructure

The strength of this correlation in a model with location fixed effects determines the predictive power of the IV estimator. Online Appendix Figure A6 zooms in to two regions of Brazil to describe how the forecasting model works. Water rich areas with steeper gradients are more likely to receive infrastructure early, but the dynamics are mediated by a constraint in the model that areas forecasted to get electricity in a prior decade do not need new infrastructure.

Ignoring the demand-side drivers of expansion forces the model to under-allocate electricity to places like São Paulo and Rio de Janeiro, which leaves extra generation capacity that must be allocated elsewhere. This leads our hypothetical maps for the early decades to display broader spatial electrification coverage than what is observed in Brazil. This weakens the predictive power of the model.

IV. Estimation Strategy

This paper examines the effect of electrification on development over the period 1960 to 2000, at the county level, using a fixed effects instrumental variables (IV) approach. We estimate the effect of electrification on development outcomes in county, c, and time (decade), t, using the following 2SLS model with county and time with county fixed effects:

$$Y_{ct} = \alpha_c^1 + \gamma_t^1 + \beta \hat{E}_{c,(t-1)} + \epsilon_{ct}, \tag{1}$$

where $\hat{E}$ is instrumented electricity provision, predicted on the basis of our model forecasting the expansion of electricity in the first stage:

$$E_{c,(t-1)} = \alpha_c^2 + \gamma_t^1 + \theta Z_{c,(t-1)} + \eta_{ct}. \tag{2}$$

E_{ct} is the proportion of grid points in the county that are electrified in period t (Figure 1, Section IIA). Z_{ct} is the proportion of grid points in a county predicted to be electrified by the forecasting model (Section III, Figure 2). Electricity provision is lagged by a decade since the development of the electricity distribution network may take several years to complete following the construction of the hydropower dams.[14] Since the data are aggregated and the number of grid points is not the same in each county, we run weighted regressing using county area as weights. Standard errors, $\epsilon_{ct,}$ are clustered at the county level due to possible serial correlation. Table 2 displays summary statistics.

The IV strategy corrects for the bias introduced by the endogenous placement of electricity infrastructure. $Z_{c,t-1}$ is predicted based solely on geographic cost considerations for hydropower plant location (steep river gradient, significant

is mirrored in the much smaller variation in the modeled networks in the 1990s and 2000s.

[14] While there may be immediate regional economic effects from the construction of dams and hydropower plants, these would be primarily short term and focused in the vicinity of the power plant. Our infrastructure data (to measure E_{ct}) includes generation plants and transmission lines, and our forecasting model also predicts the placement of both, and, therefore, we identify effects beyond the local impact of a generation plant.

Table 2—Summary Statistics

Variable	Observations	Mean	SD	Min	Max
Electricity	8,730	0.72	0.42	0.00	1.00
Percent electrified	8,730	0.60	0.33	0.00	1.00
Modeled electricity instrument	8,730	0.59	0.47	0.00	1.00
Measures of demand					
Population density	8,730	62.71	292.19	0.04	10,097.86
GDP per capita (thousands of R$)	8,728	3.77	6.06	0.08	252.13
Industrial GDP per capita	8,730	1.20	3.71	0.00	112.18
Summary measures of development					
Average value of housing (thousands)	8,730	13.05	8.37	0.43	62.53
Human development indicator (index)	8,730	0.56	0.17	0.16	0.89
HDI components					
HDI longevity	8,730	0.57	0.12	0.17	0.88
HDI salaries	8,730	0.47	0.29	0.01	1.14
HDI education	8,730	0.52	0.15	0.08	0.88
Employment					
Percent economically active	8,730	0.36	0.07	0.18	0.63
Percent employed	8,730	0.35	0.06	0.13	0.61
Urban employment	8,730	0.34	0.07	0.10	0.61
Rural employment	8,685	0.35	0.07	0.00	0.74
Human capital					
Years of schooling	8,730	2.77	1.55	0.00	9.61
Illiteracy	8,730	32.00	17.96	1.80	89.90
Salaries per capita	8,730	0.11	0.09	0.00	0.79
Human capital per capita	6,549	19.06	7.12	6.57	59.01
Population changes					
New migrants in past five years	4,366	0.07	0.04	0.00	0.33
Life expectancy	8,730	60.10	7.81	38.40	76.92
Population density	8,730	78.50	372.90	0.04	11,732.17
Percent of population in urban areas	8,730	0.52	0.24	0.02	1.00
Poverty and inequality					
Poverty	8,730	60.47	24.92	3.81	99.88
Theil index	8,730	0.46	0.13	0.10	1.26
Infant mortality	8,730	71.96	50.63	6.93	303.66
Less than four years of education	8,730	65.25	21.16	9.07	99.80
Other infrastructure					
Percent of households with running water	8,730	0.39	0.28	0.00	0.99
Percent of households with sanitation	8,730	0.19	0.27	0.00	0.98
Percent of households with cars	4,366	0.01	0.01	0.00	0.13
Geography					
Landslope	8,730	1.64	2.01	0.00	21.76

Note: For descriptions of variable definitions and sources, please see online Appendix 4.

water flow, and Amazon location) and transmission line expansion (minimizing distance), and ignoring demand-side concerns that make actual electrification endogenous. The first stage attempts to isolate the portion of the variation in grid expansion (E_{ct}) that is attributable to exogenous cost considerations. Given the county and time fixed effects, the key identification assumption is that the demand side (people or firms) does not independently move over time along the

Table 3—First-Stage Regressions
Dependent Variable: Actual Electricity Availability from Infrastucture Inventories

Modeled electricity availability	0.563***	0.323***	0.222***
	(0.03)	(0.05)	(0.05)
Year FE	Yes	Yes	Yes
County FE	No	Yes	Yes
Amazon × year dummies	No	No	Yes
R^2	0.369	0.840	0.866
Observations	8,730	8,730	8,730
F-Stat	336.3	34.71	24.6
p-value	0.00	0.00	0.00

Notes: The dependent variable is prevalence of electricity infrastructure in the county. Regressions are weighted by county area. Standard errors clustered by county in parentheses. Measures of electrification are lagged by a decade in all our second-stage regressions, and the data used for the first-stage regression therefore covers 1960–1990. The Amazon and Pantanal are referred to jointly as the Amazon.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

same spatial lines as the forecasted placement of the electricity network—from the lowest cost locations (robust water flow with a steep river gradient) in the early decades to slightly more expensive (flatter and less water-rich) locations in later years. Sections VB and VD explore the validity of the IV-based identification strategy in greater depth.

V. Results

A. *First-Stage Results*

Table 3 shows the first-stage results that predict county-level actual electricity provision using the simulated instrument produced by our model based on geographic cost factors related to hydropower. To examine cross-sectional variation in the data, column 1 controls only for year fixed effects and not county fixed effects. The point estimate is statistically significant at the 1 percent level and shows a strong correlation between modeled and actual electricity provision (the point estimate is 0.56). Controlling for county fixed effects in the second column lowers the point estimate, 0.32, but it is still highly significant and shows that there is still a strong correlation between actual and modeled electricity even within counties. The Amazon is a fundamentally different region compared to the rest of Brazil, and it plays an important role in the forecasting model. Column 3 flexibly controls for a time-varying trend in the Amazon region by interacting the Amazon indicator with dummies for each decade. This specification of the first stage is the basis for most of the two-stage regressions reported in this paper. Using only the within-county variation in the data and excluding the variation from Amazon areas, we find that areas predicted to have electricity in a given decade, by our model, are 22 percentage points more likely to have electricity in that decade. The F-statistic on the first stage is 25.

B. *Validity of the Instrument*

In order for the instrument to be valid in a time and county fixed effects IV model, the demand side (people or firms) must not move independently over time along the same spatial lines as the forecasted placement of the electricity network within a county—from the lowest cost locations (robust water flow with a steep river gradient) in the early decades to slightly more expensive (flatter and less water-rich) locations in later years. We build confidence in the validity of the instrument by presenting evidence that the expansion of settlements followed a different spatial pattern than modeled electricity provision, and that the results remain robust to limiting the source of identification of the instrument. At the extreme, we rely solely on the nonlinearities and discontinuities built in to the forecasting model through decade budgets, and exclude the direct effects of the geographic variables.

It is possible that due to water scarcity, the population moved to new counties based on water availability during our period of analysis, leading settlement of counties in Brazil to independently follow the same pattern as electricity grid expansion. While this seems unlikely since Brazil has 13 percent of the world's freshwater resources, and all inhabited land is well covered by a dense network of small rivers and groundwater (Lipscomb and Mobarak 2011), we use census population data for 1910 onward to examine whether Brazil's counties were settled before the start of the analysis in 1960 (see online Appendix Figure A7). At a low-population density cutoff of 0.5/sq-km,[15] all counties in Brazil were already settled by the starting period of our analysis, except for some counties in the Amazon. Even at a high-population density cutoff of 5/sq-km, only 23 out of 2,184 counties are settled for the first time during the analysis period of 1960–2000. Water scarcity is therefore unlikely to drive population movements and settlement patterns directly during the period of analysis.

Another way to directly examine the question of whether people and/or firms move independently over time in the same spatial pattern as our forecast of electricity grid expansion, is to create a rank ordering of locations predicted to have the highest population density or highest GDP, using the same method we use to predict the most suitable locations for dam construction. To do so, we regress population density and GDP per capita on the same geographic characteristics as those used in the electricity forecasting model: water flow, river gradient, and Amazon. We then rank-order the points predicted to have the highest population and GDP by those regressions. We examine the Spearman rank order correlation between the suitability rank for hydropower generation and the suitability rank for population density in Table 4A and the correlation between hydropower suitability and GDP in Table 4B. We find that for each one of the five major regions of Brazil, the rank order correlation for population and hydropower suitability is low, and varies between -0.03 to $+0.06$. This is a conservative test of our identification assumptions, since this region fixed effects analysis is much less stringent than the county fixed effects we employ in all our regressions. The rank order correlations for GDP

[15] This is approximately the population density of the Western US states Wyoming and Montana around 1950 (US Census 2010).

Table 4A—Spearman Correlations—Hydropower Suitability and Population Density

Region	Year 1960s	1970s	1980s	1990s	2000s
Amazon	+0.0369	+0.0368	+0.0354	+0.0594	+0.0432
North East (including Bahia, Ceara, etc.)	−0.0298	−0.0290	−0.0341	−0.0321	−0.0349
Central West (including Pantanal)	+0.0217	+0.0172	+0.0141	+0.0442	+0.0375
South East (including Minas Gerais, Rio de Janeiro, Sao Paulo)	+0.0070	+0.0124	−0.0019	−0.0247	−0.0318
South (including Parana, Rio Grande do Sul)	+0.0486	+0.0447	+0.0506	+0.0532	+0.0631

Note: Each cell presents the Spearman rank order correlation between the suitability rank for hydropower generation and the rank for population density, by region and decade.

Table 4B—Spearman Correlations—Hydropower Suitability and GDP

Region	Year 1960s	1970s	1980s	1990s	2000s
Amazon	+0.0714	+0.0082	+0.0557	+0.0700	+0.0679
North East (including Bahia, Ceara, etc.)	+0.0098	+0.0127	+0.0259	+0.0968	+0.0689
Central West (including Pantanal)	+0.0067	−0.0155	+0.0141	−0.0078	−0.0138
South East (including Minas Gerais, Rio de Janeiro, Sao Paulo)	−0.0557	−0.0631	−0.0554	−0.0826	−0.0603
South (including Parana, Rio Grande do Sul)	+0.0031	−0.0043	−0.0237	−0.0104	+0.0238

Note: Each cell presents the Spearman rank order correlation between the suitability rank for hydropower generation and the rank for GDP, by region and decade.

Table 5—Robustness Check for Reverse Causality

	Instrument for electricity infrastructure	
Lagged housing value	0.000	
	(0.00)	
Lagged county HDI		−0.045
		(0.04)
R^2	0.984	0.984
Observations	6,549	6,549

Note: Standard errors clustered by county in parentheses. All regressions have county size weights and year dummies.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

per capita rank and hydropower suitability rank for the typical region-decade is also very close to zero, and ranges from −0.06 to +0.10.

Another way to directly examine the validity of the instrument is to examine whether the placement of power plants simulated by the forecasting model can be predicted by development indicators in earlier years. Results in Table 5 show that the point estimates on decade-lagged values of development indicators that serve as our main outcome variables of interest (housing values and county HDI) are close to zero and statistically insignificant. This suggests that at least lagged development indicators do not predict the spatial allocation of hydropower dams and transmission lines, and provides some confidence that the model's simulation of cost-minimizing electrification is orthogonal to demand side factors.

Table 6—Housing Values
Dependent Variable: Average Value of Housing

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged electricity infrastructure	5.023***	1.326***	0.801***	8.468***	7.792***	8.811***
	(0.90)	(0.35)	(0.27)	(1.52)	(1.72)	(3.03)
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes
County FE?	No	Yes	Yes	No	Yes	Yes
Amazon × year dummies[a]	No	No	Yes	No	No	Yes
R^2	0.153	0.922	0.925	0.106	0.191	0.151
Observations	8,730	8,730	8,730	8,730	8,730	8,730

Notes: Standard errors clustered by county in parentheses. The dependent variable is average housing value in thousands of reais. All regressions have county size weights and year dummies. The average housing value in the sample is 13.048.

[a]Topographic factor is interacted with a full set of decade fixed effects in order to flexibly control for differential trends by that Topographic factor.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 7—Human Development Index

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged electricity infrastructure	0.036**	0.009	0.006	0.091***	0.091***	0.109**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes
County FE?	No	Yes	Yes	No	Yes	Yes
Jungle × year dummies[a]	No	No	Yes	No	No	Yes
R^2	0.657	0.960	0.960	0.640	0.931	0.930
Observations	8,730	8,730	8,730	8,730	8,730	8,730

Notes: Standard errors clustered by county in parentheses. The dependent variable is the human development index. Year dummies are included in all regressions. All regressions have county size weights. The average HDI value in the sample is 0.557.

[a]Topographic factor is interacted with a full set of decade fixed effects in order to flexibly control for differential trends by that Topographic factor.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

C. *Second Stage Results: Effects of Electrification on Development*

In Tables 6 and 7, we present the effects of electrification on two summary measures of a county's development: the average value of the housing stock and a human development index measured for the county. Across all OLS and IV specifications, the effect of electrification on subsequent changes in average housing values is large, positive, and statistically significant at the 1 percent level. The OLS regressions without fixed effects (i.e., using all cross-sectional variation in the data, including the variation between Amazon and non-Amazon regions), a 10 percent increase in electrification is associated with a 502 reais increase in average housing value, which represents a 3.8 percent increase at the mean. Adding county fixed effects in columns 2–4 reduces the magnitude of this effect to 133 reais, and controlling for time-varying Amazon trends in column 5 reduces it further to 80 reais. The IV estimates using the simulated instrument from the forecasted placement of electricity based on geographic factors in our model are larger than the corresponding OLS estimates, and just as statistically significant.

One concern with the IV results is that the Amazon dummy plays a large role in infrastructure placement in the first stage prediction, and because the Amazon and non-Amazon areas are fundamentally different, a control for the Amazon should also be included in the second-stage regressions. We address this concern by directly controlling for the Amazon indicator interacted with dummy variables for each decade in the second stage in column 6.[16] In this IV specification with county fixed effects and time-varying Amazon effects, a 10 percent increase in electrification of the county increases average housing values by 881 reais, or 6.8 percent at the mean. The elasticity of housing values with respect to electrification is thus 0.7.

As shown in Table 7, the estimated effect of electricity on the human development index is also positive. In the OLS regression with cross-sectional variation, electrification is associated with about a 3.6 point increase in the county's human development index score. With the addition of county fixed effects and Amazon trends, the magnitude reduces to a statistically insignificant 0.6–0.9 point increase. Our preferred IV specification (with county fixed effects and Amazon trends) suggests that moving from 0 to full electrification increases the county HDI score by a statistically significant (at the 5 percent level or above) 9–11 points. Since the HDI is an index score based on sample values, it is instructive to interpret this result in the following way: the nine-point increase would take the median county in Brazil in 1980 to the human development level of the sixty-ninth percentile county. This represents a significant move within the distribution of HDI.

Comparing across the OLS specifications in Tables 6 and 7, we see that cross-sectional estimates are large, and the magnitude of the effects of electrification decrease in the within-county estimates. We also see that IV estimates of the effects of electrification are substantially larger than the OLS estimates. There are three possible reasons for the downward bias in OLS relative to IV estimates in our data. First, the compliers in our IV strategy (which is based on a forecasting model for the lowest-cost generation plants) are different from the average hydropower dam in operation in Brazil, whose placement may be affected by political considerations. Areas that received electricity primarily because of the low cost of provision rather than endogenous socioeconomic, political, or other demand-side factors may yield greater rates of return, which would make the IV estimates larger than OLS.[17]

[16] This is a more conservative control than the Amazon indicator interacted with the decade budget variable (which is the exact role that the Amazon plays in our instrument construction). The linear combination of Amazon × decade dummies subsumes Amazon × budget, or any other Amazon-specific time-varying effect.

[17] This interpretation is consistent with Cadot, Roller, and Stephan (2006), which shows that transport infrastructure is highly susceptible to politically motivated allocations. Engel, Fischer, and Galetovic (2009) shows that even once developed, public works projects may not be adequately maintained because of political considerations. And in Brazil, the allocation of publicly provided health services has been shown to be subject to political considerations (Mobarak, Rajkumar, and Cropper 2011).

Second, as described in detail in Section I, the electricity network in Brazil was designed and expanded primarily by the government or government managed utility companies during the period covered by our data (1950–2000). The demand-side endogeneity bias that the IV estimation corrects may have been of the form of the government targeting poorer areas important for maintaining political support (such as the program Luz para Todos) rather than more intensive expansions in developed areas where demand is likely to be greatest. OLS estimates would be biased downward due to the government's promotion of its development objectives.

Finally, the historical electricity infrastructure variable we construct (E_{ct}) by combining paper maps and tables on inventories from various electricity companies likely suffers from measurement error, while the geography variables used to predict the placement of electricity are measured quite precisely (based on 1km by 1km satellite maps). The IV estimates may be correcting the measurement error in the independent variable and addressing the associated attenuation bias.

D. *Robustness Checks*

Controls for Other Infrastructure.—It is possible that electricity proxies for a broader package of infrastructure investments. Infrastructure is sometimes delivered as a package, and solving the logistics associated with constructing transmission lines may itself lead to a parallel road being built or other infrastructure services being delivered more efficiently. The IV strategy is designed to mitigate this concern, and it is particularly helpful that the instrument's reliance on both water (which likely lowers cost of delivery and attracts infrastructure) and gradient (which, conversely, increases costs and deters other infrastructure) makes the spatial patterns of hydropower generation and of the delivery of other types of infrastructure distinct. Nonetheless, we examine the sensitivity of the results to inclusion of other infrastructure control variables.

In particular, we control for the percent of households in the county with running water and with improved sanitation access. We would like to control for the development of roads over time. Unfortunately, we do not have a direct measure of roads. Car ownership data is available from the 2000 and 2010 census, but a longer time series is needed for inclusion in the analyses. Instead, in Table 8, we show that the growth in car ownership is positively correlated with water availability in the county interacted with a time trend, and negatively correlated with land slope interacted with a time trend. Motivated by these two correlations, we use the water trend and the land slope trend as proxies for the road network.

Table 8 presents the effects of electrification on our summary measures of local development (HDI and housing values) controlling for these other measures of infrastructure (access to running water, improved sanitation, and indirect proxies for roads). The point estimates on *lagged electricity infrastructure* remain positive, statistically significant and of similar magnitude with these new controls. Two important shortcomings of this robustness check are that the provision of other infrastructure may also be endogenous and we do not have exogenous instruments for their availability, and our proxy for the growth of the road network is indirect.

Table 8—Robustness Tests: Infrastructure Controls

	Percent HH with cars		Housing value	HDI
	OLS		IV	IV
Lagged electricity infrastructure			13.708*	0.286**
			(7.14)	(0.14)
Lagged running water			−0.352	−0.110***
			(1.60)	(0.02)
Lagged sanitation access			−3.378*	−0.124***
			(1.93)	(0.02)
Water trend	0.016***	0.052***	−0.455	−0.012
	(0.00)	(0.02)	(0.37)	(0.01)
Landslope trend	−0.002**	−0.004	0.056	−0.002
	(0.00)	(0.00)	(0.16)	(0.00)
Year dummies?	Yes	Yes	Yes	Yes
R^2			0.051	0.739
Observations	4,366	4,366	6,549	6,549
Mean of dep var:			13.048	0.557

Notes: Dependent variables are average housing value and HDI. Standard errors clustered by county in parentheses. Decade dummies are included in all regressions. All regressions have county size weights. Water trend and landslope trend are included as proxies for the evolving availability of road infrastructure, for which we do not have available data spanning the time period of interest.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table 9—Additional Robustness Tests

	HDI		Housing value	
	IV	IV[a]	IV	IV[a]
Percent of houses electrified	0.933*		44.172*	
	(0.52)		(25.77)	
Lagged electricity infrastructure		0.045		3.075
		(0.03)		(1.91)
Year dummies?	Yes	Yes	Yes	Yes
Water-year dummies	No	No	No	No
County FE?	Yes	Yes	Yes	Yes
R^2	0.782	0.939	−0.722	0.393
Observations	8,732	8,730	8,732	8,730

Note: Standard errors clustered by county in parentheses. All regressions have county size weights and year dummies.

[a] *The Electricitiy Instrument* in this regression has been constructed using data on the importance of slope and water flow in the United States.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Alternative Definition of Electricity Provision.—Due to the possibility of measurement error in *lagged electricity infrastructure*, we use the percent of households with access to electricity (*percent of houses electrified*) from the census. This variable is likely measured with less error, and provides a direct measure of household-level connectivity. Table 9 presents the results using *percent of houses electrified* for the development outcomes HDI and housing values, respectively. The instrument has a poorer fit in the first stage for both outcome variables (F-statistic of 2), probably because the instrument is designed to predict the optimal placement of electricity infrastructure, whereas variation in the extent of household electrification is more directly determined by other demand factors. However, estimates of the development effects of electrification remain qualitatively and quantitatively similar. A 9 percentage point increase in the household electrification rate[18] increases the HDI by about 9 points, almost exactly the same magnitude estimated when using *lagged electricity infrastructure* in Table 7. The effects on housing values are also similar regardless of which variable is used to measure electricity provision.

Alternative Specification of the Instrument.—The first step for forecasting the placement of electricity based on geographic factors in our model uses data from Brazil to parameterize the cost function for building hydropower dams. A possible concern with this approach is that assigning the relative importance of water flow, river gradient and the Amazon using Brazilian data may introduce an element of demand-side preferences specific to Brazil in determining hydropower dam placement. We therefore re-estimate the cost function using data from the United States rather than Brazil, recalibrate the forecast of electricity placement and generate a new instrument. The geology of hydropower generation in the US is very different from Brazil because the two countries have very different levels of water resources and differential reliance on hydropower relative to other sources of energy. Accordingly, we lose some predictive power, and the estimated effects of electrification on development are only half as large, as shown in columns 2 and 4 of Table 9.

Variation Used in the Instrument.—As explained in Section IIB, the instrument contains a number of components, including three geographic characteristics (water flow, river gradient, Amazon location), a nationwide budget for infrastructure construction in each decade, and nonlinearities and discontinuities introduced by the order in which new infrastructure is built (based on suitability rankings). In Table 10, we unpack each component of the instrument, and show that the results remain robust when we add each individual component as controls in the second stage, either in isolation or jointly, for all possible combinations of these controls. Each row in Table 10 represents the results of a different 2SLS model. The first column indicates which variables are used as controls for that regression. Only the coefficient estimate on *lagged electricity infrastructure* ($E_{c,t-1}$) is reported in the second and third columns. The dependent variable is county housing values in column 2, and county HDI in column 3. The coefficient on the instrument ($Z_{c,t-1}$, *modeled electricity availability*) in the first-stage regression is reported in column 3. All regressions still include year and county fixed-effects, land slope, and Amazon-year dummies. This table therefore represents a number of robustness tests on the results reported in Tables 3, 6, and 7.

We find that controlling for the water flow (specification 1) or the river gradient (specification 2) interacted with the decade budget, or the combination of the

[18] If a county receives electricity infrastructure, the census-based household electrification rate increases by 9 percentage points in a regression controlling for county and year fixed effects.

Table 10—Sensitivity Analysis by Directly Controlling for Geographic Factors in the Second Stage

	Specification (description of control variables added to RHS) (1)	Housing values (2)	HDI (3)	First stage (4)	First-stage F-statistic (5)
1.	Water flow × decade budget[1,2]	7.622***	0.089***	0.324***	45.770
		(1.715)	(0.029)	(0.048)	
2.	River gradient × decade budget[4]	7.770***	0.091***	0.319***	47.765
		(1.737)	(0.029)	(0.047)	
3.	Amazon location dummy × decade budget[3]	9.579***	0.126***	0.223***	25.415
		(3.104)	(0.046)	(0.045)	
4.	Water flow × decade budget and Amazon location dummy × decade budget	9.339***	0.123***	0.222***	24.822
		(3.117)	(0.046)	(0.045)	
5.	River gradient × decade budget and Amazon location dummy × decade budget	9.798***	0.130***	0.223***	25.264
		(3.080)	(0.046)	(0.044)	
6.	River gradient × decade budget and water flow × decade budget	7.620***	0.089***	0.322***	48.569
		(1.730)	(0.029)	(0.047)	
7.	River gradient × decade budget, water flow × decade budget, and Amazon location dummy × decade budget	9.630***	0.128**	0.221***	24.365
		(3.084)	(0.046)	(0.045)	
8.	Water flow × year dummies[5]	7.662***	0.089***	0.324***	45.598
		(1.713)	(0.029)	(0.048)	
9.	Amazon location dummy × year dummies	8.811***	0.109**	0.222***	24.600
		(3.025)	(0.044)	(0.045)	
10.	River gradient × year dummies	7.719***	0.091***	0.220***	47.786
		(1.747)	(0.03)	(0.046)	
11.	Water flow × year dummies and Amazon location dummy × year dummies	8.637***	0.106**	0.221***	23.903
		(3.031)	(0.044)	(0.045)	
12.	River gradient × year dummies and Amazon location dummy × year dummies	9.085***	0.112**	0.223***	24.862
		(3.062)	(0.044)	(0.045)	
13.	Water flow × year dummies and river gradient × year dummies	7.630***	0.090***	0.319***	48.29
		(1.737)	(0.294)	(0.459)	
14.	River gradient × year dummies, water flow × year dummies, and Amazon location dummy × year dummies	9.078***	0.111**	0.220***	23.888
		(3.108)	(0.045)	(0.045)	
15.	Quartic suitability rank × year dummies[6]	8.510***	0.100**	0.239***	29.070
		(2.679)	(0.397)	(0.443)	

Notes: Standard errors clustered by county in parentheses. Each row represents a different sensitivity test on the specifications reported in column 5 in Tables 6 and 7. Column 2[(3)] in this table reports the coefficient and standard error on lagged electricity infrastructure where the dependent variable is housing values (HDI). Columns 4 and 5 report the associated first-stage result. The control variables added in each row are reported in column 1.

[1] Water flow is a measure of the available water in the county.

[2] The decade budget is a measure of available funding for hydroelectric dams in the county in a given decade.

[3] Amazon location dummy is one if the area is in the Amazon or Pantanal regions.

[4] River gradient is a measure of the steepness of the river and thus the speed of water flow.

[5] Topographic factor is interacted with a full set of decade fixed effects in order to flexibly control for differential trends by that Topographic factor.

[6] Suitability rank is the key input in the grid placement forecast model, and is a nonlinear combination of water flow, the river gradient, and the Amazon location dummy variables.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

two (specification 6), directly has very little effect on the first-stage power, the second-stage point estimates, or statistical significance. This remains true when the time-varying effects of river gradient or water flow are controlled for more flexibly using interaction terms with all decade dummies rather than the decade-specific budget (see specifications 8, 10, and 13).

Directly controlling for Amazon location does reduce the first-stage power (F-stat on excluded instrument in first stage reduces to 24 from 46), and slightly affects the second-stage point estimates. However, the first- and second-stage estimates remain highly statistically significant. The effect on the estimates is about the same whether the Amazon indicator is interacted with the decade budget or a full set of decade dummies, which suggests that the Amazon-related controls are really picking up the role that the Amazon indicator plays in our forecasting model for constructing the instrument. The Amazon effects are jointly statistically significant, and this becomes our preferred specification for all regressions reported in the remainder of this paper.

Table 10 also shows that at the extreme, even when we jointly control for the time varying effects of all three topographic characteristics (specifications 7 and 14), and rely only on the nonlinearities and discontinuities in rank that the model exploits in order to simulate the hypothetical grid, the estimates remain robust, and the predictive power of the instrument remains strong. Finally, in specification 15, we directly control for a quartic in each grid point's suitability rank for hydropower generation, derived through a probit equation in the second step of our instrument construction (see Section III). This rank variable is a nonlinear combination of the three geographic characteristics (water flow, river gradient, and jungle) that serve as an important input in our predictive simulation model. Even adding a very flexible control for the suitability rank, its square, cube, and quartic does not eliminate the first-stage power necessary for the instrumental variable based identification.

VI. Mechanism Underlying the Development Effects of Electrification

In this section, we explore the effects of electrification on a range of outcomes for which long-term, time-series data are available—income, education, health, urbanization, migration—in order to gain a deeper understanding of the changes that occurred in the local economies that resulted in the development effects we have observed. It is important to gauge whether electrification led to real changes in workers' incentives or ability to invest in human capital, or whether it induced movement of people and firms, so that the gains in human development and housing values simply reflect re-sorting of productive workers and firms toward electrified areas. Skill and productivity enhancements may justify greater national-level investment in infrastructure, but general equilibrium re-sorting may not.

A. *Effects by Component of HDI*

Table 11 presents results on the effect of electrification on each of the three components of human development:life expectancy, education, and income. All regressions

Table 11—Human Development Index Components and Other Poverty Measures

	HDI: Longevity		HDI: Income		HDI: Education	
	OLS	IV	OLS	IV	OLS	IV
Panel A						
Lagged electricity infrastructure	0	−0.01	−0.03*	0.45***	0.03***	0.19***
	(0.01)	(0.05)	(0.02)	(0.15)	(0.01)	(0.06)
R^2	0.84	0.8	0.89	0.5	0.91	0.65
Observations	8,730	8,730	8,730	8,730	8,730	8,730
Mean of dep var	0.569		0.472		0.515	
	Infant mortality		Gross income PC		Poverty	
	OLS	IV	OLS	IV	OLS	IV
Panel B						
Lagged electricity infrastructure	−7.99***	−11.97	−0.01	0.11**	−0.76	−42.17***
	(2.42)	(18.08)	(0.01)	(0.05)	(1.39)	(13.84)
R^2	0.9	0.86	0.84	0.58	0.85	0.53
Observations	8,730	8,730	8,730	8,730	8,730	8,730
Mean of dep var	71.96		0.114		60.469	

Notes: Standard errors clustered by county in parentheses. Dependent variables are the component indices of the HDI. All regressions have county size weights, year dummies, and jungle × year dummies.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

control for county fixed effects and Amazon × decade-specific dummies, and we show both OLS and IV results. We find that the development gains are concentrated in the income and education sectors, and not in health. The effect of electricity on life expectancy is both statistically insignificant and very close to zero. This is consistent with the possibility that electrification has conflicting effects on health. It allows for improvements in health technology and service delivery, but it may also increase pollution and strain through expansions of heavy manufacturing industries.[19]

The estimated effect of electrification on average household income is quite large and positive in the FE-IV specification, but negative in OLS (suggesting a downward bias from the government's targeting of underdeveloped areas). These results suggest that going from no electrification in a county to full electrification takes the median Brazilian county in 1980 to the income-HDI level of the eighty-fifth percentile county.

The education component of the UN Human Development Index is comprised of literacy and school enrollment. A county gaining access to full electrification leads to a gain of 19 index points in its education score according to the FE-IV specification. This represents a move from the fiftieth percentile to the ninety-second percentile county in 1980.

The last four columns of Table 11 provide results for the effect on direct household income and poverty measures, as opposed to index values. The development gains

[19] Alternatively, it could indicate that electricity is not an important determinant of infant mortality—a main driver of life expectancy in many developing countries. However, Table 11 also shows that there is a substantial (but statistically insignificant) decrease in infant mortality associated with electrification.

Table 12—Dependent Variables: Measures of Employment Effects

	Economically active		Formal employment		Formal employment (urban)		Formal employment (rural)	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Lagged electricity infrastructure	0.011*	0.173***	0.010*	0.184***	0.009*	0.176***	0.007	0.165***
	(0.01)	(0.05)	(0.01)	(0.05)	(0.00)	(0.05)	(0.01)	(0.05)
R^2	0.759	0.349	0.689	−0.171	0.791	0.244	0.612	−0.129
Observations	8,730	8,730	8,730	8,730	8,730	8,730	8,685	8,678
Mean of dep var	0.364		0.347		0.338		0.349	

Notes: Standard errors clustered by county in parentheses. Dependent variables are employment variables. All regressions have county size weights, year dummies, and jungle × year dummies.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

are statistically significant and fairly large—a 10 percent increase in electrification increases income per capita by almost 10 percent and reduces poverty by about 7 percent at the mean.

B. *Employment Effects by Sector*

Having established that the gains from electrification are concentrated in income and education (and not in health), next we investigate whether the income effects are realized due to better employment conditions, and the sectoral distribution of the gains between formal and informal sectors, and urban versus rural areas. Table 12 reports the effects of electrification on two measures of employment: if the person was economically active,[20] and if they had formal employment. It also presents results on formal employment in urban and rural areas separately.

The results show that electrification leads to increases in both formal employment, and the broader notion of being "economically active." Specifically, a county that goes from 0 to full electrification would experience a 17–18 percentage point increase in the probability of employment based on either measure of employment. This represents a 47 percent improvement in mean employment across Brazilian counties, and corroborates the large gains in income in Table 11. In addition, these gains are distributed similarly in urban and rural areas within counties. The similarity in employment effects across formal and informal sectors, and across urban and rural areas is quite striking. It is suggestive of improvements that cut across sectors, such as worker skills, rather than industry-specific technological improvements.

Dinkelman (2011) finds that in rural South Africa, female employment rises by 9 to 9.5 percentage points in the time period shortly following electrification. We find even larger employment effects in the longer run, possibly because firms can take advantage of labor inputs when they have a longer time to adjust their production function (Morrison and Schwartz 1996).

[20] Being economically active is a broader definition of employment than formal employment. It includes employment in the formal or informal sectors as well as self-employment.

Table 13—Dependent Variables: Measures of Education Effects

	Illiteracy		Less than four years education		Years in school		Human capital	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Lagged electricity infrastructure	−2.700***	−8.350*	−0.359	−21.253***	0.062	2.022***	2.092***	11.541
	(0.72)	(4.78)	(0.90)	(7.75)	(0.08)	(0.67)	(0.41)	(7.30)
R^2	0.944	0.815	0.944	0.871	0.936	0.791	0.965	0.887
Observations	8,730	8,730	8,730	8,730	8,730	8,730	6,549	6,549
Mean of dep var	32.00		65.248		2.77		19.06	

Notes: Standard errors clustered by county in parentheses. Dependent variables are education and salary variables. All regressions have county size weights, year dummies, and jungle × year dummies.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

C. *Effects on Educational Attainment*

Improvements in employment maybe related to better educational attainment. In Table 13, we examine whether effects on education were concentrated in reducing illiteracy, improving primary education, or increasing years of education. Results show that going from no electricity to full electrification in a county leads to a reduction in the illiteracy rate of 8 percentage points (25 percent drop at the mean) and in the proportion of the population with less than four years of education of 21 percentage points (32 percent decrease at the mean). However, the largest gains were experienced in years of schooling, it increased by two years, which represents about a 72 percent increase at the mean. This suggests that more children obtained post-primary (or grade four) education, which may have led to labor productivity increases.

D. *Real Productivity Gains or the Effects of Migration and Sorting?*

We have documented positive income and employment effects of electrification across sectors and locations. These gains may have been a result of real effects of electrification (e.g., increases in efficiency and returns to education allow workers to invest more in education and quality of complementary inputs, such as capital increases, or there is better accumulation of work experience as employment conditions improve), or they may be a result of re-sorting of workers through migration into electrified areas. Our goal in this subsection is to examine whether the large development effects we have observed can be fully explained by migration and re-sorting.

To determine if migration is driving the results, we examine the effects of electrification on in-migration in each county in Table 14. Migration data is only available for the 1990 and 2000 census, which considerably shortens the panel. The IV results indicate that a 10 percent increase in electricity provision leads to a 1 percentage point increase in the influx of migrants into counties. The results are not statistically significant, possibly due to the short panel. An important point to note is that only

Table 14—Population Changes
Dependent Variables: Measures of Population Effects

	In-migration rate		Life expectancy		Population density		Urban percent of pop.	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Lagged electricity infrastructure	0.010	0.102	−0.437	−1.034	−1.107	−23.618	0.013	0.238**
	(0.03)	(0.09)	(0.32)	(2.39)	(3.74)	(19.20)	(0.01)	(0.11)
R^2	0.951	0.371	0.935	0.924	0.940	0.010	0.928	0.742
Observations	4,366	4,366	8,730	8,730	8,730	8,730	8,730	8,730
Mean of dep var	0.072		60.098		78.502		0.517	

Notes: Standard errors clustered by county in parentheses. Dependent variables are population change variables. Migration data is available only for 1990 and 2000, making the panel substantially shorter. All regressions have county size weights, year dummies, and jungle × year dummies.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

7 percent of the average county's population is comprised of recent migrants, and thus even a doubling of the in-migration rate does not change the composition of the population dramatically. A 10 percent increase in electrification would increase the migrant share of the population from 7.2 percent to 8.2 percent. It is unlikely that this increase in the migrant share could account for the 7 percent (or 4.2 percentage point) reduction in poverty and 4.7 percent improvement in employment associated with that 10 percent increase in electrification. Therefore, even taking the large (but imprecisely estimated) coefficient in the migration regression at face value, it can only explain a small portion of the development gains.

We next examine effects on county population density, and within-county urbanization, to look for further evidence of changes in population composition using variables for which we have a longer panel. The urbanization measure does provide evidence of substantial within-county sorting following electrification. Going from 0 to full electrification leads to 24 percentage points more of the county population being classified as "urban," which could either be a result of rural residents shifting toward the population centers within counties, or because the greater agglomeration leads to more of the county being classified as urban by the statistical agency. Either way, this is a within-county move, and cannot explain away the cross-county estimates of productivity gains associated with electrification.

VII. Conclusion

Unreliable energy in the developing world is a significant obstacle for firms (Straub 2008), and donors and governments have recently increased investment in large-scale electricity projects. Unreliable infrastructure imposes large costs on companies, and together with high taxes is colloquially called "custo Brazil," or the cost of doing business in Brazil.

Recent initiatives have tried to compensate for years of shortfalls in energy infrastucture investment. World Bank investments in energy nearly doubled during

2005–2008 relative to 2001–2004 (Barnes, Singh, and Shi 2010). There are many competing demands on these development funds, and understanding the returns to electricity provision can help policy makers determine spending priorities. Brazil is considering new initiatives that would lead to the privatization of the provision of infrastructure services, particularly electricity (*Economist* 2012), which makes understanding the effects of electricity access even more important.

Understanding the effect of electricity on development is challenging due to reverse causation and other endogeneity concerns. To address these concerns, we exploit variation in geography to model electricity access based on engineering cost factors that are exogenous to demand factors for electricity. The methodology of isolating the variation in infrastructure to exogenous budget and geographic cost considerations can be useful for studying the effects of hydropower investments in other countries, and the general concept can be applied to a broader range of infrastructure projects.

Applying fixed effects IV estimation to data from 1960–2000, we find large development gains from investments in electricity. Further, we show that these effects are not explained by re-sorting of productive resources in general equilibrium, that they are value-added to society. The effects are substantially underestimated when one fails to account for endogenous placement.

We now compute a back-of-the-envelope estimate of the monetary returns to electrification implied by our regression estimates, and compare it to returns calculated by traditional cost-benefit analysis (e.g., Berndt 1986; Berndt, Harper, and Wood 1989). The increase in electricity coverage in the populated (non-Amazon) areas of Brazil was 43 percentage points over our sample period (from 32 percent in 1960 to 75 percent in 2000). At an average cost of $3.5 million per MW,[21] we estimate that it costs $33 million, on average, to electrify a county that did not previously have access to electricity, and the expansion in access to electricity from 1960–2000 therefore cost an estimated $40 billion.

Our estimates suggest that electrification of a county increases housing values by approximately 9,000 reais, on average, over a decade ($4,900 at an exchange rate of 1.83 reis to $1). With an average of 12,400 households per county in 2000, this suggests an increase of $61 million in total land value, on average, for counties when they receive access to electricity. Policy in Brazil has been to charge tariffs equal to the cost of provision including a 10–12 percent rate of return for electricity companies, so the returns we calculate are after the repayment of the initial investment costs. The rate of return based on the increase in land values was therefore 184 percent *beyond* the cost of provision.[22] Traditional cost-benefit analyses of the impact of electricity provision vary widely, but the internal economic rate of return tends to be around 5–15 percent, with a few estimates around 100 percent (Munasinghe 1987). Our calculations suggest that the rate of return on electricity investments in Brazil was much higher—18.4 percent *above* the 10–12 percent return guaranteed to electricity companies through the electricity tariffs.

[21] Calculated as an average of hydropower generation costs in Brazil from the International Energy Agency's *Projected Costs of Generating Electricity* report (2010).

[22] This relies on the conservative assumption that infrastructure investment paid off over a decade. Returns are calculated over ten years.

The larger returns we estimate suggest that traditional estimates based on electricity demand tend to underestimate the benefits of electricity access. While traditional cost/benefit analyses calculate the returns through demand estimation, our longer run estimates will include external market effects as firms and workers adjust. The larger effects we estimate therefore suggest substantial external impacts of electricity access. Agglomeration is one likely source of these externalities. While the data shows little evidence of migration across regional or county boundaries, there is evidence that people are moving to urban areas within the county following electrification. This within-county migration may lead to increased human capital development and faster technological change. Further research remains to be done on the mechanisms through which agglomeration benefits resulting from electricity access may occur.

REFERENCES

Agência Nacional de Energia Elétrica. 2008 "Sistema De Informações Georreferenciadas do Setor Elètrico (SIGEL)." Agência Nacional de Energia Elétrica (ANEEL). http://sigel.aneel.gov.br/ (accessed June 1, 2008).

Ali, Ifzal, and Ernesto Pernia. 2003. *Infrastructure and Poverty Reduction: What Is the Connection?* Asian Development Bank Economics and Research Department (ERD) Policy Brief 13. Manila, January.

Asaduzzaman, Mohammad, Douglas Barnes, and Shahidur R. Khandker. 2010. "Restoring Balance: Bangladesh's Rural Energy Realities." World Bank Working Paper 181.

Aschauer, David Alan. 1989. "Is Public Expenditure Productive?" *Journal of Monetary Economics* 23 (2): 177–200.

Atack, Jeremy, Fred Bateman, Michael Haines, and Robert A. Margo. 2009. "Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850–60." National Bureau of Economic Research (NBER) Working Paper 14640.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2012. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." National Bureau of Economic Research (NBER) Working Paper 17897.

Barnes, Douglas F., Bipul Singh, and Xiaoyu Shi. 2010. *Modernizing Energy Services for the Poor: A World Bank Investment Review— Fiscal 2000–08*. World Bank Energy Sector Management Assistance Program (ESMAP). Washington, DC, December.

Berndt, Ernst R. 1986. "Electrification, Embodied Technical Progress, and Labor Productivity Growth in U.S. Manufacturing, 1889–1939." In *Electricity Use, Productive Efficiency and Economic Growth*, edited by Sam Schurr and Sidney Sonenblum, 93–114. Palo Alto: Electric Power Research Institute.

Berndt, Ernst R., Michael J. Harper, and David O. Wood. 1989. "Rates of Return and Capital Aggregation using Alternative Rental Prices." In *Technology and Capital Formation*, edited by Dale W. Jorgenson and Ralph Landau, 331–72. Cambridge, MA: MIT Press.

Cadot, Olivier, Lars-Hendrik Roller, and Andreas Stephan. 2006. "Contribution to Productivity or Pork Barrel? The Two Faces of Infrastructure Investment." *Journal of Public Economics* 90 (6–7): 1133–53.

Canambra Engineering Consultants. 1966. *1966 Power Study of South Central Brazil*. United Nations Development Programme (UNDP). New York.

Canambra Engineering Consultants. 1969. *1969 Power Study of South Brazil*. United Nations Development Programme (UNDP). New York.

Canning, David, and Esra Bennathan. 2002. "The Social Rate of Return on Infrastructure Investments." World Bank Policy Research Working Paper 2390.

Canning, David, and Peter Pedroni. 2004. "The Effect of Infrastructure Investment on Long Run Economic Growth." Williams College Department of Economics Working Paper 2004–04.

Chandra, Amitabh, and Eric Thompson. 2000. "Does Public Infrastructure Affect Economic Activity?: Evidence from the Rural Interstate Highway System." *Regional Science and Urban Economics* 30 (4): 457–90.

Companía Hidroelectria de Sao Franciso (CHESF). Online data http://www.chesf.gov.br/page/portal/paginas/sistema_chesf/sistema_chesf_sistema_transmissao/conteiner_sistema_transmissao (accessed July 2008)

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review* 101 (7): 3078–3108.

Donaldson, David. 2008. "Railroads and the Raj: The Economic Impact of Transportation Infrastructure." http://economics.mit.edu/files/6038 (accessed February 15, 2013).

Duflo, Esther, and Rohini Pande. 2007. "Dams." *Quarterly Journal of Economics* 122 (2): 601–46.

Eletrobras. 1973. "Regiaõ Sul, Usinase Linhas de Transmissaõ." Unpublished.

Eletrobras Grupo Coordenador para Operacao Interligada. 1994. "Plánejamento electro e energetico da operacao para 1994."

Energy Sector Management Assistance Program (ESMAP). 2005. *Brazil Background Study for a National Electrification Strategy: Aiming for Universal Access*. World Bank. Washington, DC, March.

Engel, Eduardo, Ronald Fischer, and Alexander Galetovic. 2009. "On the Efficient Provision of Roads." http://cowles.econ.yale.edu/~engel/pubs/efg_efficient-prov.pdf.

Esfahani, Hadi Salehi. 2003. "Institutions, Infrastructure, and Economic Growth." *Journal of Development Economics* 70 (2): 443–77.

Estache, Antonio. 2010. "A Survey of Impact Evaluations of Infrastructure Projects, Programs and Policies." European Centre for Advanced Research in Economics (ECARES) Working Paper 2010–005.

Estache, Antonio, Biagio Speciale, and David Veredas. 2005. "How Much Does Infrastructure Matter to Growth in Sub-Saharan Africa." http://www.ecares.org/ecare/personal/veredas$/David%20Veredas%20African%20infrastructures%20Version%201.pdf.

Fan, Shenggen, Linxiu Zhang, and Xiaobo Zhang. 2002. *Growth, Inequality, and Poverty in Rural China: The Role of Public Investments*. International Food Policy Research Institute Research Report 125. Washington, DC.

FURNAS. 1993. "Sistema de Transmissao do Sudeste/Centro-Oeste." Prinapais Troncos-Atual e Previsto. September.

FURNAS. 1995. "Localizacao Dos Empreendimentos. 1 Sistema de Geracao e Transmissao. November

Gall, Norman. 2002. "Blackout in Energy Policy: Brazil's Difficulties in Making Decisions." Fernand Braudel Institute of World Economics Braudel Papers 31.

Grogan, Louise, and Asha Sadanand. "Electrification and the Household." http://www.isid.ac.in/~pu/conference/dec_09_conf/Papers/AshaSadanand.pdf (accessed February 15, 2013)

Gulliver, John S., and Roger E. A. Arndt. 1991. *Hydropower Engineering Handbook*. New York: McGraw-Hill.

Hulten, Charles R., Esra Bennathan, and Sylaja Srinavasan. 2006. "Infrastructure, Externalities, and Economic Development: A Study of the Indian Manufacturing Industry." *World Bank Economic Review* 20 (2): 291–308.

Instituto Brasileiro de Geografia e Estatística. "Census." http://www.ibge.gov.br/english/download/estatistica.shtm (accessed July 2011).

Instituto de Pesquisa Econômica Aplicada. "IPEADATA." http://www.ipeadata.gov.br/ (accessed July 2011)

Ketlogetswe, C., T. H. Mothudi, and J. Mothibi. 2007. "Effectiveness of Botswana's Policy on Rural Electrification." *Energy Policy* 35 (2): 1330–37.

Khandker, Shahidur R., Douglas F. Barnes, Hussain Samad, and Nguyen Huu Minh. 2009. "Welfare Impacts of Rural Electrification: Evidence from Vietnam." World Bank Development Research Group Sustainable Rural and Urban Development Team Policy Research Working Paper 5057.

Kirubi, Charles, Arne Jacobson, Daniel M. Kammen, and Andrew Mills. 2009. "Community-Based Electric Micro-Grids Can Contribute to Rural Development: Evidence from Kenya." *World Development* 37 (7): 1208–21.

Legros, Gwenaelle, Ines Havet, Nigel Bruce, and Sophie Bonjour. 2009. *The Energy Access Situation in Developing Countries: A Review Focusing on the Least Developed Countries and Sub-Saharan Africa*. United Nations Development Programme (UNDP) and World Health Organization (WHO). New York, November.

Lipscomb, Molly, and Ahmed Mushfiq Mobarak. 2011. "Decentralization and the Political Economy of Water Pollution: Evidence from the Re-Drawing of County Borders in Brazil." http://faculty.som.yale.edu/mushfiqmobarak/decentralization%20of%20water%20brazil.pdf (accessed February 15, 2013)

Lipscomb, Molly, A. Mushfiq Mobarak, and Tania Barham. 2013. "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil: Dataset." *American Economic Journal: Applied Economics*. http://dx.doi.org/10.1257/app.5.2.200.

Michaels, Guy. 2008. "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System." *Review of Economics and Statistics* 90 (4): 683–701.

Mobarak, Ahmed Mushfiq, Andrew Sunil Rajkumar, and Maureen Cropper. 2011. "The Political Economy of Health Services Provision in Brazil." *Economic Development and Cultural Change* 59 (4): 723–51.

Morrison, Catherine J., and Amy Ellen Schwartz. 1996. "State Infrastructure and Productive Performance." *American Economic Review* 86 (5): 1095–1111.

Munasinghe, Mohan. 1987. *Rural Electrification for Development: Policy Analysis and Applications*. London: Westview Press.

Operador Nacional do Sistema Elétrico. 2008. "Sistema De Informações Geogràficas Cadastrais do SIN (SINDAT)." Operador Nacional do Sistema Elétrico (ONS). http://www.ons.org.br (accessed June 1, 2008).

Organisation for Economic Co-operation and Development, and International Energy Agency. 2010. *Projected Costs of Generating Electricity*. Organisation for Economic Co-Operation and Development (OECD) Nuclear Energy Agency and International Energy Agency (IEA). Paris, June.

Paulo, São. 2012. "Facing Headwinds, Dilma Changes Course: The Government Announces Plans to Privatize Infrastructure and Disappoints Striking Bureaucrats." *Economist*, August 18. http://www.economist.com/node/21560565.

Rud, Juan Pablo. 2012. "Electricity Provision and Industrial Development: Evidence from India." *Journal of Development Economics* 97 (2): 352–67.

Sistema de Informações Empresariais do Setor de Energia Elétrica (**SIESE**)**.** 1987. *Relatório Estatistico De Linhas De Transmissão e Subestações 1970, 1975, 1980–86*. Rio de Janeiro: Eletrobras.

Sistema de Informações Georreferenciadas do Setor Elétrico (**SIGEL**), **Agencia Nacional de Engia Electrica** (**ANEEL**). http://sigel.aneel.gov.br (accessed July 2008).

Sistema de Informções do SIN (**SINDAT**), Operador Nacional de Energia Electrico, ONS http://www.ons.org.br/conheca_sistema/dados_tecnicos.aspx (accessed July 2008).

Straub, Stephane. 2008. "Infrastructure and Development: A Critical Appraisal of the Macro Level Literature." World Bank Policy Research Working Paper 4590.

Strobl, Eric, and Robert O. Strobl. 2011. "The Distributional Impact of Large Dams: Evidence from Cropland Productivity in Africa." *Journal of Development Economics* 96 (2): 432–50.

US Census. 2010. "Resident Population Data: Population Density." US Census. http://2010.census.gov/2010census/data/apportionment-dens-text.php (accessed September 10, 2011).

US Energy Information Administration. 2010. *International Energy Outlook 2010*. US Energy Information Administration. Washington, DC, July.

US Geological Survey. 1996 "HYDRO1k Elevation Derivative Database." US Geological Survey (USGS). http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/hydro (accessed July 7, 2010).

World Bank. 2003. *Infrastructure Action Plan*. World Bank. Washington, DC, July.

Yeaple, Stephen R., and Stephen S. Golub. 2007. "International Productivity Differences, Infrastructure, and Comparative Advantage." *Review of International Economics* 15 (2): 223–42.