Infrastructure Development Can Benefit the Environment: Electrification, Agricultural Productivity and Deforestation in Brazil

Juliano Assunção
PUC-Rio

Molly Lipscomb
University of Virginia

Ahmed Mushfiq Mobarak
Yale University

Dimitri Szerman
CPI and PUC-Rio

April 15, 2015

Abstract

A common forest protection strategy used in countries where legal enforcement of deforestation laws is weak aims to increase agricultural productivity and intensification, in order to reduce the pressure to clear forests for new land. We show that such productivity-enhancing strategies can have ambiguous effects on forest protection in theory: it can expand the scope of farming, which is detrimental to the forest, but it can also induce farmers facing factor-market constraints to switch away from environmentally destructive cattle grazing and into less-harmful crop cultivation. We examine these predictions using five waves of the Brazil Agricultural census, and by re-constructing the evolution of the electricity grid for Brazil for the period 1960-2000. We show that electrification increased crop productivity (partly through irrigation investments), and farmers' land use decisions subsequently change. The productivity increase leads to farming expansion and frontier land conversion, but also shifts land use within farms away from land-intensive cattle ranching and into crop cultivation. The latter allows farmers to retain more native vegetation within rural settlements, and overall we estimate that electrification causes a net decrease in deforestation in Brazil. We address the endogeneity of infrastructure placement by developing a model that forecasts hydropower dam placement based on topographic attributes of each location, and produces hypothetical maps that show how the electrical grid would have evolved had infrastructure allocation been based solely on exogenous cost considerations. The maps isolate the exogenous portion of the panel variation in electrification, and we use them to

instrument for the evolution of the actual electricity grid, while controlling for location fixed effects.

Keywords: Electricity, Hydro-power, Agriculture, Productivity, Deforestation, Brazil, Amazon

1 Introduction

The rapid loss of major tropical forest ecosystems has been one of the major environmental disasters of the last century. Nearly 20% of recent global greenhouse gas emissions are attributed to tropical deforestation (Stern, 2008). The vast biodiversity along the Amazon river basin coupled with rapid economic growth makes Brazil the single most important player in the tension between development and environmental sustainability. While there has been a deceleration in the rate of deforestation in Brazil recently, 78,564 km2 of forest cover was lost in the last seven years, and the scale of the problem therefore remains enormous (MMA, 2013).

Deforestation is intricately tied to decisions on land use, and specifically, land for agricultural production in Brazil. While logging is often the proximate cause of land clearing, re-growth occurs in most moist tropical forests. For areas to remain deforested in the longer term, the propensity to convert the land to agricultural use matters most. Agricultural productivity in frontier areas, and the type of farming that gets practiced -– crop cultivation versus cattle grazing – are therefore the key determinants of long-run deforestation. Brazil produces over a hundred billion dollars annually through large-scale crop cultivation, and ranks among the world's three largest producers of sugarcane, soybeans and maize. Crop output increased 365% between 1996 and 2006, and Brazil has been dubbed "the farm that feeds the world" (The Economist, August 26, 2010). Brazil is also the world's largest exporter of beef, with a ten-fold increase in exports during that decade. Many farmers engage in both cultivation and cattle-grazing simultaneously.

Cattle grazing and crop cultivation pose very different risks for deforestation, and this margin of land use decisions will be central to our analysis. Cattle grazing is extremely land intensive with limited use of confinement in Brazil, and the average stocking ratio reported in the 2006 agricultural census was less than 1 head per hectare. In contrast, crop cultivation accounted for only 10.6% of total farm area, but 60% of the value of output in 2006 (IBGE). Increasing crop productivity can therefore have ambiguous effects on deforestation in theory. It has the potential to curb deforestation by inducing land conversion away from grazing and into more intensive cropping, but it could also induce expansion of agriculture into frontier lands.

We first examine these theoretical possibilities using a model in which farmers engage in both land-intensive cattle grazing and in crop cultivation. The farmer faces a factor market constraint that limits growth. Any productivity shock biased towards cropping will induce farmers to switch into cultivation and decrease the land allocation to grazing. The shift away from the activity that is more land intensive decreases overall land use, and benefits native vegetation. On the extensive margin, increased agricultural productivity induces new people to move into farming, which has the opposite effect on deforestation. The overall effect on deforestation is therefore theoretically ambiguous, and this motivates our empirical inquiry.

For empirical identification, we use the impressive expansion of the electricity grid in Brazil during the period 1960-2000 that electrified many frontier areas and farms as a measure of a shock to agricultural productivity. We first document that electricity access increases cropping productivity more than cattle, and then investigate whether production structure and farmers' land use decisions change as a result.

To address the endogeneity issues inherent in infrastructure data (where investment may follow demand), we use the IV estimation strategy developed in Lipscomb et al. (2013). We forecast hydropower dam placement and transmission grid expansion based on exogenous topographic attributes of each location. The forecasting model produces hypothetical maps that show, given the constrained budget of generation plants for each decade, how the electrical grid would have evolved had infrastructure allocation been based solely on cost considerations, ignoring demand-side concerns. The maps isolate the portion of the panel variation in electricity grid expansion that is attributable to engineering cost considerations, and thereby provide exogenous variation in electricity access which we use as an instrument for actual electrification.

Our empirical strategy takes advantage of the fact that Brazil relies almost exclusively on hydropower to meet its electricity needs. 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. The forecasting model includes location fixed effects to control for the fixed geographic attributes, so that the identification comes only from discontinuities in the ranking of different locations' suitability for dam construction, given the decade-specific budget constraints.

We show that electrification improves productivity in crop-based agriculture more than in cattle ranching. One mechanism is that electricity permitted investments in irrigation. We show empirical evidence consistent with the two opposing implications from the theory: (a) the increase in agricultural productivity leads to an expansion in farming and induces frontier land conversion, and (b) land use within farms shifts away from cattle ranching (which is the most environmentally destructive), and into crop cultivation. This allow farmers to retain more native vegetation within rural settlements.

Overall, we estimate that electrification causes a net decrease in deforestation in Brazil. This is because the change in land use patterns within rural settlements (away from grazing and into cultivation) is the more quantitatively dominant of the two opposing effects.

The fact that increased agricultural productivity can protect forests and native vegetation has very important implications for policy. Rainforest protection in Indonesia, Africa and Latin America is hampered by regulators' inability to enforce fines or bans on deforestation. Forest protection measures therefore have to account for the preferences of potential users of the common pool resource, and focus on strategies that are in the economic interest of user groups. Governments and other environmental organizations are increasingly experimenting with approaches such as direct "payments for ecosystem services," or interventions that improve farm productivity (and therefore does not require outside funds, and are viewed as more sustainable).[1]

An assumption underlying strategies aimed at reducing deforestation through increased investment in capital for rural farms is that increased productivity will lead to agricultural intensification which reduces the pressure to clear forests for new land, rather than expand the scope of farming. Our model indicates that this a potentially dangerous strategy, because it has ambiguous effects on land use in theory. Our empirical results indicate that productivity-enhancing strategies did allow for decreased net deforestation in Brazil over the period 1970-2000. Improvements in agricultural productivity appear to be a promising avenue for environmental protection when regulators' capacity to enforce forest protection laws is weak. Our analysis suggests that the net effect on deforestation will depend on the type of activity that gets displaced when agriculture becomes more productive. In Brazil, the proliferation of land-intensive cattle grazing makes improved cultivation beneficial for the environment.

The beneficial environmental effect of the expansion in electricity infrastructure stands in contrast to Pfaff (1999), Cropper et al. (1999) and Cropper et al. (2001), who show that road infrastructure aids deforestation in Brazil and Thailand, respectively. Stavins and Jaffe (1990) find that flood-control infrastructure projects account for 30 percent of forested wetland depletion in the Mississippi Valley by affecting private land use decisions. Our nuanced findings on the opposing effects of infrastructure development on deforestation contribute to a long-standing literature on the non-monotonic relationship, also known as the Environmental Kuznets Curve (EKC), between economic growth and environmental outcomes, starting with Grossman and Krueger (1991, 1995). The existing empirical evidence on the EKC is mixed and mostly based on cross-country regressions, see Foster and Rosenzweig (2003) and Cropper and Griffiths (1994).

Our paper is also related to the literature on technology adoption in agriculture (BenYishay and Mobarak, 2014), (Conley and Udry, 2010). It is most closely related to papers on causes and consequences of irrigation technology in the United States (Hornbeck and Keskin, 2014) and in India (Sekhri, 2011). We also contribute to a rapidly growing literature on the effects of electrification (Dinkelman, 2011; Rud, 2012; Lipscomb et al., 2013)) and other forms of infrastructure (Duflo and Pande, 2007; Donaldson, ming) on development.

The paper is organized as follows: section 2 discusses historical land use in Brazil, the vast

[1]Several developing countries are already beginning to implement payments for environmental services, including Costa Rica, Brazil, Vietnam, and Uganda(Porras, 2012).

growth in the electricity network during the period 1960-2000, and the expansion of the use of irrigation. Section 3 discusses a simple theoretical model which we use to investigate the contrasting impacts of increased agricultural productivity on land use: the increased intensity of agricultural productivity, versus the potential for expansion across increased land area as agriculture becomes more profitable. Section 4 discusses the three key datasets that we use–the Census of Agriculture in Brazil, Electricity Data from various historical archives in Brazil and elevation maps from USGS, and rainfall data. Section 5 discusses our estimation strategy and the instrumental variable technique we employ. Section 6 discusses the empirical results, and section 7 concludes.

2 Background

2.1 Land Use in Brazil: Cattle Grazing and Crop Cultivation

Figure XX shows the trends of land use within farmland in Brazil from 1970 to 2006. Overall, farmland expanded XX percent in the period. Within farmland, the main categories of land use are pastureland, cropland and native vegetation. Pastureland, which is almost entirely use for cattle grazing, declined both in absolute and relative numbers. At the same time, producers have increased the fraction dedicated to crops – both annual and perennial.[2] Finally, the amount of native vegetation has also markedly increased over time. The presence of native vegetation within farmland is a feature of many Latin American countries (REFERENCES) and can be explained by many factors.

The operation of a Brazilian livestock farm in terms of investment and capital stock is very different from a crop farm. In particular, the typical livestock farm uses more land and less capital than a crop farm. Data from the 2006 Census of Agriculture show that, in a typical livestock farm, the value of machinery and equipment per hectare is one-sixth of that of a typical crop farm. Investment per farm is also lower in livestock farms, although an equal fraction of farms in each activity make investments in machinery and equipment. These figures are not surprising when one notes that only 4 percent of cattle farms used confinement in 2006, and that only 0.2 percent of producers pasteurize the milk they sell. In short, the typical cattle grazing farm requires low levels of capital investments within farm gate when compared to crop farming.

One common feature of rural economies in developing countries is the frictions, and ensuing constraints in inputs markets faced by producers. In Brazil, producers typically face constraints in inputs other than land.

2.2 Electrification in Brazil

2.3 Irrigation in Brazil

[2]Not surprisingly, these changes mirror changes in productivity (or yield) gains: whereas gross yields in agriculture have quadrupled over the 1970-2006 period, cattle grazing yields, measured as heads per hectare, only doubled.

3 Conceptual Framework

In this section we build a simple theoretical framework inspired by the salient features of farming and land use in Brazil, with the goal of generating predictions on how a productivity shock in crop cultivation will affect farming choices and deforestation. We focus on farming because that represents the biggest threat to native vegetation in Brazil. Our model allows farmers to engage in both crop cultivation and cattle grazing because these are the two major categories of agricultural activities, as indicated in the previous section and in the agricultural census data.

The economy is endowed with total land of $\bar{H}$ which is initially completely covered by native vegetation. A continuum of individuals reside in this economy, and each decides whether to become a farmer and convert land to agricultural use. These agents only differ by their outside option θ, which is their individual-specific opportunity cost of operating a farm. $\theta \sim \Gamma$, with pdf γ). Profits from farming activities is common across farmers and is denoted Π, and the set of farmers is therefore defined by those with $\theta \leq \Pi$, which has mass $\Gamma(\Pi)$.

Each farm is a tract of land of size H, which is fully covered by native vegetation before farming activities commence. Each farmer can engage in both crop cultivation and cattle grazing, and the areas allocated to each type of activity are denoted H_c and H_g, respectively. We assume that the production functions for the two activities are similar, except that there is a factor which is more useful in crop cultivation, which we will denote K. Electrification, denoted by Ω, improves the productivity of K. Our modeling choice reflects the fact that electrification enhances the productivity of crop cultivation more than cattle grazing. We assume the following forms for the production functions for crops and cattle grazing: $C = \Omega K F(H_c)$ and $G = F(H_g)$, with $F_H > 0$, $F_{HH} < 0$ and $F_H(0) = \infty$.[3]

Land and the factor K can be bought in the market at prices p and r, respectively. Farmers are credit constrained and need to fund their expenditures with capital and land from their own resources, M. We normalize the prices of C and G to 1. Thus, each farmer's problem can be written as:

$$\max_{K,H_c,H_g} \Pi = \Omega K F(H_c) + F(H_g) - rK - p(H_c + H_g) \tag{1}$$

subject to

[3] The factor K only entering the production function for crop but not cattle is merely a modeling simplification. The results we derive only require that electrification benefits crop cultivation relatively more.

$$rK + p(H_c + H_g) \leq M, \tag{2}$$
$$H_c + H_g \leq H. \tag{3}$$

We will focus on the case where the resource constraint (eq. 2) is binding, because the majority of farmers in Brazil are small and medium holders who face some factor market constraints in capital, credit or labor which affects their ability to generate K. Land will therefore not be the limiting factor, and the land constraint (eq. 3) will typically not bind. This focus reflects reality (farming in Brazil expanded into frontier lands that just needed to be cleared and occupied during our period of study), and also makes the model interesting and informative. The credit constraint always binds because the profit function is linear with respect to K and $F_H(0) = \infty$.

Optimal land use and production choices for farmers are characterized by the following set of first-order conditions:

$$\begin{aligned}
\Omega K F_H(H_c) &= (1+\lambda)p & (4)\\
F_H(H_g) &= (1+\lambda)p & (5)\\
\Omega F(H_c) &= (1+\lambda)r & (6)\\
\lambda(rK + p(H_c + H_g) - M) &= 0 & (7)
\end{aligned}$$

where λ is the Lagrange multiplier associated with equation (2)

Equations 4 and 6 imply that $K = F(H_c)/F_H(H_c)$ and thus $dK/dH_c > 0$. Since the factor K is more useful in cultivation than in grazing, both K and the land allocated to cultivation will move in the same direction in this model. K is the limiting factor for farmers, and next we explore the implications of this credit constraint:

$$\left(1 + \frac{r}{p}\frac{dK}{dH_c}\right)\frac{dH_c}{d\Omega} = -\frac{dH_g}{d\Omega} \tag{8}$$

If farmers are credit constrained, and if electrification increases investments in K (and therefore in land allocated to cultivation), then it must decrease the land allocated to cattle grazing.

Equations 5 and 6 determine that:

$$\frac{r}{p}F_{HH}(H_g)\frac{dH_g}{d\Omega} - \Omega F_{HH}(H_c)\frac{dH_c}{d\Omega} = F_H(H_c) \tag{9}$$

Substituting 8 into 9, we can see that $dH_c/d\Omega > 0$:

$$-\frac{r}{p}F_{HH}(H_g)\Big(1+\frac{r}{p}\frac{dK}{dH_c}\Big)\frac{dH_c}{d\Omega} - \Omega F_{HH}(H_c)\frac{dH_c}{d\Omega} = F_H(H_c) \tag{10}$$

Electrification therefore increases the productivity of K and induces farmers to invest in more K. K is useful for cultivation, which increases the land allocated to cultivation. This necessarily leads credit constrained farmers to lower land allocated to cattle grazing, because a larger share of their budget is spent on cultivation.

The net effect on native vegetation within the farm will depend on farmers' total land demand across cultivation and grazing. We define the farmer's total land demand as $H_f = H_c + H_g$, equation 8 can be rearranged to:

$$\frac{dH_f}{d\Omega} = \frac{dH_c}{d\Omega} + \frac{dH_g}{d\Omega} = -\frac{r}{p}\frac{dK}{dH_c}\frac{dH_c}{d\Omega} < 0 \tag{11}$$

The total land demand for all forms of agricultural activities decreases, because farmers have to spend more money on K. In summary, the model predicts that electrification (i.e. increasing the productivity of the limited factor) will: (i) increase use of K, (ii) induce farmers to shift land use from land-intensive cattle grazing to K-intensive cultivation; and (iii) reduce farmers' total land demand;

The net effect of electrification on deforestation will depend not only on intensive-margin changes in land demand within each farm, but also on how the productivity shock induces extensive-margin changes in the decision to enter the agricultural sector. To analyze this net effect, we define the total area of native vegetation as

$$\begin{aligned} H_v &= \bar{H} - \int_{\theta<\Pi} H_f d\Gamma(\theta) \\ &= \bar{H} - H_f\Gamma(\Pi). \end{aligned} \tag{12}$$

The total derivative of the forest with respect to electrification displays two opposing effects:

$$\begin{aligned} \frac{dH_v}{d\Omega} &= \quad -\frac{d\Gamma(\Pi)}{d\Pi}\frac{d\Pi}{d\Omega}H_f \quad -\Gamma(\Pi)\frac{dH_f}{d\Omega} \\ &= \underbrace{-\gamma(\Pi)KF(H_a)H_f}_{<0} \quad \underbrace{-\Gamma(\Pi)\frac{dH_f}{d\Omega}}_{>0} \end{aligned} \tag{13}$$

The first (negative) term shows that the positive productivity shock associated with electrification increases the farm profits and inducse an increase in the number of farmers, which would lead to some deforestation as native vegetation is cleared for new farms. The second term shows that electrification reduces the land demand for each farmer by inducing farmers to shift away from land-intensive cattle grazing activities. The net effect on the forest is therefore theoretically ambiguous; it will depend on the relative magnitudes of the two opposing effects, including the mass of citizens who are on the margin of participation in agriculture. We will examine each of the two (intensive and extensive margin) effects in the data, and also compare the relative magnitudes of these estimated effects to infer the net implication of the productivity shock for deforestation.

The mechanism highlighted in this model makes a few assumptions about the agricultural production function that we can examine in the data, and the model yields a few further predictions that we can also test:

1. Electrification increases productivity of crop cultivation more so than cattle grazing productivity. 2. Electrification should lead to greater investments in capital. 3. Electrification should induce relatively more investment in the types of capital that raise cultivation productivity relatively more. 4. Electrification should induce new farmers to enter agriculture. The number of farms should increase.

4 Data

We combine three datasets in order to observe the impact of the vast expansion of the electricity network across Brazil from 1960-2000 on agricultural productivity, agricultural investments, and deforestation. First, we use county-level data from the Brazilian Census of Agriculture in order to track amount of land under cultivation, agricultural inputs, and total harvests. We use data assembled by Lipscomb et al. (2013) for measures of electricity infrastructure in each decade and an instrumental variable which provides the exogenous variation in electricity access. Finally, we use rainfall data compiled by Matsuura and Willmott (2012). Table 1 presents summary statistics from these datasets.

4.1 Census of Agriculture

Definition of a rural establishment and level of aggregation The Brazilian Census of Agriculture is a comprehensive and detailed source of data on the universe of rural establishments in the country. The definition of a rural establishment is constant across the waves we use, and is similar to what would be commonly thought of as a farm: a continuous plot of land under a single operator, with some rural economic activity – crop, vegetable or flower farming, orchards, animal grazing or forestry. There are no restrictions on the size of the plot, tenure, or market participation. Common lands are excluded from this definition, as are domestic backyards and gardens. Throughout the paper, we refer to a rural establishment simply as a *farm*. We use county-level data from the following 5 waves of the Census of Agriculture: 1970, 1975, 1985, 1996 and 2006.[4] During this period, there were significant changes in the borders and number of Brazilian municipalities. We follow the methodology of Reis et al (2010), who construct minimum comparable geographical areas that are constant over this period, allowing for meaningful comparison across years. We loosely refer to these areas as *counties*.

Outcome variables: Area Three sets of outcome variables are central to our analysis. First is the farm area in each of three land use categories: cropland, pastures, and native vegetation. Together, these three land use categories account for between 81% and 90% of the total land in farms in Brazil during the period 1970-2006. The remaining farm area is bundled in a fourth "other" category, which includes orchards, planted forests, buildings and facilities, water bodies and non-arable land.[5] Cropland excludes area for perennial crops (many of which are in

[4] This selection was made in order to match the other available sources of data. The first wave was carried in 1920. From 1940 to 1970 the Census of Agriculture was decennial. From 1970 to 1985 it was carried in 5-year intervals. The last two waves were carried in 1996 and 2006.

[5] For our purposes in this paper, we explicitly separate planted forests from native forests. The area in planted forests is small, and bundling the tow categories makes no quantitative difference in or results.

orchards) and includes forage-land. Pastures can be either natural or planted.

Outcome variables: Productivity Second, we construct measures to capture farm productivity as well as the productivity of crop farming and cattle grazing separately. We measure farm productivity by their gross production value divided by total farm area (*production per hectare*). Gross production value is the the market value of all goods produced in farms, including production for own consumption. Crop farming productivity is measured in an analogous way: gross crop production value of divided by cropland (*crop production per hectare*). Our main measure of cattle grazing productivity is the farm inventory of cattle heads divided by hectares of pastureland (*heads per hectare*). We also breakdown the total cattle herd into beef cattle and dairy cattle, and measure dairy cattle productivity as milk production per head of dairy cattle.

Outcome variables: Capital and Inputs A third set of outcome variables is related to the capital stock, irrigation and input usage in farms. For capital stock, we use the number of tractor in a country. For irrigation, we have the number of farms that use irrigation as well as the irrigated area within farms. Finally, we use spending on fertilizers and pesticides as measures of input usage.

4.2 Electricity Data

We use the measure of electricity access in counties across Brazil constructed in Lipscomb et al. (2013). The measure is derived from archival tables of the construction date and location of hydropower plants and the related transmission lines. As in Lipscomb et al. (2013), we focus on transmission lines and generation plants as these are the highest cost components of the infrastructure network and the most dependent on geographic costs. Distribution networks are very closely linked with areas where demand for electricity is highest.

The measure of access to electricity infrastructure is generated as follows: Brazil is divided into 33,342 evenly spaced grid points. 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 –it is estimated that on average the distribution networks stretch one-hundred kilometers across. The grid points are then aggregated to the county level, and the electricity access variable is defined as the proportion of electrified grid points within a county.

The instrumental variable for electricity infrastructure is the geographic lowest cost areas for expansion in each decade as proposed by Lipscomb et al. (2013). This instrument is further explained in section 5.1. The instrument is developed using on geographic data collected from the USGS Hydro1k dataset. The Hydro1k dataset is a hydrographically accurate digital elevation map developed from satellite photos of the earth. This data is then matched to each of the 33,342 evenly spaced gridpoints for use in the model, and then predicted access is aggregated to the level of the 2,184 standardized counties across Brazil.

We match Census data to electricity data with a small time lag between the two. As explained in Section 4.2, the electricity data is county-decade panel, not a county-year panel. For that reason, we match the 1970 Census data to the electricity data for the 1960s; the 1975 Census data to the 1970s electricity data; the 1985 Census data to the 1980s electricity data; the 1995 Census data to the 1990s electricity data; and 2006 Census data to the 2000s electricity data. This gives farms a short period of time to react to new electricity access so that we observe the changes resulting from expansion in infrastructure.

4.3 Climate Data

Finally, we use the rainfall data compiled by Matsuura and Willmott (2012) to construct various indicators of drought, dryness and rainfall volatility for each county. This dataset provides monthly precipitation estimates at each node of a 0.5×0.5 degree grid. These estimates are obtained by interpolating data from local weather stations.

To construct indicators of drought, dryness and rainfall volatility, we start by identifying all grid nodes inside each county. If there are less than four nodes with precipitation data inside the county, we then find the four closest nodes to the county's borders. For each county, we then take an weighted average of this set nodes, using the inverse of the distance to county's centroid as weights.

We define rainfall volatility of county c as the standard deviation of the residuals of the following regression:

$$r_{cmy} = \beta_0 + \theta_m + \delta_y + \epsilon_{cmy},$$

where r_{cmy} is rainfall in county c, in month m and year y, θ_m is a month fixed effect and δ_y is a year fixed-effect. In words, we calculate rainfall volatility over and above seasonality and common shocks. We then define high (low) volatility counties as those whose volatility index is above (below) the median.[6]

[6]We calculate other volatility measures, as well as indexes of droughts and dryness. We are still working on results using those other measures, and future versions of this paper should include such results either on its body or in the appendix.

5 Estimation Strategy

In order to identify the impact of access to electricity on deforestation and farm productivity, we use variation in electrification from 1960-2000 and data from the agricultural census on farm productivity and data on deforestation over that period. The principal identification concern in estimating the effect of access to electricity on farm productivity is that demand variables that attract the government to install new electricity infrastructure in some counties will also be related to farm productivity and deforestation. For example, quickly growing nearby cities may increase the demand for electricity, pushing the government to increase the power network in the area, but it could also increase the demand for agricultural products and increase the level of capital investments in agriculture because of high local demand. This would create an omitted variable bias, and we therefore need an instrumental variable which includes only variation exogenous to farm productivity and deforestation.

5.1 Predicting Electricity Expansion Based on Geographic Costs: the design of the Instrument

Our instrument is the predicted electricity availability at each grid point based on minimization of construction cost using only geographic characteristics in each decade. We use the national budget for electricity plants in each decade and predict where these are likely to be placed given where electricity plants and transmission networks have been placed in past decades. In the construction of the instrument, we use only topographic characteristics of the land (flow accumulation and slope in rivers) to estimate likely locations for new electricity access. This instrument is the same as instrument created in Lipscomb et al. (2013).

As described in Lipscomb et al. (2013), there are three key steps to the creation of our instrument: first we calculate the budget for plants in each time period based on actual construction of major dams in each decade across Brazil. Second, we generate a cost variable that ranks potential locations by geographic suitability in each period given the locations of existing dams in each period. Finally, following the prediction on estimated construction site for each dam, we generate an estimated transmission network flowing from the new plants.

The budget of electricity plants is generated based on the actual construction of major electricity plants in Brazil over the period. This allows us to model greater expansion of electricity in years in which the national government decided to expand production of electricity, and reduced expansion in years in which the government budgeted for fewer new plants.

In order to rank the suitability of the different sites, we generate hydrographic variables using the USGS Hydro1k dataset. We generate weights for hydrographic variables using the actual placement of hydropower plants in Brazil (for robustness we have compared these weights to those generated using US hydropower plants, and we arrive at similar results). The cost parameters are derived using probit regressions using an indicator for whether a location has a dam built on it at the end of the sample period (2000) with the topographic measures as the explanatory variables. Steep gradients and high water availability are key factors reducing dam costs.

Following the estimate of placement of new plants, the model predicts transmission lines flowing out from each plant. All plants are assumed to have the same generation capacity, and we make no assumption on demand in various areas, so we make the simplifying assumption that each plant has two transmission substations attached to it. We minimize the cost of the transmission lines based on data and land slope. We then assume that all grid points within 50km of a predicted plant or predicted transmission substation are covered by distribution networks.

In later decades, we take the existing predicted network as given and estimate additional plants and transmission lines as locating in the next lowest cost areas. We then estimate the coverage of electricity access in a county by estimating average coverage of grid points with predicted electricity across the county.

The key potential identification concern related to this instrumental variables estimation strategy would be if the geographic costs for expanding electricity access also affected the productivity of agriculture or the attractiveness of deforesting new areas. While variables like water access and slope in rivers could affect agricultural productivity in a cross-sectional framework, our identifying variation results from variation in whether the cost parameter of a gridpoint is low enough to make it among the low cost budgeted points in a given decade. This non-linearity in threshold chosen gridpoints is different from a simple ranking of lowest to highest cost gridpoints, and creates variation in electricity access which is exogenous to agricultural productivity even though agricultural productivity is affected by water access and land slope.

5.2 Estimation Strategy

We estimate the effect of electrification on the productivity of rural establishments over the period 1960 to 2000 using county-level data. We are interested in running regressions of the form

$$Y_{ct} = \alpha_c + \gamma_t + \beta E_{c,t} + \varepsilon_{ct}, \tag{14}$$

where Y_{ct} is the outcome of interest in county c at time t, α_c is a county fixed-effect, γ_t is a time fixed-effect, and $E_{c,t}$ is the proportion of grid points in county c that are electrified in period t –

that is, $E_{c,t}$ is our measure of actual electricity infrastructure.

The main concern with (14) is that, even controlling for time and year fixed-effects, the evolution of electricity infrastructure is likely to be endogenous to a various factors also affecting the evolution of farm productivity. This causes OLS estimates to be biased.

We therefore use an instrumental variable (IV) approach, making use of the instrument described in Section 5.1. Specifically, we use a 2SLS model where the first stage is

$$E_{ct} = \alpha_c^1 + \gamma_t^2 + \theta Z_{c,t} + \eta_{ct}, \tag{15}$$

where Z_{ct} is the fraction of grid points in county c predicted to be electrified by the forecasting model (relying only on the exogenous variation from the geographic cost variables changing according to the budgeted amount of infrastructure in each decade) at time t. The second stage is

$$Y_{ct} = \alpha_c^2 + \gamma_t^2 + \beta \widehat{E}_{c,t} + \varepsilon_{ct}^2, \tag{16}$$

where $\widehat{E}_{c,t}$ is obtained from the first stage regression (15). Note that both $Z_{c,t}$ and $E_{c,t}$ are constructed by aggregating grid points within the county. Since the number of grid points vary in each county, we weight regressions using county area as weights. In all specifications, we cluster standard errors at the county level due to possible serial correlation.

Our IV strategy corrects the for the bias introduced by the endogenous placement of electricity infrastructure by isolating the supply determinants of the electricity grid evolution. These supply-side determinants are river gradient, water flow and Amazon location. The use of Amazon location in the instrument is subject to criticism, because it may introduce spurious variation in the instrument. We include as controls interactions of Amazon location and year dummies in all our specifications.[7]

[7] Lipscomb et al. (2013) discuss the introduction of these interactions, and present a series of robustness checks.

6 Empirical Results

6.1 First-stage results

Table 2 shows the first-stage results of our main analysis. As explained in section 4, our instrument is based on a engineering model that takes various inputs. Columns (1)-(3) show different specifications controlling directly for some of these inputs. In addition to county fixed-effects, which are included in all specifications in Table 2, Column (1) uses year-fixed effects. The modeled electricity availability is highly correlated with actual electricity infrastructure, and this correlation is significant at the 1 percent level. Column (2) adds Amazon-specific year dummies to flexibly control for the region's time trend, which has significantly differed from that of the rest of the country. The point estimates decreases from from 0.296 to 0.192, but remains significant at the 1 percent level. Column (3) adds interactions of our water flow and river gradient measures with year dummies. The change in the point estimate is negligible. In Columns (4) and (5), we check that both our modeled instrument and measure of electricity infrastructure are indeed correlated with actual electricity provision as captured by the Census of Agriculture. Although the point estimates in Column (4) are lower than those in Columns (1)-(3), the correlation is still significant.

6.2 The effects of electricity on agricultural productivity

Based on the discussion in section 2 and the model presented in section 3, we interpret the arrival of electricity as a positive productivity shock to agriculture and in particular to crop cultivation. The results presented below support our interpretation that the arrival of electricity can be thought of as a productivity shock to crop cultivation.

Table 3 reports the main effects of increasing electricity infrastructure on agricultural productivity. Columns (1)-(3) show respectively the OLS, reduced form and IV estimates when the dependent variable is the log of agriculture production value per hectare of farmland. The IV estimates are larger than the OLS estimates and imply that that a 10 percent increase in electricity availability increases agricultural productivity by 16 percent, and this result is significant at the 5 percent level.

To further understand which activity benefits relatively more from new electricity infrastructure, we analyze separately the effects on crop and cattle grazing productivity. Columns (4)-(6) show results when the dependent variable is the log of crop production value per hectare of cropland. The IV point estimate implies that a 10 percentage-point increase in electricity infrastructure increases crop productivity by 20 percent, and this effect is significant at the 1 percent level. The high impact of electricity on crop productivity is mirrored by a low impact on cattle grazing productivity. Columns (7)-(9) show the effects of electricity on the number of cattle heads per hectare of pastureland, or *heads per hectare*. The IV estimate in column (9) implies that a 10 percentage-point increase in electricity leads to a 0.07 increase in heads per hectare, a 6.5 percent effect on the mean. Not only this is a lower impact than that for crop cultivation, it is not statistically significant at conventional levels.

In sum, the arrival of electricity infrastructure in a county significantly increases crop productivity, but not cattle grazing productivity. This result gives us confidence Not only this result is interesting in itself, it also gives us confidence in our model, in particular where it assumes that electricity is a positive productivity shock to crop cultivation productivity.

6.2.1 Identification concerns

As explained in section 4, the instrument uses cross-sectional variation from geographical factors, and time-series variation from the national budget for construction of electricity infrastructure and suitability ranks that introduce discontinuities on the order in which new infrastructure is built. To mitigate concerns that any of this variation is not good variation for dealing of the endogeneity problems of grid placement, Table 4 presents results of a series of sensitivity tests where we use all possible combinations of our instrument's components as explicit controls in the second-stage regressions, on top of county fixed-effects and decade dummies. Each row of Table 4 reports a different specification of a 2SLS regression where the dependent variable is crop productivity, in column (1), or cattle grazing productivity, in column (2). We also report the corresponding first-stage statistics in columns (3) and (4). As can be seen, both the first stage and the main result of Table 3 – that electricity affects crop cultivation productivity, but not cattle grazing productivity – survive all the different specifications.

6.3 Changes in Land Use and Production Decisions

Given that electricity increases overall agricultural productivity, it is natural to expect that it will lead to an expansion of farmland, as producers' will want to do more agriculture. But electricity also changes the relative productivities of crop cultivation and cattle grazing, which implies that producers should shift away from cattle grazing into crop cultivation. In this section we explore in more details these changes in producers' decisions.

Table 5 shows the effects of electrification on land allocation within farms. Columns (1) and (2) show how farmland expands following more electricity infrastructure. The IV estimate implies that the share of farmland in the typical county increases by 1.3 percentage points following a 10 percentage point increase in electricity infrastructure. This coefficient however is not precisely estimated and hence is not statistically significant. In the remaining columns we look at changes in the shares of pastureland, cropland and area in native vegetation within farms. In Columns (3) and (4), we see that the share of pastureland in the county's farmland decreases with electricity infrastructure. The IV estimate in column (4) implies that the share of pastures in farmland decrease by 4.5 percentage points following a 10 percentage point increase in electricity, an effect of 9.5 percent on the mean. Columns (5) and (6) show the same analysis for total cropland. The IV estimate in column (6) implies that the share of cropland decreases by 2.4 percentage points, a 11.5 percent effect on the mean. In columns (7) and (8) we look at changes annual cropland. The effect of new electricity infrastructure is small in magnitude and not statistically significant. Finally, in columns (9) and (10) we look at the share of farmland that remains in native vegetation. The IV estimate implies that a 10 percentage-point increase in electrification induces producers to increase native vegetation within rural establishments by 4.7 percentage points, a mean effect of 29 percent.

The results so far suggest that, by increasing crop farming productivity relative to cattle grazing productivity, the arrival of electricity induce producers to reduce the share of land they allocate to pastures relative to (annual) cropland. What is more surprising is the large effect of electricity on the share of native vegetation within farms. Such large effect raises two questions. First, assuming that land in native vegetation brings no private benefits for producers, why would producers ever choose to keep land in native vegetation? As outlined in our model of land use choices, producers may face constrains on input factors other than land – capital and/or labor. Since crop farming requires more of these inputs, the reduction in land demand for cattle grazing can be larger than the increase in land demand for crop farming. When the intensification effect is larger than the expansion effect, more land is left in native vegetation[8].

Second, what is the effect of electricity in overall native vegetation, not only within farmland?

[8] Of course, there may be other mechanisms. We discuss these in section X.

Table 5 reveals two opposing effects. On one hand, new infrastructure induces an expansion of farmland, which potentially has negative effects on native vegetation *outside* farms. On the other hand, there is a direct positive effect on native vegetation *inside* farms. To be able to calculate the net effect, we would ideally have data on native vegetation outside farms. To the best of our knowledge there are no countrywide reliable sources for most of our period of analysis.[9] We therefore need to make assumptions on what was the state of native vegetation outside farms prior to the arrival of electricity. Assuming that all non-farmland is covered (not covered) with native vegetation yields the lower (upper) bound of 0.22 (0.35). That is, a 10 percentage point increase in electricity infrastructure increases the share of native vegetation in the typical county by 2.2 – 3.5 percentage points. See the appendix for details on how to calculate this estimate from the numbers presented in Table 5.

As a robustness check, we lag land use outcomes forward by one decade (results in the Appendix Table 8. Results remain largely unchanged, suggesting that these are not just short-run effects.

6.4 Testing mechanisms and other model predictions

We now empirically evaluate our model's predictions to build confidence that it can explain the mechanisms underlaying our results. First, in our model there are two opposing forces following a productivity shock to agriculture productivity – an intensification and one expansion effects. We argue that these opposing effects come from different groups of producers, as it would be inconsistent for one single group of producers to display both forces. The mechanism we outline is one where new, incoming operators open new farms (or, equivalently, fewer operators leave). In Table 6, columns (1)-(2) show a large effect on the number of farms above 10 hectares in a county. We exclude very small farms from the dependent variable for conceptual reasons.[10] Our model's prediction is about new operators attracted by an increase in productivity. While our model's prediction is silent on farm size, very small farms are typically operated by families for subsistence, and therefore do not fit into our model.

Second, we test one important link between electricity and agricultural productivity – irrigationI. One of our model's implications is that producers respond to an increase in the availability of electricity by making crop-related investments. Irrigation is a strong candidate, as

[9]By design, the Census of Agriculture collects farmland data. Good countrywide remote sensing data is available starting in late 1990's and early 2000's.

[10]While farms of 10 hectares may be considered large for some countries, like Bangladesh or India, they are considered small in Brazil where the average farm size ranged from 60 to 73 hectares in our sample period. For example, land redistribution programs grant no less than 5 hectares for a family to produce at subsistence levels. Depending on the county, the estimated minimum size of plot of land for subsistence may be as large as 100 hectares.

explained in section 2. Columns (3) and (4) show that both the number of farms as well as irrigated farm area grow substantially with an increase in electricity infrastructure. A 10 percentage point increase increase in electricity leads to 27 percent more farms with irrigation and a 70 percent increase in irrigated land.

More investments in capital Next, Table 6 presents the effects of electrification usage of inputs. In the IV specification, a 10 percent increase in electrification leads to a 53 percent in crease in fertilizer spending (Column (2)), and a 28.7 percent increase in pesticides (Column (4)), and both effects are significant at the 1 percent level. Electrification also leads to more tractors being used, as shown in columns (5) and (6).

6.5 Other mechanisms by which electricity could be linked to forests

Demand for Forest products One alternative explanation for the positive link between electricity and forests, is through a rise in demand for forestry products induced by an increase in income.[11] Foster and Rosenzweig (2003) argue that such demand mechanism was central to explain the positive association between income and forest in India, as well as in a panel of countries. One important condition for this mechanism to be captured empirically is that local demand for forestry products must be met by local supply. Thus, in their panel of countries Foster and Rosenzweig (2003) find that a positive association between income and forest growth for closed economies – Brazil included – but not for open economies. Brazil being a closed economy, we as the question: did the shift in land use toward forests come from increases in demand for forest products?

We answer this question in Table 7. If consumption of forest products increases, production should also increase. In columns (1) and (2) the dependent variable is the log of the total value of forestry goods produced. Both the OLS and IV estimates are negative, and the IV estimate is not statistically significant, indicating that production of forestry products does not increase with electricity, despite the increase in native vegetation documented in Table 5. Forestry goods however are very heterogeneous, ranging from wild fruits to timber. In columns (3) and (4) we focus on the production of wood-related products – fuelwood, charcoal and timber. The IV estimate is now positive, but not statistically significant. In columns (5) and (6) we ask whether producers make a more intensive use of the forests in their property — a natural thing to do when faced with rising demand for forestry products – and use the log of the production value of forestry produces per hectare of forest area. The negative OLS and IV estimates suggest that the rise in forest area within farmland outpaces their direct economic exploration. Finally, we

[11](Lipscomb et al., 2013) find positive links between electricity and income.

ask whether producers actively plant more forests, presumably to meet demand for products that cannot be produced with native species, and use is the share of planted forests in farmland as the dependent variable in columns (7) and (8). Both the OLS and IV estimates are small in magnitude and non-significant. To sum up, we find no evidence that the demand channel to be driving the growth of native vegetation in Brazil for the period we analyze.

7 Conclusion

Table 1: Sample Descriptive Statistics

	Mean	Std. Dev.
Electricity Infrastructure	0.75	0.40
Modeled electricity instrument	0.60	0.47
Fraction of Farms with Irrigation	0.06	0.11
Fraction of Farms with Electricity	0.34	0.33
Average farm size (ha)	94.32	167.01
Fraction of county area in...		
Farmland	0.72	0.31
Pastures	0.37	0.24
Native Vegetation	0.11	0.10
Cropland	0.17	0.19
Productivity variables		
Production Per Hectare (log)	9.78	3.34
Crop Production per Hectare (log)	3.97	3.37
Heads of Cattle per Hectare	1.13	2.99
Input usage		
Spending in Fertilizers Per Hectare (log)	1.45	2.98
Spending in Pesticides Per Hectare (log)	0.77	2.26
# of Tractors (log)	3.94	1.90
Number of AMCs	2,181	
Number of observations	10,905	

Notes: Monetary variables measured in thousands of reais in 2002.

Table 2: First-Stage Results

Dependent Variable	Electricity Infrastructure			% of Farms with Electricity	
	(1)	(2)	(3)	(4)	(5)
Modeled electricity availability	0.296***	0.192***	0.190***	0.107***	
	[0.0475]	[0.0441]	[0.0444]	[0.0356]	
Electricity Infrastructure					0.104***
					[0.0227]
Year FE?	Yes	Yes	Yes	Yes	Yes
Jungle × year dummies	No	Yes	Yes	Yes	Yes
Water flow × year dummies	No	No	Yes	No	No
River gradient × year dummies	No	No	Yes	No	No
Observations	10,905	10,905	10,905	10,894	10,894
Mean dep. var.	0.750	0.750	0.750	0.348	0.348
F-stat	38.8	18.9	18.3		
p-value	0.000	0.000	0.000		

Notes: Standard errors clustered at county level in brackets. All specifications include county fixed effects and year fixed effects, and use county area weights.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: The Effects of Electricity on Agricultural Productivity

	log Production Per Hectare ($)			log Crop Production Per Hectare ($)			Heads Per Hectare		
	(1) OLS	(2) Reduced Form	(3) IV	(4) OLS	(5) Reduced Form	(6) IV	(7) OLS	(8) Reduced Form	(9) IV
Electricity Infrastructure	0.191**		1.662**	0.279***		2.009***	0.158		0.733
	[0.0964]		[0.730]	[0.0715]		[0.593]	[0.214]		[0.854]
Instrument		0.318**			0.385***			0.140	
		[0.152]			[0.103]			[0.182]	
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Jungle x year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,892	10,892	10,892	10,879	10,879	10,879	10,878	10,878	10,876
R^2	0.973	0.973	0.967	0.976	0.976	0.971	0.411	0.411	0.018
Mean dep. var.	9.78	9.78	9.78	3.97	3.97	3.97	1.13	1.13	1.13

Notes: Standard errors clustered at county level in brackets. All specifications include county fixed effects. The dependent variable in columns (1)-(3) is the log of total farm production value divided by total farmland. The dependent variable in columns (4)-(6) is the log of total crop production value divided by total cropland. The dependent variable in columns (7)-(9) is the number of cattle heads per hectare of pastureland.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Sensitivity Analysis by Directly Controlling for Geographic Factors in the Second Stage

Specification (description of control set added to RHS)	(1) log Crop Production Per Hectare	(2) Heads Per Hectare	(3) First stage
1. Water flow × decade budget	2.035*** [0.462]	-0.216 [0.659]	0.295*** [0.0426]
2. River gradient × decade budget	2.024*** [0.463]	-0.219 [0.673]	0.291*** [0.0410]
3. Amazon dummy × decade budget	2.327*** [0.777]	-0.466 [1.014]	0.150*** [0.0368]
4. Water flow × decade budget and Amazon dummy × decade budget	2.345*** [0.788]	-0.468 [0.980]	0.149*** [0.0368]
5. River gradient × decade budget and Amazon dummy × decade budget	2.392*** [0.795]	-0.477 [1.053]	0.149*** [0.0369]
6. River gradient × decade budget and water flow × decade budget	2.030*** [0.464]	-0.218 [0.660]	0.292*** [0.0410]
7. River gradient × decade budget, water flow × decade budget, and Amazon dummy × decade budget	2.459*** [0.826]	-0.487 [1.004]	0.146*** [0.0371]
8. Water flow × year dummies	2.032*** [0.462]	-0.208 [0.654]	0.296*** [0.0426]
9. Amazon dummy × year dummies	2.009*** [0.593]	0.733 [0.854]	0.192*** [0.0395]
10. River gradient × year dummies	2.005*** [0.462]	-0.211 [0.650]	0.289*** [0.0409]
11. Water flow × year dummies and Amazon dummy × year dummies	2.014*** [0.601]	0.772 [0.828]	0.190*** [0.0396]
12. River gradient × year dummies and Amazon dummy × year dummies	2.037*** [0.596]	0.798 [0.904]	0.193*** [0.0394]
13. Water flow × year dummies and river gradient × year dummies	2.015*** [0.464]	-0.207 [0.642]	0.291*** [0.0408]
14. River gradient × year dummies, water flow × year dummies, and Amazon dummy × year dummies	2.085*** [0.617]	0.917 [0.908]	0.190*** [0.0397]
15. Quartic suitability rank × year dummies	1.989*** [0.557]	0.392 [0.718]	0.209*** [0.0392]

Notes: Standard errors clustered at county level in brackets. Each row represents a different sensitivity test to the inclusion of controls which are used in the construction of the instrument. All specifications include county fixed effects. The dependent variable in column (1) is the log of gross crop production value divided by cropland. The dependent variable in column (2) is the number of cattle heads per hectare of pastureland. Column (3) reports the first-stage coefficient associated with the instrument, and its clustered standard error.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: The Effects of Electricity on the Allocation of Land within Farms

Land Use Category	$\frac{\text{Farmland}}{\text{County Area}}$		$\frac{\text{Pastures}}{\text{Farmland}}$		$\frac{\text{Cropland}}{\text{Farmland}}$		$\frac{\text{Native Vegetation}}{\text{Farmland}}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Electricity Infrastructure	0.014	0.13	0.034*	−0.45**	−0.023**	−0.0062	0.022	0.47***
	[0.014]	[0.15]	[0.020]	[0.18]	[0.011]	[0.063]	[0.022]	[0.16]
Observations	10,905	10,905	10,894	10,894	10,894	10,894	10,894	10,894
Mean dep. var.	0.70	0.70	0.48	0.48	0.16	0.16	0.16	0.16

Notes: Standard errors clustered at county level in brackets. All specifications use county area weights and include county fixed effects, year fixed effects, and Amazon-year dummy interactions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Testing other Model Predictions: IV Estimates

	Number of Farms (log)		Irrigation		Capital and Inputs		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\geq$ 10 ha	$\geq$ 20 ha	Number of Farms (log)	Irrigated Area (log)	Number of Tractors (log)	Fertilizer Per Ha. (log)	Pesticids Per Ha.
Electricity Infrastructure	1.424*	1.831**	2.715**	7.060***	4.071**	3.655***	2.406**
	[0.789]	[0.881]	[1.228]	[2.732]	[1.656]	[1.370]	[1.035]
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Jungle x year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,894	10,894	10,905	10,905	10,905	10,894	10,894
Mean dep. var.	6.44	6.15	3.01	4.14	3.94	2.23	1.56

Notes: Standard errors clustered at county level in brackets. All specifications include county fixed effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: The Effects of Electricity on Forestry Production

	log Forestry Production Value		log Production Value of Wood Products		log Forestry Production Value Per Hectare of Forest		Share of Planted Forests on Farmland	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Electricity Infrastructure	-0.279^{*}	-0.469	-0.117	0.236	-0.381^{**}	-3.747^{**}	-0.00315	-0.0263
	[0.153]	[1.690]	[0.182]	[1.595]	[0.155]	[1.768]	[0.00234]	[0.0162]
Observations	10,905	10,905	10,905	10,905	10,876	10,876	10,894	10,894
Mean dep. var.	8.953	8.953	4.229	4.229	0.161	0.161	0.0176	0.0176

Notes: Standard errors clustered at county level in brackets. All specifications use county area weights and include county fixed effects, year fixed effects, and Amazon-year dummy interactions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: The Effects of Electricity on the Allocation of Land within Farms: Long Run Effects

Land Use Category	$\frac{\text{Farmland}}{\text{County Area}}$		$\frac{\text{Pastures}}{\text{Farmland}}$		$\frac{\text{Cropland}}{\text{Farmland}}$		$\frac{\text{Native Vegetation}}{\text{Farmland}}$	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Electricity Infrastructure	-0.016	0.14	0.0056	-0.36^{**}	-0.017	-0.022	0.047^{**}	0.43^{***}
	[0.011]	[0.14]	[0.018]	[0.15]	[0.011]	[0.072]	[0.020]	[0.16]
Observations	8,724	8,724	8,713	8,713	8,713	8,713	8,713	8,713
Mean dep. var.	0.72	0.72	0.48	0.48	0.15	0.15	0.15	0.15

Notes: Standard errors clustered at county level in brackets. All specifications use county area weights and include county fixed effects, year fixed effects, and Amazon-year dummy interactions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8 Appendix

8.1 Calculating the effect of electricity on native vegetation

Column (10) of Table 5 gives us the effect of electricity on native vegetation *inside* farms. Because we do not have explicit data on native vegetation *outside* farms, we need one assumption to be able to back out the effect of electricity on overall native vegetation. To see this, first note that $V(e) = V_I(e) + V_O(e)$, where V is total native vegetation, V_I denotes native vegetation inside of farms, and V_O denotes native vegetation outside of farms. Denoting county area by C, we are interested in

$$\frac{\partial}{\partial e}\frac{V(e)}{C} = \frac{\partial}{\partial e}\frac{V_I(e)}{C} + \frac{\partial}{\partial e}\frac{V_O(e)}{C} \tag{17}$$

From the numbers in Table 5 we can back out the term $\frac{\partial}{\partial e}\frac{V_I(e)}{C}$ using the chain rule. But since we do not have data on $\frac{V_O}{C}$, we must make one assumption to fully recover the effect of electricity on overall native vegetation. Note that we can write $V_O(e) = k(C - F(e))$, that is, native vegetation outside of farms is a fraction k of the county area (C) that is not in farms ($F(e)$). Plugging that on the second term on the RHS of (17),

$$\frac{\partial}{\partial e}\frac{V_O(e)}{C} = \frac{\partial}{\partial e}\frac{k(C - F(e))}{C} = -k\frac{\partial}{\partial e}\frac{F(e)}{C} \tag{18}$$

We have $\frac{\partial}{\partial e}\frac{F(e)}{C}$ from column (2) of Table 5. The only information we do not have is k, the fraction of the county area outside of farms that is in native vegetation. Assuming $k = 1$ ($k = 0$) gives us a lower (upper) bound on the effect of electricity on overall native vegetation.

To calculate $\frac{\partial}{\partial e}\frac{V_I(e)}{C}$, note that the IV coefficient in column (10) of Table 5 gives us $\frac{\partial}{\partial e}\frac{V_I(e)}{F(e)}$ which, by applying the chain rule, is $\frac{\frac{\partial V_I}{\partial e}\cdot F(e) - V_I(e)\cdot\frac{\partial F}{\partial e}}{F(e)^2}$. Solving for $\frac{\partial V_I}{\partial e}$ and dividing both sides by C yields

$$\frac{\partial}{\partial e}\frac{V_I}{C} = \frac{\partial}{\partial e}\frac{V_I(e)}{F(e)}\cdot\frac{F(e)}{C} + \frac{V_I}{F(e)}\cdot\frac{\partial}{\partial e}\frac{F(e)}{C} \tag{19}$$

Plugging equations (18) and (19) into (17),

$$\frac{\partial}{\partial e}\frac{V(e)}{C} = \frac{\partial}{\partial e}\frac{V_I(e)}{F(e)}\cdot\frac{F(e)}{C} + \left(\frac{V_I}{F(e)} - k\right)\cdot\frac{\partial}{\partial e}\frac{F(e)}{C} \tag{20}$$

If we evaluate this derivative at the sample means, Table 5 gives all the terms on the RHS of this equation except for k. The effect of electricity on the share of farmland in native vegetation, $\frac{\partial}{\partial e}\frac{V_I(e)}{F(e)}$ is 047. The fraction of the county area in farmland, $\frac{F(e)}{C}$ is 0.70 for the typical municipality. The fraction of farmland in native vegetation is $\frac{V_I(e)}{F(e)}$ is 0.16 for the typical municipality. Finally, the effect of electricity on farmland 0.13. Assuming $k = 1$ gives a lower (upper) bound of 0.22

(0.35). That is, a 10 percentage point increase in electricity infrastructure increases the share of native vegetation in the typical county by 2.2 – 3.5 percentage points.

References

BenYishay, A. and Mobarak, A. M. (2014). Social Learning and Communication. NBER Working Papers 20139, National Bureau of Economic Research, Inc.

Conley, T. G. and Udry, C. R. (2010). Learning about a new technology: pineapple in Ghana. *American Economic Review*, 100(1):35–69.

Cropper, M. and Griffiths, C. (1994). The Interaction of Population Growth and Environmental Quality. *American Economic Review*, 84(2):250–54.

Cropper, M., Griffiths, C., and Jyotsna (2001). Predicting Spatial Patterns of Deforestation: The Role of Roads and Protected Areas in North Thailand. *Land Economics*, May.

Cropper, M., Griffiths, C., and Mani, M. (1999). Roads, Population Pressures, and Deforestation in Thailand, 1976-1989. *Land Economics*, 75 (1):58–73.

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

Donaldson, D. (forthcoming). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*.

Duflo, E. and Pande, R. (2007). Dams. *The Quarterly Journal of Economics*, 122(2):pp. 601–646.

Foster, A. D. and Rosenzweig, M. R. (2003). Economic growth and the rise of forests. *The Quarterly Journal of Economics*, 118(2):pp. 601–637.

Grossman, G. M. and Krueger, A. B. (1991). Environmental impacts of a north american free trade agreement. *National Bureau of Economic Research Working Paper 3914, NBER, Cambridge MA.*

Grossman, G. M. and Krueger, A. B. (1995). Economic growth and the environment. *Quarterly Journal of Economics*, 110:353–377.

Hornbeck, R. and Keskin, P. (2014). The historically evolving impact of the ogallala aquifer: Agricultural adaptation to groundwater and drought. *American Economic Journal: Applied Economics*, 6(1):190–219.

Lipscomb, M., Mobarak, A. M., and Barham, T. (2013). Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil. *American Economic Journal: Applied Economics*, 5(2):200–231.

Matsuura, K. and Willmott, C. (2012). Terrestrial precipitation: 1900-2010 gridded monthly time series (1900 - 2010) (v 3.01 added 6/14/12). University of Delaware. http://climate.geog.udel.edu/ climate/.

MMA (2013). Sistema Nacional de Unidadades de Conservação, SNUC. Database, Ministério do Meio Ambiente.

Pfaff, A. (1999). What Drives Deforestation in the Brazilian Amazon? Evidence from Satellite and Socioeconomic Data. *Journal of Environmental Economics and Management*, 37(1):26–43.

Porras, I. (2012). Payments for environmental services in costa rica: from rio to rio and beyond. *iied Briefing*.

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

Sekhri, S. (2011). Public provision and protection of natural resources: Groundwater irrigation in rural india. *Amercian Economic Journal: Applied Economics*, 3(4):29–55.

Stavins, R. N. and Jaffe, A. B. (1990). Unintended Impacts of Public Investments on Private Decisions: The Depletion of Forested Wetlands. *The American Economic Review*, 80(3):pp. 337–352.

Stern, N. (2008). The economics of climate change. *American Economic Review*, 98(2):1–37.