Electricity, Agricultural Productivity, and Deforestation: Evidence from Brazil*

Dimitri Szerman
Universität Mannheim

Juliano Assunção
PUC-Rio

Molly Lipscomb
University of Virginia

Ahmed Mushfiq Mobarak
Yale University

September 18, 2024

Abstract

We examine the impact of improved access to electricity on deforestation through the incentives it creates for farms to switch from cattle grazing to croplands. We generate a model which elucidates the conditions under which increases in agricultural productivity can increase the incentives to intensify rather than expand land use and examine the predictions of that model using county-level data from five waves of the Brazilian Census of Agriculture and satellite-based measures of land use. We estimate the impact of rural electrification in Brazil from 1960-2000 and the resulting increases in agricultural productivity. We show that locations suitable for hydropower generation experienced improvements in crop yields incentivizing credit-constrained farmers to shift away from land-intensive cattle-grazing and into cropping. As a result, agricultural land use declined, more native vegetation was protected, and these effects persist 25 years later in both census and satellite data. Brazil's deforestation rate would have been almost twice as large between 1970 and 2000 without the increase in agricultural productivity that resulted from electrification. The conservation benefits of electrification are comparable to prominent forest conservation policies implemented in Brazil.

*We thank Arthur Bragança, Francisco Costa, Taryn Dinkelman, Teevrat Garg, Torfinn Harding, Seema Jayachandran, Dana Kassem, Melanie Morten, Richard Hornbeck, Terence Johnson, Alex Pfaff, Eustáquio Reis, Mark Rosenzweig, Edson Severnini, Jillian Stallman, Joe Shapiro, Gabriel Ulyssea, and seminar participants at Stanford University, the NBER Summer Institute (Energy and Environmental Economics), Yale University, PUC-Rio, Duke University, PUC-Chile, Universidade de São Paulo, FGV/EPGE in Rio, FGV/EESP in São Paulo, Universidade Federal do Rio de Janeiro, UC–Santa Barbara, UC-San Diego, University of Zurich, the Econometric Society Summer European Meeting, Association of Environmental and Resource Economics (AERE) Annual Meetings, Barcelona Summer Institute, CESIfo Energy and Climate Area Conference, NEUDC, and World Bank for valuable comments and discussions. We also thank Steven Helfand for kindly sharing data from the Brazilian Census of Agriculture.

1 Introduction

The rapid loss of tropical forests is one of the most important environmental disasters of the last century. Tropical deforestation is the second largest source of anthropogenic greenhouse gas emissions and severely impacts the world's biodiversity and water regimes (Smith et al., 2014; Dasgupta, 2021). Around 90 percent of forest loss in the tropics is from clearing to make space for crops and cattle (Curtis et al., 2018), which highlights a fundamental tension between conservation and development goals (Frank and Schlenker, 2016). Improving agricultural productivity is thought to be necessary for economic development (World Bank, 2007), but whether increased productivity would exacerbate or alleviate the trade-off between production and conservation is an open question.

This paper posits a new mechanism through which improvements in agricultural productivity could decrease deforestation. Our theory is inspired by the institutional details of agriculture in Brazil, one of the world's largest producers of agricultural commodities. Farms in Brazil engage in two activities that differ in their factor intensities – "cattle grazing" which is land-intensive, and "crop cultivation" which is more capital-intensive. Productivity shocks biased toward cropping induce credit-constrained farmers to switch to cultivation and decrease the land allocated to grazing. The shift away from the land-intensive activity decreases overall land demand for agriculture, ultimately protecting forests. However, increased agricultural productivity may induce entry into farming on the extensive margin, which makes the overall effect on deforestation theoretically ambiguous. This motivates our empirical inquiry.

We explore these dynamics of agricultural land use and deforestation taking advantage of the large-scale expansion of the electricity grid into rural areas in Brazil during the period 1960-2000, which improved agricultural productivity. Electricity directly benefits agricultural production by enabling farmers to use techniques and equipment that would otherwise be more costly or infeasible to implement.

To address the endogeneity of infrastructure placement, our identification is based on an instrument leveraging geographical differences in construction costs of generation plants and changes in the impact of these costs over time. This strategy, developed by Lipscomb et al. (2013), takes advantage of the fact that most of the electricity expansion in Brazil was based on hydropower generation, the costs of which depend on topographic factors such as water flow and river gradient.[1] This allows us to isolate the variation in grid expansion in Brazil that is attributable to exogenous cost considerations (holding fixed the geographic attributes of each location) and use it as an

[1] An analogous strategy has been used by the literature that estimates the economic effects of roads (see e.g., Faber, 2014; Burgess et al., 2015; Baum-Snow et al., 2017; Banerjee et al., 2020).

instrument to estimate the effects of farm electrification.

We construct a decennial panel dataset of Brazilian counties from 1960 to 2000, combining (a) historical measures of electricity infrastructure to reconstruct the evolution of the electrical grid by decade (Lipscomb et al., 2013), (b) five waves of the Brazilian Census of Agriculture to track farmland expansion, land use within farms, agricultural output, and input use, and (c) satellite images to characterize land use both inside and outside farms.

We first document – using data from multiple rounds of the census of agriculture – that farm electrification improved crop productivity but not the productivity of cattle grazing. We find that the expansion of the electricity grid led to an increase in agricultural productivity and a slowdown of deforestation in Brazil over the period 1970-2006. There is a 14 percentage point increase in farm electricity access in a county when the infrastructure arrives, allowing farmers to move away from land-intensive cattle grazing and shift into cropping. This protects forests by decreasing agricultural land use overall. Our estimates indicate that a 10 percent increase in a county's electricity availability increases the total proportion of land covered by native vegetation by 3.9 percentage points. Brazil's forest cover decreased by about 7 percentage points between 1985 and 2006 (from 76% to 69%), which means that without the increase in agriculture productivity brought about through the electrification of rural Brazil, the rate of deforestation could have been 50% larger.

To understand the mechanism underlying these results, we use data from the agricultural census to document how agricultural practices change once farms are electrified. Electrification leads to increased investment in crop-related capital goods that benefit from electricity access, such as grain storage facilities (which require humidity control and temperature control) and irrigation equipment. Crop yields increase as a result. The relative shift towards cropping is also evident in increased investment in plows and planting and harvesting machines. In contrast, access to electricity has no effect on cattle grazing productivity. Land allocated to crop production expands while land allocated to cattle grazing contracts. Since cattle grazing is land-intensive, the increased land-use in crops is not sufficient to offset the decline in pastures.

We also use the data to interrogate a few critical assumptions made in the theoretical model. We use rainfall shocks to show that many Brazilian farmers are indeed credit-constrained, and data on bank presence to show that our land use results are driven by the more constrained farmers. We use longer-run data to show that the changes in land use are not short-lived; they are evident even 25 years after the productivity shock. Satellite imagery also confirms these patterns.

Our primary contribution is to the longstanding academic and policy debates regarding the link between agricultural productivity and deforestation. A prominent article published in *Science*, Phalan et al. (2016), begins,

> "One potential way to reduce [environmental impacts from agriculture] is to increase food production per unit area on existing farmland, so as to minimize farmland area and to spare land for habitat conservation"

This yield intensification effect is sometimes referred to as "Borlaug's hypothesis". It is supported by Foster and Rosenzweig (2003) which shows that income increases, particularly income gains from agricultural productivity increases, lead to decreases in deforestation. Abman et al. (2020) shows that an agricultural extension program aimed at increasing agricultural productivity reduced local deforestation by 14% in Uganda under similar factor market constraints. Abman and Carney (2020) find that by increasing agricultural yields, a fertilizer subsidy program reduced pressure on expansion of farms, and reduced deforestation in Malawi, an environment in which credit constraints are also known to hold (Giné et al., 2012). We add to the literature by showing that electrification contributed to the transformation of agriculture from large grazing lands to higher-intensity cropping in Brazil. Our findings contribute to the understanding of the mechanisms through which deforestation occurs: the decreases in deforestation from increased agricultural productivity found in these papers are dependent on a constraining factor such as low access to credit.

On the other hand, improving agricultural productivity could instead incentivize individual producers to *expand* farming in order to increase profits and put greater pressure on forests, an argument that has been labeled "Boserup's hypothesis" (Angelsen and Kaimowitz, 2001). Improvements in the transport infrastructure, reducing the transport costs of agricultural goods, may be an example of this (Pfaff, 1999; Cropper et al., 1999, 2001; Asher et al., 2020; Souza-Rodrigues, 2019). Increased income may also lead to deforestation through increases in demand for land-intensive goods; for example, Alix-Garcia et al. (2013) shows that income transfers from Progressa lead to increased deforestation. There is a long-standing "environmental Kuznets curve literature on the trade-offs between economic development and environmental conservation concerns (Grossman and Krueger, 1991, 1995).

We bring to this debate a plausible theoretical mechanism by which Borlaug's hypothesis could materialize when farmers are credit-constrained, and then devise an empirical test of this mechanism. It is important to understand whether productivity improvements in agriculture offer a plausible path out of the fundamental tension between economic development and environmental conservation goals, because different world regions have followed alternative pathways to agricultural output growth. Figure 1 illustrates how South Asia followed an *intensification* path, whereby most of the gains in food production stemmed from increasing yields. At the other extreme, Sub-Saharan countries followed the *extensification* path in which increases in food required expanding into new agricultural land. Latin American agriculture encroached on new land until the 1980s, and then shifted towards intensification.

Our contribution is to show a channel through which the intensification of farming leads to decreased deforestation: differential improvements in crop productivity relative to land-intensive cattle production decreases land demand from farms in the presence of credit constraints. In related work Bustos et al. (2016) evaluates the impact of increases in agricultural productivity on economic growth using an instrumental variable based on the suitability of different agricultural lands to the introduction of genetically modified seeds. They find that the seeds had strong impacts on agricultural productivity, and this increased agricultural productivity led to increased manufacturing productivity. The introduction of genetically modified seeds in Brazil in 1996 post-dates most of the variation in electricity access in our sample, so this increase in growth from genetically modified seeds was not related to the productivity increases from electrification. The reduction in deforestation in our context is primarily related to switching in agricultural production as a result of *differential* improvements in crop relative to cattle grazing. This type of switching to optimize productivity is similar to impacts from irrigation technology in the United States (Lewis and Severnini, 2020; Hornbeck and Keskin, 2014) and in India (Sekhri, 2011). Badiani and Jessoe (2013) show that when irrigation technology arrives, initially farmers increase production by irrigating existing crops, and eventually by planting more profitable highly irrigated crops. Other channels may also be relevant, such as increases in manufacturing productivity and the resulting migration of farm workers to nearby cities and changes in the development of non-farm lands.

Our paper also illuminates a creative policy tool to pursue forest conservation. Preventing deforestation via fines, bans, or designating areas as 'protected' is challenging due to leakage that displaces deforestation towards unregulated areas. Regulators find it difficult to enforce fines or bans, especially in developing countries (Balboni et al., 2021; Burgess et al., 2012, 2019; Harding et al., 2021; Harstad and Mideksa, 2017; Gonzalez Lira and Mobarak, 2019), though protected conservation areas can be effective in some cases (Oldekop et al., 2019; Assunção et al., 2023). Against that backdrop, we find that a seemingly-unrelated policy of extending electrification to rural areas changes farmers' economic incentives in ways that inadvertently protect the forest without any monitoring or policing requirement. The magnitude of the effect on forest cover appears to be as large as the documented effect of the most prominent package of direct conservation policies ever implemented in the Amazon called PPCDAM (Assunção et al., 2015). This suggests that conservation policies should maintain a large focus on

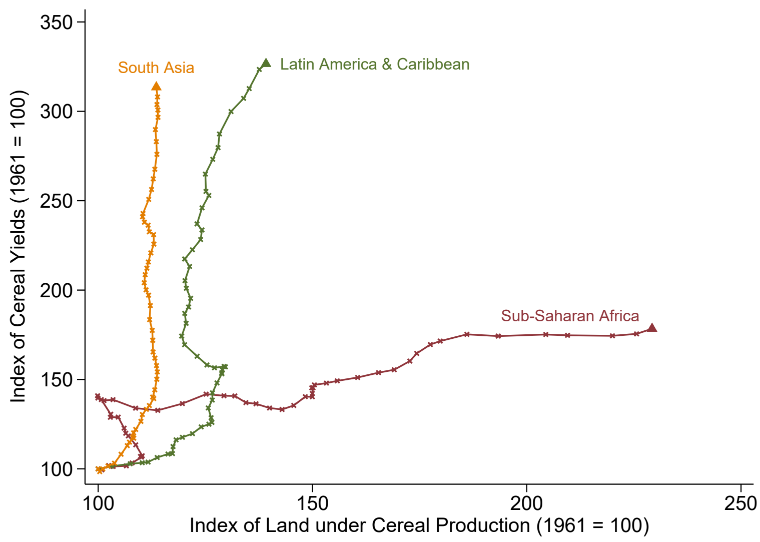

Figure 1: Cereal Yields and Land under Cereal Production: 1961–2014

Notes: In the figure, each data point is one region-year, starting in 1961 (the left-most points) and ending in 2014 (the top-most points). In South Asia, most of the increase in production came from rising yields, with little addition of new agricultural land. In Sub-Saharan Africa, agricultural land increased more than yields. Latin American followed the Sub-Saharan path until the 1980s and then shifted towards the Asian model of agricultural growth. The figure reflects the Green Revolutions in Asia and Latin America and suggests that, at a continental scale, Borlaug's hypothesis holds true. Original data source is the Food and Agriculture Organization (FAO). Data was downloaded from the World Bank's World Development Indicators.

the economic interests of user groups. Policies such as direct payments for ecosystem services (Porras et al., 2012; Jayachandran et al., 2017; Jayachandran, 2013), policies that encourage substitution of land use away from cattle grazing (Araujo et al., 2020), or interventions that improve farm productivity fall in this camp.

We also contribute to the literature on the effects of electrification. Lewis and Severnini (2020), Chakravorty et al. (2016), and Dinkelman (2011) document economic benefits of rural electrification in the U.S., the Philippines and South Africa. Lee et al. (2020c) and Burlig and Preonas (2016) find no significant economic gains from *household* electrification. As Lee et al. (2020b) discuss, the main difference between these studies and our approach is that we study the impact of new electricity infrastructure at an aggregate level, which benefits not only households but also firms, farms, and entire communities. Another important difference is the target population and time frame—while Lee et al. (2020c) focuses on immediate impacts of electrification on poor rural households, we study changes in land use over a 50-year horizon. A third difference is that, as electricity rolled out, ready-to-implement technologies and processes useful for farming that rely on electricity (like irrigation, drying, storage, and processing) were already available (Fluck, 1992). Our setting shares the spirit of that of Usmani and Fetter (2020), who show that rural electrification produces significant economic effects in

India only when preexisting economic opportunities are in place.

2 Context: Agriculture in Brazil

In this section, we provide stylized facts about land use and agriculture during our sample period. These stylized facts motivate the conceptual framework presented in section 4, which guides our empirical analysis.

Land Use Table 1 provides an overview of farmland and land-use within farms during our study period. According to the Census of Agriculture, farms occupied 35 to 44 percent of the country's territory between 1970 and 2006. Within farms, the land is divided into pastures (mostly for cattle), cropland, and native vegetation. These three land-use categories account for over 80 percent of farmland. "Native vegetation" is the Census of Agriculture's land use categorization for vegetation that has not been cultivated. This category refers to "forests" in most of Brazil, but savannas and steppes are native to some areas.[2] The share of farmland dedicated to pastures declined between 1970 and 2006, while the shares of cropland and native vegetation increased. Farmers keep native vegetation on their properties not only to extract forestry income, but also because frictions in labor, credit, or rental markets may make some of the land unprofitable to cultivate.

Table 1: Land Use in Brazil: 1970–2006

Year	$\frac{\text{Farmland}}{\text{Country Area}}$	Shares of Land Use within Farmland			
		Pastures	Cropland	Native Vegetation	Other
	(1)	(2)	(3)	(4)	(5)
1970	0.35	0.52	0.11	0.19	0.18
1985	0.44	0.48	0.14	0.22	0.16
2006	0.39	0.48	0.16	0.29	0.08

Source: Brazilian Census of Agriculture. The table shows country-wide figures published by IBGE between 1970 and 2006. Column 1 shows the total land in farms (*farmland*) divided by the country area. Column 2 shows pastureland divided by farmland. Column 3 shows cropland used for annual crops, including fallowing, divided by farmland. Column 4 shows the area under native vegetation (*matas naturais*) divided by farmland. Column 5 shows the share of land in all other uses, which include perennial crops; planted forests; farmsteads, buildings, livestock facilities, ponds, roads, and wasteland; and other non-cultivated land.

Crop Cultivation and Cattle Grazing Cattle grazing is a relatively land-intensive activity in Brazil, while crop cultivation is more intensive in physical and human capital. To illustrate, table 2 shows measures of capital-to-land and labor-to-land ratios in crop

[2] The presence of trees within farms is a common feature in the tropics. For example, Zomer et al. (2014) finds that 50 percent of agricultural land in Central America had at least 30 percent of tree cover in year 2000.

and livestock farms. For example, in 2006, the value of machinery per hectare in the typical livestock farm was one-fifth of that of a typical crop farm. Furthermore, crop farms had three times as many workers per hectare as livestock farms. These figures reflect the fact that only 4 percent of cattle farms use confinement, whereas over 60 percent of the harvested area of maize is mechanized, as is virtually all of the country's soybean production (IBGE, 2006). In short, cattle grazing requires low levels of capital when compared to crop farming.

Table 2: Labor and Capital intensivity in Crop and Livestock Farms: 1970–2006

Year	Workers Per 1,000 Hectares		Value of Machinery, Equipments, and Vehicles Per Hectare	
	Crop farms	Livestock farms	Crop farms	Livestock farms
	(1)	(2)	(3)	(4)
1970	125	19	59	17
1985	120	31	415	118
2006	92	31	550	114

Source: Brazilian Census of Agriculture. The table shows country-wide figures published by IBGE between 1970 and 2006.

Credit and Labor Market Constraints Like many rural economies in the developing world, farmers in Brazil face labor and capital constraints, and access to credit is far from universal (Conning and Udry, 2007). Between 1970 and 2006, at least 80 percent of Brazilian farmers had no access to external financing, and access to other financial products such as insurance was even less common (IBGE, 2006). Agricultural labor markets in developing economies are also plagued by search and informational frictions (Fink et al., 2020; Jeong, 2020), on top of the notoriously stringent employment regulations in Brazil that make hiring and firing very costly (Ulyssea, 2010).

3 Electricity and Agriculture

Rural Electrification in Brazil The electricity grid expanded massively during the second half of the twentieth century in Brazil, particularly into rural areas. While 75 percent of urban households had access to electricity in 1970, only 4 percent of farms were electrified. By 1995, the proportion of farms with access to the grid increased tenfold. Hydropower, which requires intercepting large amounts of water at high velocity, supported virtually all of this expansion.

Electricity and Crop Production Electricity is valuable for agricultural production, as it enables farmers to use techniques and equipment that would otherwise be more costly or infeasible to implement (Fluck, 1992; Pimentel, 2009). Post-harvest storage and processing of grains requires humidity control and machinery for drying grains, including ventilators, which are cheaper to operate when electricity is available. Lewis and Severnini (2020) analyze data from a rural electrification experiment in the 1920s in the US where farm machines, excluding water pumping devices, consumed 30 percent of the electricity provided to farms, with consumption peaking during the harvest season. Furthermore, electricity provides energy that is necessary to pump and distribute groundwater for irrigation. During the Green Revolution, electricity permitted widespread irrigation and enabled Indian farmers to fully realize the productivity benefits of new seed varieties (Rud, 2012).[3]

Electricity and Livestock Production Livestock production can also benefit from electricity in theory, depending on the specific products the sector focuses on. Mechanized milking, pasteurizing, and cooling of dairy products are energy-intensive processes (Lewis and Severnini, 2020).[4] However, 85 percent of the Brazilian cattle herd is for beef production, and only one-fifth of the milk-producing farmers pasteurize their milk, which is typically done in facilities outside the farm-gate (IBGE, 2006). Given the nature of farming in Brazil, electricity has been far more important for crop production than for cattle grazing. This will drive an important assumption in our model, and we will later verify in the data that electricity was indeed more beneficial for crop production than for cattle.

Indirect effects of Electricity on Agricultural Production Electrification can affect farm production through other less direct channels. For example, electricity can be labor-saving in rural areas (Dinkelman, 2011). Since crop cultivation in Brazil is more labor-intensive than cattle grazing, this channel would presumably also benefit crop production more than livestock production, in line with our argument. Furthermore, large-scale electrification may support the supply of services available to farmers, such as banking, telecommunications, private sector supply chains, and extension services. Rural electrification increases human capital accumulation (Lipscomb et al., 2013), and rural amenities, which can attract more skilled people to rural areas. All of this can improve farm productivity, and again may have differential impacts on crop production and animal husbandry.

[3]While diesel pumps are an alternative, diesel was especially expensive during oil price shocks in the 1970s (World Bank, 1990; Rud, 2012).

[4]Industrial poultry farming is also energy-intensive, and benefits from electrification (Lewis and Severnini, 2020). However, poultry requires very little land and plays a minor role for the land use and deforestation decisions we study in this paper. For this reason, we use the terms "livestock production" and "cattle grazing" interchangeably in this paper.

4 Conceptual Framework

In this section we build a simple, partial-equilibrium theoretical framework inspired by the salient features of farming in Brazil, with the goal of generating predictions on how a positive productivity shock to crop cultivation will affect farming choices and deforestation. To mirror the language in our empirical exercise, we refer to the key productivity parameter in our model as "availability of electricity" and denote it by Ω.

Set up The economy is endowed with total land of $\overline{H}$ which is initially covered by native vegetation. A continuum of individuals reside in this economy, and each decides whether to become a farmer. These individuals differ in their opportunity cost of operating a farm, $\theta \sim \Gamma$, which can be thought of as the wage rate in the non-farm sector. The availability of electricity can affect farmers' outside options, and so we assume that Ω shifts Γ in the sense of first-order stochastic dominance; the direction of this shift depends on how electrification affects returns to farming relative to non-farm activities.[5] The equilibrium profit from farming, Π^*, is common across farmers; the mass of farmers is therefore $\Gamma(\overline{\theta})$, where $\overline{\theta} = \{\theta : \theta \leq \Pi^*\}$.

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, allocating H_c and H_g units of land to crops and pasture, respectively, with land that is not used for either remaining as forestland. The production functions for the two activities are similar, except that crop cultivation requires an additional factor, denoted by K. We think of K as capital, labor, or a combination of both. We make two simplifying assumptions. First, we assume that electrification improves the productivity of crop cultivation, but not cattle grazing. Second, we assume that only crop cultivation requires K. Specifically, 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 $F_H(0) = \infty$.[6]

Land and capital can be bought at prices p and r, respectively. Farmers are credit constrained and need to fund their expenditures on capital such as tractors or irrigation equipment and land from their own resources, M. We normalize the prices of C and G

[5]Formally, an increase in electrification can cause the distribution of opportunity costs to dominate or be dominated by the original distribution. If electrification improves farmers' outside options more (less) than it improves the returns to farming, then $\Gamma(\theta; \widehat{\Omega}) \leq (\geq) \Gamma(\theta; \widetilde{\Omega})$, for all $\widehat{\Omega} > \widetilde{\Omega}$.

[6]The model's substantive implications require only that crop cultivation is more capital-intensive, and benefits relatively more from electrification, and both these assumptions are supported by the data.

to 1.[7] Thus, the farmer's problem can be written as:

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

$$\text{subject to} \quad rK + p(H_c + H_g) \leq M, \tag{1b}$$

$$H_c + H_g \leq H. \tag{1c}$$

Predictions Since the profit function is linear in K and $F_H(0) = \infty$, the resource constraint (1b) always binds. This is merely a modeling device, and what is essential in this model is that farmers are constrained in their ability to hire K. Land will therefore not be the limiting factor, and the land constraint (1c) will not bind. The solution to this problem yields optimal land use and production choices with the following properties:

$$\frac{\partial K^*}{\partial \Omega} > 0 \tag{2a}$$

$$\frac{\partial H_c^*}{\partial \Omega} > 0 \tag{2b}$$

$$\frac{\partial H_g^*}{\partial \Omega} < 0 \tag{2c}$$

$$\frac{\partial (H_c^* + H_g^*)}{\partial \Omega} < 0 \tag{2d}$$

The intuition behind equations (2a)–(2d) is straightforward. Since capital and land allocated to crops become more productive with electrification, K and H_c move in the same direction as Ω, as shown in equations (2a) and (2b). However, since the credit constraint binds, the farmer can only increase land allocated to crop cultivation and hire more K if she decreases land allocated to cattle grazing (equation 2c). The total demand for agricultural land within the farm, $H_c^* + H_g^*$, decreases in response to increases in electrification (equation 2d): as farmers substitute cropland for pastureland, they also spend more on K and hence must give up more of H_g than they can increase H_c.[8]

The net effect of electrification on deforestation depends 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, H_v, as the difference between the economy's total land endowment and the farmer's total land demand

[7] To the extent that commodity prices are exogenous to local conditions, this normalization is innocuous. In any case, making prices endogenous to the (local) productivity shock would not add predictive power to this framework.

[8] In reality, the price of cropland is typically higher than the price of pastureland, which would make the negative effect on H even stronger. That would strengthen our prediction, but our model ignores that margin. In the same vein, we assume that electrification does not increase input prices. If electrification increases relative price p/r, farmers would adjust by spending more in K and less in $H_c + H_g$.

for agricultural purposes:

$$H_v = \overline{H} - \int_{-\infty}^{\overline{\theta}} (H_c^* + H_g^*) d\Gamma(\theta) \tag{3}$$

The derivative of the total area of native vegetation with respect to electrification has two components:

$$\frac{dH_v}{d\Omega} = \underbrace{-\frac{d(H_c^* + H_g^*)}{d\Omega}\Gamma(\overline{\theta})}_{>0} - \underbrace{(H_c^* + H_g^*)\Gamma(\overline{\theta})\frac{d\overline{\theta}}{d\Omega}}_{\lesseqgtr 0} \tag{4}$$

The first term relates to the intensive-margin adjustment, through which electrification reduces the demand for agricultural land within farms. The second term is the extensive-margin effect: electrification changes $\overline{\theta}$—the threshold in the distribution of farming opportunity costs below which individuals decide to farm. Whether this threshold increases or decreases with electrification depends on the changes in the returns to farming relative to returns in the non-farm sector.

The net effect on the forest is therefore theoretically ambiguous. We will examine each of the two (intensive and extensive margin) effects separately in the data, and then combine them quantitatively to infer the total effect of a productivity shock on deforestation.

To summarize, this model makes a few assumptions about the agricultural production function that we can examine in the data, and yields a few further testable predictions. First, we make the testable assumption that electrification increases crop cultivation productivity more than cattle grazing productivity. Second, we assume that farmers face constraints in factor markets. Beyond the support for this assumption that we have already provided in section 3, we will formally test its implications using variation in rainfall and credit access. Third, the model predicts that electrification should lead to greater investments in capital, specifically in capital that is useful for crop farming. Fourth, the model predicts that positive productive shocks induce farmers to shift land use from cattle grazing to crop cultivation. Finally, the model highlights that electrification has intensive- and extensive margin effects on the demand for agricultural land. On the intensive margin, it reduces demand for agricultural land through reductions in land demand for cattle grazing. Increases in land demand for crop cultivation are not enough to offset the reduction in land demand for cattle grazing. The effects on the extensive margin are ambiguous. Demand for farmland may or may not increase depending on the relative magnitude of farms' profits vis-à-vis farmers' outside options.

5 Data Sources and Variable Definitions

We build a decennial panel dataset of Brazilian counties from 1960 to 2000 that combines data from three main sources. First, we use historical measures of electricity infrastructure and the evolution of the grid by decade compiled by Lipscomb et al. (2013), who also provide the basis of our instrument for the growth of electrification in Brazil (further described in section 6). Second, we use five waves of the Brazilian Censuses of Agriculture to track farmland expansion, land use within farms, agricultural output, and inputs. Finally, we use data derived from satellite imagery, which has the advantage of measuring land use for the entire county, although it cannot distinguish between land use within and outside of farms.

The number of counties (or *municípios*) in Brazil increased from 2,766 to 5,564 during our study period due to redistricting. The concordances provided by Reis et al. (2011) and Ehrl (2017) enables us to construct a balanced panel of minimum comparable areas that remain constant over this period. After merging data from all sources, our panel contains 2,172 "standardized counties". Table 3 presents summary statistics for the data we use. We now describe these data.

Electricity Infrastructure Lipscomb et al. (2013) compile digital maps of the electricity grid in Brazil for the period 1960-2000 based on archival research that identifies the locations and construction dates of hydro-power plants, transmission lines, and substations.[9] The authors obtained reports, inventories, and maps from major regional electricity companies in Brazil operating during that period, and consolidated that information into the status of the electricity grid in each decade. Following Lipscomb et al. (2013), we divide Brazil into 32,578 evenly spaced grid points; in counties with a power plant or substation, we assume that all grid points within a 50-kilometer radius of the county's centroid have access to electricity.[10] Our county-level electricity access variable is defined as the county's proportion of grid points that are electrified.

Agricultural Data The Brazilian Census of Agriculture is a comprehensive source of data on the universe of farms in the country.[11] We use county-level aggregations from

[9]Transmission lines transport electricity from power plants to substations in regions where the electricity will be used. Substations reduce voltage, making the electricity suitable for the distribution lines that ultimately supply electricity to firms, farms, and households.

[10]We proceed in this way because the authors compiled historical data on generation plants and transmission stations, and it was not feasible to collect direct data on the many distribution companies operating during the period. Distribution networks stretch one-hundred kilometers on average, which is why we choose the 50-km radius.

[11]The Census of Agriculture surveys all rural establishments in the country. The definition of a rural establishment 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 plot size, tenure, or market participation.

the following five waves of the Census of Agriculture: 1970, 1975, 1985, 1996, and 2006. These waves match agricultural outcomes to the electricity infrastructure data with a five-year lag. We use the census data to investigate the effects of electrification on three sets of outcomes: agricultural output, land use, and capital use.

When looking at agricultural output, we measure crop yields by dividing total grain production by harvested area of all major grains: maize, cotton, soybeans, beans, rice, and wheat. For cattle grazing, we use the stocking ratio, i.e., heads of cattle divided by hectares of pastureland. We measure changes in the composition of farm production using the cattle share of farm production value, which captures the relative variation in physical output and prices.

To examine the effects on land use, we start by using the proportion of county area in farms (*farmland*). We then examine the within-farmland shares of three land-use categories—*cropland*, *pastures*, and *native vegetation*[12]. *Cropland* is the harvested area of all major grains: maize, cotton, soybeans, beans, rice, and wheat, which on average corresponds to 87 percent of the harvested area of all annual crops during our sample period.

Finally, to test our model's prediction that electrification increases investment in crop-related capital, we use the per-hectare number of tractors, planting and harvesting machines, plows, and grain storage facilities as measures of farm capital stock. We also use the proportion of farms that use irrigation.

Satellite Data We complement our analysis using satellite-based classification of land use. Specifically, we use data from collection 3 of the Brazilian Annual Land Use and Land Cover Mapping Project (MapBiomas), which provides land-use classifications based on annual composites of LANDSAT images from 1985 to 2017. MapBiomas classifies each 30m x 30m pixel into 20 land use categories using algorithms specific to each of the six Brazilian biomes.[13] We downloaded per-county aggregations with total area in each land use category for 1985, 1995, and 2005, to match the years in the Census of Agriculture data. We then aggregate the 20 land-use categories in the satellite data into four classes that match well with the agricultural census data: natural vegetation, pastures, cropland, and others. Finally, we match these data to the electricity infrastructure data with a 25-year lag, on average. The main advantage of satellite measurements is that they contain information on land use in the whole county, not only within farms.

[12] The remaining farm area includes orchards, annual crops other than grains, planted forests, buildings and facilities, water bodies and non-arable land. The area in planted forests is small, and bundling it with native vegetation makes no quantitative difference in our results.

[13] To our knowledge, MapBiomas is the only satellite data product with land use classification for the entire Brazilian territory and this time period. The data and classification methods are publicly available at `mapbiomas.org/en`.

Table 3: Sample Descriptive Statistics

	Mean	Std. Dev.	Min	Max
Electricity				
Electricity Infrastructure	0.75	0.40	0.00	1.00
Modeled electricity instrument	0.74	0.43	0.00	1.00
Percent of Farms Electrified	0.33	0.35	0.00	2.05
Land Use				
Farmland/County Area	0.71	0.27	0.00	6.26
Pastures/Farmland	0.47	0.24	0.00	0.95
Cropland/Farmland[a]	0.11	0.15	0.00	3.14
Native Vegetation/Farmland	0.18	0.15	0.00	0.99
Capital usage				
Percentage of Farms with Irrigation	6.16	10.94	0.00	93.28
Grain Storage Facilities[a]	27.34	85.90	0.00	3,263.17
Planting and Harvesting Machines[a]	10.11	23.69	0.00	669.86
Plows[a]	20.46	38.35	0.00	587.97
Tractors[a]	22.94	38.50	0.00	868.05
Agricultural output				
Grain Yields (log)[a]	0.07	0.79	-2.77	3.01
Heads of Cattle Per Hectare	1.07	0.76	0.12	13.24
Share of Cattle Production Value	0.29	0.22	0.00	0.99
Land Use - Satellite Data				
Pastures/County Area	0.48	0.28	0.00	1.00
Cropland/County Area	0.09	0.17	0.00	0.91
Native Vegetation/County Area	0.36	0.27	0.00	1.00
Other				
Bank Branches[a]	0.64	3.10	0.00	112.1
Number of Counties	2,172			
Number of observations	10,860			

Notes: The data is a decennial panel of standardized counties from 1960 to 2000. [a] Per 10,000 hectares of county area. See Appendix B for more detailed descriptions of variable definitions and sources.

6 Empirical Strategy

We estimate the effect of electrification on land-use decisions using a decennial panel of counties. Assuming a linear and additive structure, we are interested in producing an unbiased estimate of θ in an equation of the following form:

$$Y_{c,t+1} = \alpha_c + \gamma_t + \rho E_{ct} + v_{ct}, \quad (5)$$

where $Y_{c,t+1}$ is the outcome of interest (e.g., land use) in county c at time $t+1$, α_c is a county fixed effect, γ_t is a time fixed effect, and E_{ct} is our measure of electricity infrastructure—the proportion of grid points in county c that are electrified at time t.

Estimating (5) by OLS is likely to produce biased estimates of ρ because the placement of electricity infrastructure is not random. For example, infrastructure may expand to areas where the demand and economic returns are high, or may be allocated by the government to under-served areas. Since these factors are unobserved to us, they will be in v_{ct}, and may affect the demand for agricultural land. The correlation between E_{ct} and v_{ct} could create a bias in the OLS estimator in either direction.

The Instrument To overcome these identification concerns, we use an instrumental variable (IV) approach. We instrument E_{ct} with the output of an electricity infrastructure expansion model that is based only on engineering cost considerations. Our approach, originally developed in Lipscomb et al. (2013), takes advantage of two facts. First, hydropower accounts for the majority of electricity generation in Brazil. The cost to build a hydropower plant is largely determined by geography. Hydropower plants require a steep river gradient and a large amount of water flow to create pressure from the water descending through the turbines. Although creating these conditions artificially is possible through large investments, naturally steep areas with high water flow have a clear cost-advantage. Second, the expansion of the electrical grid in Brazil was led by the federal government. We are thus able to identify the budget for electricity generation in each decade that determined the expansion. We use this information on the country-wide scale of expansion to introduce panel variation in the instrument, similar to the approach in Duflo and Pande (2007). Our basic strategy is to forecast how the grid would have evolved decade-by-decade, if it expanded from low to high-cost locations.

To construct our instrument—predicted electricity availability—we mimic the (counterfactual) decision-making of a planner who focuses only on engineering costs (and ignores demand) to determine the expansion of the electricity network. Using a spatial grid of 32,578 evenly spaced points, we build a model that predicts electricity availability at each grid point-decade, and we then aggregate it to the county-decade level.

The model works in three steps.

The first step provides a ranking of the cross-section of grid points based on their *suitability* for a hydropower plant. To do that, we create a hydropower suitability index for every grid point based on topographic factors (water flow accumulation and river gradient) and the presence of dense forests, which are also known to considerably increase construction costs. The suitability index is the predicted probability of the point receiving a hydropower plant from a probit regression in which the dependent variable indicates whether the location has a dam built as of the year 2000 (the end of our sample period). Appendix table B.1 shows the result of this probit regression.

The second step determines which point-decades are allocated a hydropower plant by the hypothetical planner seeking to minimize cost. For example, we know that the first decade's budget allowed for the construction of 53 new plants. Therefore, the 53 lowest-cost locations are allocated generation plants in the first decade in the instrument we construct. The second decade's budget allowed for the construction of 36 new plants, which means that the next 36 lowest-cost locations would receive a generation plant, and so on for every subsequent decade. In the third and final step, we use simulated grid annealing to optimize the lowest cost placement of two transmission substations per generation plant and assume that all grid points within 50 kilometers of a predicted plant or substations are electrified imposing the constraint that areas already electrified don't need additional infrastructure. The simulation leaves some variation in the predicted grid in each decade, so we deviate slightly from Lipscomb et al. (2013) in that we average the predicted network across 500 runs of the Matlab code in order to improve the signal to noise ratio. The steps for estimation of the instrument are explained in greater detail in Lipscomb et al. (2013).

Estimation Our instrument is the fraction of grid points in each county-decade predicted to be electrified by this model; it is designed to capture how the grid *would have evolved* over decades had investment decisions been made solely on the basis of geography-driven construction cost considerations. Denoting the instrument by Z_{ct}, we estimate the following system of equations by two-stage least squares:

$$E_{ct} = \alpha_c^1 + \gamma_t^1 + \delta Z_{c,t} + \theta^1 X_{c,t} + \eta_{c,t} \tag{6}$$

$$Y_{c,t+1} = \alpha_c^2 + \gamma_t^2 + \beta \widehat{E}_{c,t} + \theta^2 X_{c,t} + \varepsilon_{c,t} \quad , \tag{7}$$

where $\widehat{E}_{c,t}$ are the fitted values from the first stage regression (6). Note that both $Z_{c,t}$ and $E_{c,t}$ are county-level averages of grid point values. We therefore weight regressions by the county's number of grid points. We cluster standard errors by county. In our preferred specification, we include interactions between the hydropower suitability index and time fixed effects in both estimating equations, which flexibly control for geography-specific trends. As we discuss below, these controls help to reinforce the credibility of our exclusion restriction.[14]

Identification Our identification is based on the hypothetical grid's expansion to less attractive, higher cost locations for hydropower over time. Because we have time variation in (actual and modelled) electrification within counties, we can include county fixed-effects in our IV estimation. This means that time-invariant geographic characteristics, such as elevation, slope, and water availability, do not contribute to identifying the (local average) treatment effects that we estimate. This is important because a key threat to identification would otherwise be that geographic factors like elevation and water flow directly affect agriculture.

In our fixed-effects framework, the exclusion restriction concerns the decade-by-decade process of *evolution* of our hypothetical electricity grid, which produces discontinuous jumps in the probability of early-versus-late electrification in certain locations. The 53 lowest-cost locations are electrified in the first decade, but locations with topographic factors that barely miss the cut (that rank 54, 55, 56, etc, in the first step of instrument construction) have to wait for the next decade to receive power. Location fixed-effects and trends specific to the suitability ranking[15] absorb much of the geographic variation in the instrument, thereby isolating these discontinuities in our cost-based forecast for a county to be electrified as the key source of identification. Our argument is that these discontinuous jumps in receiving power early-versus-late are unlikely to be directly related to the continuous evolution of farmland expansion, pasture-to-cropland conversion, and deforestation, and should not violate the exclusion restriction in our IV strategy.

Moreover, our model of grid expansion predicts that electrification evolves from areas with high slopes and water volume to areas with low slopes and no water. Historically, the expansion of agricultural land and population tracks water availability. However, agricultural land and population also evolve from areas with low slopes to areas with higher slopes. The interplay of these two geographic factors in opposite directions implies that the dynamics of our instrument should not be collinear with the dynamics of agriculture and population movements, which strengthens the confidence in our exclu-

[14] Our panel data contains five distinct time periods. We use the subscripts $Y_{c,t+1}$ and E_{ct} in equation (7) to denote the fact that outcome variables Y collected from various waves of the agricultural census are matched to the electricity infrastructure data that is typically lagged about 5 years. For example, in the first period defined in our dataset, the 1970 agricultural census is matched to the electricity grid for the 1960s reported in Lipscomb et al. (2013), while the last period is the 2006 census matched to the electricity grid in the early 2000s.

[15] We flexibly control for a polynomial in the suitability index interacted with time-fixed effects in our preferred specification, to alleviate concerns about spurious common trends between our predicted grid expansion instrument and spatial patterns of agricultural expansion and deforestation.

sion restriction. The fact that our instrument is not collinear with population dynamics also alleviates concerns that it predicts the expansion of other types of infrastructure such as sanitation or roads. That said, road-building and electric-grid development often go hand-in-hand. However, even if our instrument predicts that broader set of infrastructure, we are only interested in using that as an agricultural productivity shock to explore subsequent changes in land use. Whether the productivity shock comes from rural electrification alone, or a combination of electricity and roads, is not a first-order concern for this paper. It could change the precise policy implication for governments interested in conserving forests ("improve agricultural productivity" as opposed to "invest in electrification"); these are important subtleties in interpretation.

A weakness of our IV approach, as we detail in appendix D.2, is that our instrument becomes weak (F-Stat of 7.3-8.1 depending on the specification) in the first stage if we directly control for trends specific to each component of the geographic factors used in the instrument: water flow, river gradient and the Amazon.[16] To avoid the weak instrument problem, we instead control for trends specific to the hydropower suitability index, which combines those three inputs into a single index. Appendix D.2 shows that inclusion of the Amazon-specific time trends is the main culprit that makes our first-stage weak, so that appendix also explores the sensitivity to removing the Amazon region entirely from the analysis. The Amazon region accounts for 59% of the land area of Brazil, and therefore 59% of the grid points used for instrument construction fall within the Amazon. Our IV approach breaks down when excluding these 59% of the effective (weighted) sample because the instrument becomes too weak. However, we retain adequate first-stage power at conventional levels (F-stat of 16.9-26.8) when excluding up to two-thirds of the Amazon territory. Specifically, we divide Amazonian counties into three groups with approximately equal area – counties either belong to the state of Pará, Amazonas state, or other states. Excluding any two of these groups at a time keeps the second-stage results largely unchanged, and the main conclusions of our paper are retained.

Instrument Validity Tests We perform two additional exercises to explore the validity of our empirical strategy. First, we check whether the spatial dynamics of farm production follows the same pattern as our forecast of electricity grid expansion. To do that, we compute correlations between the rank order of a county's hydropower suitability and the rank order for two indicators—farm production value per hectare and the county's share of farmland—within regions of Brazil and within years.[17] This is a stringent test of our identification assumptions because in our analysis we use

[16] This issue is further detailed in an online-only note (Lipscomb et al., 2021).

[17] Lipscomb et al. (2013) perform a similar exercise and find low correlations between the suitability index and population density and GDP.

Table 4: Spearman Correlations: Hydropower Suitability and Agriculture

	Year				
Region	1960s	1970s	1980s	1990s	2000s
Panel A: Hydropower Suitability and Farm Production Value Per Hectare					
North	−0.185	−0.176	−0.171	−0.209	−0.134
North East	−0.129	−0.106	−0.028	−0.059	−0.011
South East	−0.026	−0.021	−0.040	−0.057	−0.067
South	0.071	0.041	0.055	0.058	0.056
Center West	0.384	0.543	0.536	0.445	0.106
Panel B: Hydropower Suitability and Share of Farmland					
North	0.177	0.185	0.184	0.197	0.191
North East	0.119	0.094	0.070	0.116	0.081
South East	0.013	0.015	0.001	−0.015	−0.022
South	0.061	0.066	0.056	0.065	0.053
Center West	0.479	0.529	0.539	0.455	0.439

Notes: Each cell presents a Spearman rank order correlation, by region and decade. In Panel A, the correlation is computed between the suitability rank for hydropower generation and the rank for the farm production value per hectare. In Panel B, the correlation is computed between the suitability rank for hydropower generation and farmland share of the county area.

county fixed-effects, and not region fixed-effects. Table 4 shows the results. For most regions and decades, these correlations range from 0.01 to 0.21 in absolute value. The exceptions are in the Center-West region, where the magnitudes hover between 0.10 and 0.54. This can be a source of concern, and in appendix D.3, we show that the main results of our paper do not change if we exclude the center-west region from the analysis.

Second, we check whether placement of power plants and transmission lines simulated by the forecasting model can be predicted by farm production indicators in earlier years by estimating equations of the form $instrument_{it} = \beta outcome_{i,t-1} + \lambda_t + \theta_i + \xi_{it}.$. Under the null that our instrument is as good as random the estimates of β should be statistically equal to zero. Results are shown in Appendix Table C.1. Almost all coefficients are small in magnitude, suggesting that pre-existing production factors are not a major factor in determining electricity placement.

7 Empirical Results

7.1 First-stage results and Electricity Adoption by Farms

Table 5 shows whether our instrument — the hypothetical grid expansion based on cost considerations — predicts the actual evolution of Brazil's electricity grid. The specification in column 1, which uses county as well as time fixed effects, confirms a strong and significant correlation between modeled electrification and actual electricity infrastructure. If our model predicts that a county should get electrified in a certain decade, the likelihood of that county actually receiving electrification increases by 46 percentage points. Column 2 adds flexible controls for geographic-specific trends by including a quartic polynomial of the hydroelectric suitability index interacted with time fixed-effects. The point estimate decreases from 0.46 to 0.39, and the partial F-statistic decreases from 67.0 to 36.3. We use this more conservative specification from column 2 in the remaining results reported in this paper.[18]

While these F-statistics are above the thresholds traditionally used as a rule of thumb as suggested by Staiger and Stock (1997), recent work by Lee et al. (2020a) shows that conventional clustered standard errors may be biased downward when the F-statistic is below 104.7. With an F-statistic of 36.3, the tF critical value is 2.247 (Lee et al., 2020a, Table 3), which implies that the 95% tF-corrected confidence intervals are 15 percent wider than conventional confidence intervals. In appendix table D.1, we reproduce all second-stage results along with the tF-corrected 95% CIs suggested in Lee et al. (2020a). Note that Lee et al. (2023) points out that these intervals are in most cases overly conservative. Only one coefficient (investment in planting and harvesting machines) has tf-corrected confidence intervals that include zero when conventional confidence intervals do not.

Column 3 explores whether the instrument predicts electrification of farms, based on questions directly asked to farm operators in the agricultural census. Our instrument not only predicts the arrival of grid infrastructure, but also adoption of electricity by farms (column 3) controlling for county, year, and the quartic suitability rank interacted with decade fixed effects. Column 4 shows that the arrival of the infrastructure grid increases the proportion of farms electrified by 13.7 percentage points (pp). We use this to calibrate the effects throughout the paper–we highlight changes in productivity and deforestation as a result of a 10 pp increase in electrification, to keep the arithmetic simple and benchmark the magnitudes to something like an intent-to-treat effect of electrification policy.

7.2 Does electrification affect agricultural productivity?

We now test the central assumption of our theoretical framework, namely that electricity improves agricultural productivity, and that the productivity shock is biased

[18] Note that the instrument is an estimated prediction based on simulated grid annealling. As a result, there are small fluctuations in the instrument based on the seed. We simulate the grid 500 times and average the grid point predictions before aggregating to the county (amc60) level. This improves on the strategy used in (Lipscomb et al., 2013) by providing more stable estimates and by improving the signal to noise ratio.

Table 5: First-Stage Results

Dependent Variable	Electricity Infrastructure		Proportion of Farms with Electricity	
	(1)	(2)	(3)	(4)
Instrument	0.456***	0.385***	0.278***	
	[0.056]	[0.064]	[0.049]	
Electricity Infrastructure				0.137***
				[0.022]
Year dummies	Yes	Yes	Yes	Yes
Quartic suitability rank × year dummies	No	Yes	No	Yes
Observations	10,860	10,860	10,860	10,860
Number of Counties	2,172	2,172	2,172	2,172
Mean of Outcome	0.750		0.334	0.334
Partial F-stat	67.0	36.3		

Notes: In columns 1 and 2 the dependent variable is the prevalence of electricity infrastructure in the county measured from infrastructure inventories. Adding a quartic polynomial of the suitability index interacted with year dummies does not change the coefficient substantially. We keep the specification in column 2 as our preferred specification throughout the paper. In columns 3 and 4, the dependent variable is the fraction of farms with electricity in the county, measured from the Censuses of Agriculture. These columns show that both the instrument and our measure of electricity infrastructure correlate with a measure of rural electrification contained in our main dataset. All specifications include county fixed effects and are weighted by the number of grid points contained in the county. Standard errors clustered at the county level in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: The Effects of Electricity on Agricultural Production

Dependent Variable	(1) Grain Yields (log)	(2) Heads of Cattle Per Hectare of Pastureland	(3) Cattle Production Value/Total Production Value
IV	0.948***	−0.399	−0.242***
	[0.231]	[0.261]	[0.085]
OLS	0.109***	−0.011	0.027*
	[0.040]	[0.067]	[0.015]
Observations	10,700	10,648	10,860
Mean dep. var.	0.07	1.07	0.29

Notes: This table shows that electricity infrastructure is a positive productivity shock to crop cultivation (column 1) but not to cattle grazing (column 2). A corollary is that farm production shifts away from cattle grazing hough not statistically significantly so (column 3). The dependent variable in column 1 is the log of grain yields (production, in thousands of tons, divided by grain harvested area in the county). Grains include maize, soybeans, cotton, wheat, beans, and rice. The dependent variable in column 2 is the number of cattle heads per hectare of pastureland in the county. The dependent variable in column 3 is the ratio of cattle production value to total farm production value. All regressions are weighted by the number of grid points in the county, and include county fixed effects and interactions between time fixed effects and a quartic polynomial on the suitability index. Standard errors clustered at the county level in brackets.

towards crop farming rather than livestock-rearing. We use our IV strategy to estimate the effects of electricity on measures of crop and cattle production. Table 6 shows the results. Column 1 shows that electricity increases crop yields: the IV estimate implies that a 10 percentage point increase in electricity infrastructure increases yields by 9.5 percent (p-value < 0.01). In contrast, column 2 shows that the cattle stock density does not increase with electricity—both the OLS and IV estimates are negative, small, and not statistically significant. These results support the interpretation that electrification is a productivity shock that benefits crop cultivation more than cattle grazing.

Electricity could impact farm production in ways that do not necessarily improve yields. For instance, electrification could raise incomes and increase the local demand for beef. Electrification could thus be a demand shock which benefits cattle grazing without necessarily affecting the density of the cattle herd. Under those conditions, farmers would find it profitable to expand cattle production instead of crop cultivation, which is the opposite of what our model predicts. Alternatively, the price of grains may decrease in response to higher yields, maintaining cattle as a larger component of the revenue of the farm relative to grains. We therefore test crop versus cattle production's contributions to farm revenue in column 3, which uses the cattle share of the total farm production value as the dependent variable. The IV estimate implies that a 10 pp increase in electricity infrastructure leads to a 2.4 percentage point reduction in cattle production as a proportion of the total value of production on the farm. Farmers increase crop production substantially more than cattle production in response to an increase in electrification.

7.3 Changes in Land Use Decisions

Our model predicts that farmers will increase the share of farmland allocated to crop cultivation and decrease land allocated to cattle pasture after receiving the positive shock to crop cultivation productivity. Further, the increase in cropland will not fully offset the decrease in pastureland because of differences in land-intensity and factor market constraints. Within farms, the demand for total land across all agricultural uses will fall as a result. Electrification and increased crop productivity will produce land-sparing effects.

In contrast to these clear predictions on land use *within* farms, the model delivers ambiguous predictions regarding the effects of electrification on the expansion of farming as a whole. Farmland may expand or contract depending on whether electricity benefits farm or non-farm sectors more.

Table 7 shows the effects of electrification on land use. Columns 1–3 show changes in the share of land *within* farms allocated to pastures, cropland, and native vegetation, respectively. The coefficients' signs are in line with our model's predictions. When a county gets electrified, there is a clear shift away from pastures in favor of cropland. A 10 percentage point increase in electrification lowers the share of farmland in pastures by 3.7 percentage points, and increases the share allocated to grains by 1.5 percentage points. The cropland expansion is smaller than the reduction in pastures, corroborating the land-sparing prediction of our model. This leads to a net positive effect on native vegetation increasing its land-share within farms by 3.9 percentage points (column 3). In sum, a positive productivity shock to crop cultivation has a land-sparing effect within farms, which protects native vegetation.

Next, column 4 shows the effects of electrification on farmland expansion. The negative coefficient in column 4 implies that farmland expands less rapidly once counties get electrified. A 10 percentage point increase in electricity infrastructure lowers the county's share of farmland by 4.7 percentage points. Our model would explain this negative coefficient as evidence that access to electricity improves farmers' outside options more than the returns to agriculture, on average.[19].

7.3.1 Net Effect on Total Native Vegetation Inside and Outside Farms

The Census of Agriculture, by definition, does not report land-use outside farms, which prevents us from directly estimating the effect of improving crop productivity on total (county-wide) native vegetation. However, we can combine the estimates in columns

[19] One might also be interested in the direct effect of the change in agricultural productivity following increased access to electricity. We also provide IV results with electricity as an IV for endogenous productivity–log grain production per hectare–in appendix E.

Table 7: The Effects of Electricity on Land Use

Dependent Variable	(1) $\frac{\text{Pastures}}{\text{Farmland}}$	(2) $\frac{\text{Cropland}}{\text{Farmland}}$	(3) $\frac{\text{Native Vegetation}}{\text{Farmland}}$	(4) $\frac{\text{Farmland}}{\text{County Area}}$
IV	-0.371^{***}	0.146^{***}	0.385^{***}	-0.472^{***}
	[0.125]	[0.052]	[0.128]	[0.134]
OLS	0.028	0.000	0.033	-0.063^{**}
	[0.020]	[0.007]	[0.028]	[0.027]
Observations	10,860	10,860	10,860	10,860
Mean dep. var.	0.47	0.11	0.18	0.71

All regressions are weighted by the number of grid points in the county and include county fixed effects and a quartic polynomial on the suitability index interacted with time fixed-effects. Standard errors clustered at the county level in brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3 and 4 of Table 7 with other data moments to estimate total effects on county-wide native vegetation. Let V_T denote total native vegetation, V_I native vegetation inside farms, and $k(C-F)$ native vegetation outside of farms, expressed as a fraction $k \in [0,1]$ of the difference between county area C and farmland F. This notation gives us an accounting identity for native vegetation: $V_T = V_I + k(C-F)$, which we can use to derive an expression for the effect of electrification on total native vegetation[20]:

$$\frac{\partial(V_T/C)}{\partial\Omega} = \underbrace{\frac{\partial(F/C)}{\partial\Omega} \cdot \left(\frac{V_I}{F} - k\right)}_{\text{extensive-margin effect}} + \underbrace{\frac{\partial(V_I/F)}{\partial\Omega} \cdot \frac{F}{C}}_{\text{intensive- margin effect}} , \tag{8}$$

All terms on the right-hand side of (8) are reported in Table 7 other than the share of native vegetation outside of farms, k. Column 4 provides estimates of $\frac{\partial(F/C)}{\partial\Omega}$ and the mean $\frac{F}{C}$, whereas column 3 provides estimates of $\frac{\partial(V_I/F)}{\partial\Omega}$ and the mean $\frac{V_I}{F}$. We use these to calculate the total effect on native vegetation in table 8 under different assumptions about the value of k.

Column 2 presents our preferred estimate, in which we infer the value of k by combining data from the census of agriculture with satellite data.[21] Under this assumption of $k = 0.86$, our regression estimates imply a 10 percentage point increase in electrification increases the county's share of native vegetation by 5.95 percentage points.[22]

[20]See appendix A.2 for the derivation.

[21]The census data imply that native vegetation inside farms accounts for 0.71 × 0.18 = 13% of the typical county's area. In the satellite data, native vegetation accounts for 38% of the typical county. Therefore, the share of native vegetation outside of farms should be roughly (0.38 - 0.13)/(1 - 0.71) = 86%.

[22]To calculate the standard error of this linear combination of regression coefficients, we treat the sample averages $\frac{F}{C}$ and $\frac{V_I}{F}$ as constants. To obtain the covariance between the estimates of $\frac{\partial(F/C)}{\partial\Omega}$ and $\frac{F}{C}$, we stack observation on outcomes and fully interact equations 6 and 7 with an indicator for the outcome. Alternatively, we can calculate the standard errors by bootstrapping, which yields very similar results to the ones we report.

Table 8: Effect of Electrification on Total Native Vegetation

	Assumed native vegetation cover outside of farms			
	(1) Full Cover (k = 1)	(2) Implied by data (k = 0.86)	(3) Equal Cover (k = 0.18)	(4) No cover (k = 0)
IV	0.661***	0.595***	0.274***	0.189*
	[0.108]	[0.097]	[0.091]	[0.104]
OLS	0.075***	0.066***	0.023	0.012
	[0.025]	[0.023]	[0.020]	[0.021]

This table presents the effects of electrification on total native vegetation for different assumptions regarding the state of native vegetation outside of farms, which we do not observe with the Census of Agriculture data. The effects are weighted averages of the marginal effects in farmland and native vegetation inside farms, shown in Table 7. The weights are, respectively, the share of farmland (sample mean: 0.71) and the difference between the shares of native vegetation inside (sample mean: 0.18) and outside of farms (denoted by k in equation (8)). We have no data on this last share, so we make four assumptions: in column 1, we assume that non-farmland is fully covered by native vegetation (k = 1); in column 2, we assume that k = 0.86, which is what satellite and census of agriculture data jointly imply; Column 3 assumes that non-farmland and farmland have the same share of native vegetation (k = 0.18); and Column 4 assumes that non-farmland has no native vegetation (k = 0). To obtain the covariances needed for computing the standard errors, we re-estimate the regressions in Table 7 stacking the outcomes and estimating the both marginal effects jointly in one regression.

Since k is unknown, columns 1 and 4 provide upper and lower bounds on this estimate. Since the estimate of $\frac{\partial(F/C)}{\partial\Omega}$ is negative, assuming that all land outside of farms is covered by native vegetation, or $k = 1$, gives an upper limit to the effect on total native vegetation. Assuming $k = 0$ yields a lower limit. The upper limit displayed in column 1 implies that a 10 percentage point increase in electrification increases the share of native vegetation by 6.6 percentage points. If land outside farms is entirely deforested ($k = 0$), the effect decreases to 1.9 percentage points, and is statistically significant only at the 10% level (column 4). Column 3 assumes that the share of native vegetation is equal within and outside farms, and this yields an effect of 2.7 percentage points (p<0.01).

To assess the magnitude of a 5.95 percentage point increase in native vegetation, it is helpful to consider to overall state of native vegetation in Brazil. In 1985, the midpoint of our study period, 76% of Brazil's area was covered by native vegetation according to the satellite-based land use classification that we use (see section 5). By 2006, that figure had dropped to 69%; a loss of 7 percentage points. This implies that without the increase in agriculture productivity brought about by the expansion of rural electrification in Brazil between 1970 and 2000, the rate of deforestation (or loss in natural vegetation) would have been almost twice as large. Our estimates suggest that instead of a 7 percentage point decrease in native vegetation between 1985 and 2005, we might have seen a (5.9+7) = 12.9 percentage point loss, if farms had not received that access to electricity which permitted them to move away from cattle grazing.

This effect size is comparable to estimates of the effects the most prominent package of conservation policies implemented in the mid-2000s to curb deforestation in the Amazon that is known by the Portuguese acronym PPCDAM. This policy package included heavier penalties for illegal deforestation and tighter enforcement of environmental regulations(Burgess et al., 2019; Assunção et al., 2015, 2019a). Assunção et al. (2015) estimate that deforestation rates would have more than doubled in the absence of this package of policies.

7.3.2 Long run

In the results presented thus far, the outcome variables are lagged by five years on average to allow for the effects of electrification on farms to materialize. Our model argues that farmers relinquish cattle grazing to focus on growing crops. This occurs because farmers are capital-constrained and cannot simultaneously expand their business along multiple margins. This raises the possibility that perhaps the gains to native vegetation we observe are short-lived, and once farmers' incomes increase and their credit constraints are relaxed, the land use effects would revert. To examine this possibility, we re-estimate the effects of electrification on land use in a similar way as Table 7, but increase the lag between the measure of electricity infrastructure and the land use outcomes to approximately 25 years.

The results show that the short-run effects on land use do not dissipate over time. Within farms, the effect on pastures (column 1) is similar to the short-run results: A 10 percent increase in electrification lowers the share of farmland allocated to pastures by 3.5 percentage points 25 years later. The effect on cropland (column 2) drops to 0.9 from 1.5 percentage points, and loses statistical significance. The point estimate of the effects on native vegetation (column 3) are now even larger than the short-run effects (though not statistically significantly larger): a 10 percent increase in electrification leads to a 4.8 percentage point increase in the proportion of farmland kept under native vegetation 25 years later. Finally, the point estimate of the effects on farmland expansion (column 4) are smaller: the coefficient drops from 4.7 to 3.7 percentage points for a 10 percent increase in infrastructure, but is still statistically significant (the point estimates are not statistically significantly different) . In summary, an increase in crop productivity still has a land-sparing effect even 25 years later, and the patterns of land use changes between pasture, grains and native vegetation is exactly as predicted by the model.

7.3.3 Effects on Satellite-based Measures of Land Use

We now study the effects of electrification on land use as measured by classification of satellite images taken between 1985 and 2005. The main advantage of satellite measurements is that they contain information on land use in the entire country, not only within farms. They also allow us to study long-run changes, as we merge the electric-

Table 9: Long Run Effects of Electricity on the Allocation of Land: Census of Agriculture Data

Dependent Variable	(1) Pastures / Farmland	(2) Cropland / Farmland	(3) Native Vegetation / Farmland	(4) Farmland / County Area
IV	-0.346^{***}	0.089	0.476^{***}	-0.374^{***}
	[0.080]	[0.064]	[0.128]	[0.107]
OLS	-0.023	-0.020^{**}	0.088^{***}	-0.095^{***}
	[0.019]	[0.008]	[0.025]	[0.022]
Observations	6,516	6,516	6,516	6,516
Mean dep. var.	0.48	0.11	0.16	0.75

This table is similar to Table 7, except that the dependent variables are forward-lagged by two decades. As a result, the number of observations drops because we loose two periods of our panel of counties. The corresponding first-stage coefficient is 0.406 (s.e. 0.083; partial F-statistic 24.11). This Table confirms that the findings of Table 7 do not apply just in the short run. All regressions are weighted by the number of grid points in the county, and include county fixed effects, year fixed effects, and directly control for a quartic polynomial on the suitability index. Standard errors clustered at the county level in brackets.

ity infrastructure data to satellite images captured 25 years later. However, unlike the agricultural census data, these images cannot distinguish between areas within and outside farms, and are therefore not as informative about the mechanisms posited in our theory of land use. When working with satellite data, we therefore measure land use as proportion of the county area rather than fractions of farmland.

Table 10 displays the results. These regressions show strong evidence of pasture-to-crop conversion. The share of county area allocated to pastures decreases by 1 percentage point with a 10 pp increase in electrification (column 1), while the share of cropland increases by 1.3 percentage points. The coefficient estimate on column 3 suggests that the share of forest land decreases by 0.09 percentage points, but this effect is very small and not statistically significant. The satellite data largely corroborates the changes in the land use patterns we observed in the waves of the agricultural census data. Improvements in crop productivity brought about through farm electrification facilitate a shift away from cattle grazing and into crop-farming and indirectly protects the forest, and that protection persists over 25 years.

7.4 Effects on capital usage in crop cultivation

In our land-use model, electrification increases the marginal product of capital—specifically, capital necessary for crop farming. This capital could be one of a number of fixed cost factors of production used more intensively in crop farming than grazing–tractors, irrigation, human capital, etc. Consistent with an increase in the marginal product of capital used for crop production, Table 11 shows evidence that electrification increases usage of capital goods employed mostly in crop cultivation. Columns 1 and 2 show

Table 10: Long Run Effects of Electricity on the Allocation of Land: Satellite Data

Dependent Variable	(1) $\frac{\text{Pastures}}{\text{County Area}}$	(2) $\frac{\text{Cropland}}{\text{County Area}}$	(3) $\frac{\text{Forest}}{\text{County Area}}$
IV	-0.097^{***}	0.127^{***}	-0.009
	[0.021]	[0.035]	[0.028]
OLS	-0.010^{*}	0.008^{**}	0.002
	[0.006]	[0.004]	[0.004]
Observations	6,516	6,516	6,516
Mean dep. var.	0.50	0.07	0.37

This table uses remote sensing data made available by the Mapbiomas project, which classifies pixels from Landsat images yearly from 1985 through 2017 into land use categories. Outcomes are forward-lagged by 25 years with respect to the electricity data. The sample is the same as the sample used in Table 9. All regressions are weighted by the number of grid points in the county, and include county fixed-effects, and directly control for a quartic polynomial on the hydropower suitability index interacted with time fixed-effects.
Standard errors clustered at the county level in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

that farmers invest more in crop-related electricity-intensive capital equipment–grain storage facilities (which requires temperature and humidity control) and irrigation–as electricity becomes available. These two capital goods are particularly informative about the specific mechanism posited in our model, as not only they are related to crops, they also directly require energy which is facilitated by electrification. The mean effects are large: a 10 percent increase in electricity infrastructure leads to a 36-percent increase (at the mean) in the number of grain silos and a 19-percent increase in the proportion of farms with irrigation.

Columns 3–5 show that electrification increases the usage of other crop-related capital goods that do not directly rely on electricity. This is consistent with farmers intensifying their cropping activities after electricity arrives. The IV point estimates imply mean effects of 8 percent for planting and harvesting machines, 10 percent for plows, and 2.6 percent for tractors (the coefficient for tractors is not statistically significant). These results are consistent with the intensification of cropping highlighted in our model, which is the specific mechanism by which demand for agricultural land decreases within farms. Increased electricity infrastructure enables farmers to adopt technologies and employ capital that would not be feasible otherwise.

7.5 Testing the Capital Constraints Assumption

A key assumption in our theoretical framework is that farmers' ability to respond to productivity shocks is limited because of factor market imperfections. We model these imperfections as credit constraints, which is a well documented feature of developing economies, especially in rural contexts (see, e.g., Banerjee and Duflo, 2005; Burgess and Pande, 2005; Conning and Udry, 2007). Furthermore, as discussed in section 2, access

Table 11: The Effects of Electricity on Crop-related Capital Use

Dependent Variable	Energy-intensive capital goods		Other crop-related capital goods		
	(1) Grain Storage Facilities	(2) Percentage of Farms with Irrigation	(3) Planting and Harvesting Machines	(4) Plows	(5) Tractors
IV	102.976***	12.738***	7.955**	21.476***	6.561
	[17.530]	[2.820]	[3.631]	[5.570]	[4.528]
OLS	13.647***	2.597***	0.231	1.333	1.300
	[2.827]	[0.453]	[1.113]	[1.475]	[1.302]
Observations	10,860	10,860	10,860	10,860	10,860
Mean dep. var.	27.34	6.16	10.11	20.46	22.94

Notes: The table shows evidence, consistent with our model's assumptions and predicitons, that electricity infrastructure increases the usage of capital that supports crop cultivation, in particular grains. Grain storage facilities, machines, plows, tractors are normalized by county area (measured in 10,000 hectares). In each column, the first row shows results from our preferred IV-fixed effect specification, and the second row shows OLS, fixed-effects results. All regressions are weighted by the number of grid points in the county, and include county fixed effects, year fixed effects, and directly control for a quartic polynomial on the suitability index (see Table 5). Standard errors clustered at the county level in brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

to credit by Brazilian farmers has been historically low. We now provide evidence of credit constraints faced by Brazilian farmers during our sample period, using two different approaches and data sources.

First, we study how farming investments respond to rainfall shocks in the recent past, which we treat as an exogenous source of income for farmers that effectively mimics the experimental ideal of granting cash to randomly selected farmers (Rosenzweig and Wolpin, 1993). In the absence of credit market imperfections, past fluctuations in rainfall (or cash drops) should not affect current investment decisions (Karlan et al., 2014). We collect historical precipitation data from Matsuura and Willmott (2012) and construct a measure of rainfall shocks based on deviations from a county-specific historical means (details in appendix C). We then run regressions of the form:

$$y_{ct} = \sum_{k=-4}^{0} \beta_k r_{c,t-k} + \delta X_{c,t} + \theta_c + \alpha_t + \epsilon_{c,t} \quad , \tag{9}$$

where $r_{c,t-k}$ is our measure of a rainfall shock in county c, period $t-k$.

Table 12 shows that transitory rainfall shocks have persistent effects on yields, and on investment in capital goods. Investment in tractors, plows, planting and harvesting machines and grain storage facilities all initially decrease, but then rebound to increase by more than the initial fall in periods following positive rainfall shocks. Grain yields are larger two decades after the rainfall shock, even after we control for contemporaneous rainfall. These persistent effects on investments are inconsistent with perfect credit markets.

Table 12: Testing Credit Constraints using Rainfall Shocks

	(1) Grain Yields (log)	(2) Value of Production Per Hectare (log)	(3) Gross Production Value (log)	(4) Tractors	(5) Plows	(6) Planting and Harvesting Machines	(7) Grain Storage Facilities
shock at t	0.010***	0.011**	0.019***	-1.102^{***}	-0.474^{**}	-1.033^{***}	-7.507^{***}
	[0.003]	[0.005]	[0.005]	[0.163]	[0.222]	[0.142]	[0.667]
shock at t-1	0.045***	0.015**	0.009	1.783***	-0.078	0.239	-1.861^{**}
	[0.004]	[0.006]	[0.006]	[0.192]	[0.233]	[0.199]	[0.889]
shock at t-2	0.017***	0.029***	0.013***	1.844***	3.186***	1.348***	12.245***
	[0.003]	[0.005]	[0.005]	[0.172]	[0.300]	[0.150]	[0.979]
shock at t-3	-0.019^{***}	0.021***	0.029***	-0.298	0.593**	-0.149	3.744***
	[0.003]	[0.006]	[0.005]	[0.189]	[0.236]	[0.166]	[0.837]
shock at t-4	-0.004	0.007	0.017***	1.147***	0.787***	1.792***	-1.058
	[0.004]	[0.006]	[0.006]	[0.193]	[0.300]	[0.206]	[0.654]
Observations	10,701	10,855	10,855	10,855	10,855	10,855	10,855
Mean dep. var.	0.0657	12.54	16.61	22.94	20.46	10.11	27.35

This table tests our model's assumption that farmers are constrained in (at least) one production factor other than land. We use past rainfall shocks as a measure of (as good as random) capital drops to farmers. Columns 1–3 show that past (positive) rainfall shocks affect current production, even after we control for current shocks. The remaining columns show that these past rainfall shocks affect investment: Grain storage facilities, machines, plows, tractors are divided by county area (measured in 10,000 hectares). Taken together, these results indicate that exogenous variation in past income affect current production decisions. This is not consistent with an economy without frictions in credit markets. To construct our measure of rainfall shock, we use historical precipitation data from Matsuura and Willmott (2012). See appendix C for details. All regression include county fixed effects and year fixed effects. Controlling for a quartic polynomial on the suitability index interacted with year dummies (like we do in the analysis using the IV-fixed effect approach) does not change the results. Standard errors clustered at county level in brackets.

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

Table 13: The Effects of Electricity on the Allocation of Land: The role of banks

Dependent Variable	(1) Pastures / Farmland	(2) Cropland / Farmland	(3) Native Veg / Farmland	(4) Farmland / County Area
Panel A: IV estimates				
Electricity Infrastructure	-0.384^{***}	0.151^{***}	0.395^{***}	-0.493^{***}
	[0.130]	[0.054]	[0.132]	[0.140]
Electricity Infrastructure x Bank Branches	0.336^{*}	-0.120	-0.264^{*}	0.573^{**}
	[0.188]	[0.087]	[0.155]	[0.243]
Panel B: Effects at the 75th percentile of Bank Branches distribution (0.55 banks per 10,000 hectares)				
	-0.199^{**}	0.085^{**}	0.250^{***}	-0.177
	[0.092]	[0.042]	[0.088]	[0.118]
Observations	10,860	10,860	10,860	10,860
Mean dep. var.	0.47	0.11	0.18	0.71

This table shows how the heterogeneity of treatment effects on land-use by how credit-constrained farmers are. We use the density of bank branches (number of bank branches per 10,000 hectares) to proxy for access to credit. Panel A shows the IV coefficients. We our modeled infrastructure variable interacted with the number of bank branches per 10,000 hectares as an additional instrument for the interaction term. Consistent with the implications of the model presented in section 4, the interaction term is roughly the size of the coefficient on electricity infrastructure: ie 1 bank branch per 10,000 hectares reduces the effect of full expansion of electricity access to nearly 0. In panel B, we compute the implied effects of electricity infrastructure for a county at the 75th percentile of the distribution of bank branches density (0.55 bank branches per 10,000 hectares. The 25th percentile is zero bank branches per 10,000 hectares. All regressions are weighted by the number of grid points in the county, and include county fixed effects, year fixed effects, and directly control for a quartic polynomial on the suitability index (see Table 5). Standard errors clustered at the county level in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We construct a second test of credit market imperfections using the idea that in our model, farmers who are not credit constrained should not have to substitute pastures for cropland upon receiving a positive productivity shock to crop cultivation. We test that prediction by adding an interaction term between electricity infrastructure and the density of bank branches in the county to our main regression.[23] If the number of bank branches is a good proxy for local credit availability,[24] then we should observe smaller substitution away from pastures in such areas when electricity arrives.

Table 13 shows that the presence of banks change the effect of electrification on land-use in exactly that direction. In Panel A, columns 1 and 2 show that an increase in credit availability leads to less substitution of pastureland for cropland. The coefficients on the interaction terms are opposite-signed to the main effects, which implies that electrification has smaller effects on pastures and cropland in areas where credit is more available to farmers. As a result, the productivity shock has a less pronounced effect on native vegetation, as indicated by the negative coefficient on the interaction term

[23] In this specification, we directly control for the density of bank branches, and instrument the interaction term using the IV of our modeled infrastructure variable interacted with bank branches.

[24] This exercise is inspired by Burgess and Pande (2005), who study a government-led expansion of bank branches in India and show that the presence of banks expands access to credit in rural areas. Like India, the expansion of banking in Brazil was also led by public banks Sanches et al. (2018). Our exercise is analogous to that of Jayachandran (2006), who studies how the sensitivity of agricultural wages in India with respect to crop yields changes with access to banking.

in column 3. Panel B shows that in counties at the 75th percentile of the distribution of bank branches density (0.55 bank branches per 10,000 hectares), electrification produces only two thirds the effect on native vegetation compared to counties at the mean. Finally, column 4 shows that counties with more bank branches per hectare display a faster expansion of farmland than counties with less access to capital, highlighting the importance of credit to agriculture. These results are in line with the mechanism we posit: that capital constraints are important for restricting the expansion of cattle grazing. Assunção et al. (2019b) study the effects of a policy that restricts Brazilian farmers' credit access, and also finds that it reduces deforestation.

8 Conclusion

We show that electrification and the resulting agricultural productivity gains to cropping in agriculture slowed the rate of deforestation in Brazil. We present evidence that the underlying mechanism for these effects is changes in land use within farms, as well as a slower expansion of farmland once productivity improves via rural electrification. Access to electricity helped to modernize agriculture in Brazil as it enabled farmers to abandon land-intensive practices, and operate with higher capital-to-land ratios. Importantly, this helped preserve an important natural resource. This decrease in deforestation relies on two key contextual features highlighted by our model: (1) the productivity shock differentially impacted cropping over cattle grazing which requires more land, and (2) factor (credit in our context) constraints which made it difficult for farmers to respond to productivity improvements on the extensive margin.

Our analysis has some limitations. First, forest cover is not the only relevant environmental outcome when studying the intensification of agriculture. The use of pesticides and fertilizers, typical in intensive agriculture, can also impose external costs on plant, animal, and human populations (Dias et al., 2019; Dasgupta, 2021). In that sense, the aggregate, net environmental benefits of intensification may be smaller than our estimates suggest. On the other hand, our analysis ignores the (likely) positive effects of improving productivity of Brazilian farmers on forests in other countries. Brazil is a major commodity producer–"the world's food basket"–, it is likely that an increase in Brazilian agricultural productivity has the potential to spare land elsewhere through general equilibrium effects as global beef prices respond to decreases in land allocations for grazing (Baylis et al., 2013). In addition, electricity may cause structural change in the manufacturing sector impacting the demand for labor and/or capital which could also impact the agricultural sector–these general equilibrium effects are not observable given our data and are outside of the scope of our study.

These limitations notwithstanding, our results have important implications for environmental policymaking. Many popular conservation policies, such as designating areas for protection, can cause "leakage" by displacing deforestation to unregulated areas where enforcing fines and bans are difficult (Burgess et al., 2012, 2019; Harstad and Mideksa, 2017). Our results show that conservation policies can be successful if they account for the economic interest of user groups. Governments and other environmental organizations are increasingly experimenting with approaches such as direct payments for ecosystem services (Porras et al., 2012; Jayachandran et al., 2017) or interventions that improve farm productivity. Our findings suggest that support to increase farm productivity has the potential to reduce deforestation.

References

Abman, R. and Carney, C. (2020). Agricultural productivity and deforestation: Evidence from input subsidies and ethnic favoritism in Malawi. *Journal of Environmental Economics and Management*, 103.

Abman, R., Garg, T., Pan, Y., and Singhal, S. (2020). Agriculture and deforestation. *Available at SSRN 3692682*.

Alix-Garcia, J., McIntosh, C., Sims, K. R. E., and Welch, J. R. (2013). The Ecological Footprint of Poverty Alleviation: Evidence from Mexico's Oportunidades Program. *Review of Economics and Statistics*, 95(2):417–435.

Angelsen, A. and Kaimowitz, D., editors (2001). *Agricultural Technologies and Tropical Deforestation*. CABI Publiching.

Araujo, R., Costa, F., and Sant'Anna, M. (2020). Efficient forestation in the brazilian amazon. *SocArXiv. December*, 10.

Asher, S., Garg, T., and Novosad, P. (2020). The Ecological Impact of Transportation Infrastructure. *The Economic Journal*, 130(629):1173–1199.

Assunção, J., Gandour, C., and Rocha, R. (2023). Deter-ing deforestation in the amazon: environmental monitoring and law enforcement. *American Economic Journal: Applied Economics*, 15(2):125–156.

Assunção, J., Gandour, C., and Rocha, R. (2015). Deforestation slowdown in the Brazilian Amazon: prices or policies? *Environment and Development Economics*, 20(6):697–722.

Assunção, J., Gandour, C., Rocha, R., and Rocha, R. (2019a). DETERring Deforestation in the Amazon: Environmental Monitoring and Law Enforcement. Climate Policy Initiative.

Assunção, J., Gandour, C., Rocha, R., and Rocha, R. (2019b). The Effect of Rural Credit on Deforestation: Evidence from the Brazilian Amazon. *The Economic Journal*, 130(626):290–330.

Badiani, R. and Jessoe, K. (2013). The impact of electricity subsidies on groundwater extraction and agricultural production. *Department of Agriculture and Resource Economics Working Paper, University of California Davis. Retrieved*.

Balboni, C., Burgess, R., and Olken, B. A. (2021). The origins and control of forest fires in the tropics. Technical report, Tech. Rep.

Banerjee, A., Duflo, E., and Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics*, 145:102442.

Banerjee, A. V. and Duflo, E. (2005). Growth Theory through the Lens of Development Economics. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1 of *Handbook of Economic Growth*, chapter 7, pages 473–552. Elsevier.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). Roads, Railroads, and Decentralization of Chinese Cities. *The Review of Economics and Statistics*, 99(3):435–448.

Baylis, K., Fullerton, D., and Shah, P. (2013). What drives forest leakage. *Work. Pap., Dep. Agric. Consum. Econ., Univ. Ill., Urbana-Champaign*.

Burgess, R., Costa, F. J. M., and Olken, B. (2019). The Brazilian Amazon's Double Reversal of Fortune. Mimeo.

Burgess, R., Hansen, M., Olken, B. A., Potapov, P., and Sieber, S. (2012). The Political Economy of Deforestation in the Tropics. *The Quarterly journal of economics*, 127(4):1707–1754.

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A., and Padró i Miquel, G. (2015). The Value of Democracy: Evidence from Road Building in Kenya. *American Economic Review*, 105(6):1817–51.

Burgess, R. and Pande, R. (2005). Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment. *American Economic Review*, 95(3):780–795.

Burlig, F. and Preonas, L. (2016). Out of the darkness and into the light? Development effects of Rural Electrification. Working Paper 268, Energy Institute at Haas.

Bustos, P., Caprettini, B., and Ponticelli, J. (2016). Agricultural Productivity and Structural Transformation: Evidence from Brazil. *American Economic Review*, 106(6):1320–65.

Chakravorty, U., Emerick, K., and Ravago, M.-L. (2016). Lighting Up the Last Mile: The Benefits and Costs of Extending Electricity to the Rural Poor. Discussion Paper 16-22, Resources for the Future. Available at http://dx.doi.org/10.2139/ssrn.2851907.

Conning, J. and Udry, C. (2007). Rural Financial Markets in Developing Countries. In Evenson, R. and Pingali, P., editors, *Handbook of Agricultural Economics*, volume 3 of *Handbook of Agricultural Economics*, chapter 56, pages 2857–2908. Elsevier.

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

Cropper, M., Puri, J., and Griffiths, C. (2001). Predicting the Location of Deforestation: The Role of Roads and Protected Areas in North Thailand. *Land Economics*, 77(2):172–186.

Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A., and Hansen, M. C. (2018). Classifying drivers of global forest loss. *Science*, 361(6407):1108–1111.

Dasgupta, P. (2021). *The Economics of Biodiversity: the Dasgupta Review.* HM Treasury, London.

Dias, M., Rocha, R., and Soares, R. R. (2019). Glyphosate Use in Agriculture and Birth Outcomes of Surrounding Populations. Discussion Paper 12164, IZA.

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

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

Ehrl, P. (2017). Minimum comparable areas for the period 1872-2010: an aggregation of Brazilian municipalities. *Estudos EconÃ´micos*, pages 215–229.

Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies*, 81(3):1046–1070.

Fink, G., Jack, B. K., and Masiye, F. (2020). Seasonal Liquidity, Rural Labor Markets and Agricultural Production. *American Economic Review*, forthcoming.

Fluck, R. C. (1992). Energy in World Agriculture . In Fluck, R. C., editor, *Energy in Farm Production* , Energy in World Agriculture. Elsevier, Amsterdam.

Foster, A. D. and Rosenzweig, M. R. (2003). Economic Growth and the Rise of Forests. *The Quarterly Journal of Economics*, 118(2):601–637.

Frank, E. G. and Schlenker, W. (2016). Balancing economic and ecological goals. *Science*, 353(6300):651–652.

Giné, X., Goldberg, J., and Yang, D. (2012). Credit market consequences of improved personal identification: Field experimental evidence from malawi. *American Economic Review*, 102(6):2923–2954.

Gonzalez Lira, A. and Mobarak, A. M. (2019). Slippery fish: Enforcing regulation under subversive adaptation.

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(2):353–377.

Harding, T., Herzberg, J., and Kuralbayeva, K. (2021). Commodity prices and robust environmental regulation: Evidence from deforestation in Brazil. *Journal of Environmental Economics and Management*, 108:102452.

Harstad, B. and Mideksa, T. K. (2017). Conservation contracts and political regimes. *The Review of Economic Studies*, 84(4):1708–1734.

Hidalgo, F. D., Naidu, S., Nichter, S., and Richardson, N. (2010). Economic determinants of land invasions. *The Review of Economics and Statistics*, 92(3):505–523.

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.

IBGE (2006). Censo agropecuário 2006.

Jayachandran, S. (2006). Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy*, 114(3):538–575.

Jayachandran, S. (2013). Liquidity Constraints and Deforestation: The Limitations of Payments for Ecosystem Services. *The American Economic Review*, 103(3):309–313.

Jayachandran, S., de Laat, J., Lambin, E. F., Stanton, C. Y., Audy, R., and Thomas, N. E. (2017). Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science*, 357(6348):267–273.

Jeong, D. (2020). Creating (Digital) Labor Markets in Rural Tanzania. Mimeo.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). Agricultural Decisions after Relaxing Credit and Risk Constraints. *The Quarterly Journal of Economics*, 129(2):597–652.

Lee, D. S., McCrary, J., Moreira, M. J., and Porter, J. (2020a). Valid t-ratio Inference for IV. Papers 2010.05058, arXiv.org.

Lee, D. S., McCrary, J., Moreira, M. J., Porter, J. R., and Yap, L. (2023). What to do when you can't use'1.96'confidence intervals for iv. Technical report, National Bureau of Economic Research.

Lee, K., Miguel, E., and Wolfram, C. (2020b). Does Household Electrification Supercharge Economic Development? *Journal of Economic Perspectives*, 34(1):122–44.

Lee, K., Miguel, E., and Wolfram, C. (2020c). Experimental Evidence on the Economics of Rural Electrification. *Journal of Political Economy*, 128(4):1523–1565.

Lewis, J. and Severnini, E. (2020). Short-and long-run impacts of rural electrification: evidence from the historical rollout of the us power grid. *Journal of Development Economics*, 143.

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.

Lipscomb, M., Mobarak, A. M., Barham, T., and Szerman, D. (2021). "another look at the precision of iv estimates of the development effects of access to electricity in brazil". available at $https : //www.mollylipscomb.com/_files/ugd/a200af_629245100d8f46f592f171ada715c492.pdf$.

Matsuura, K. and Willmott, C. (2012). Terrestrial Precipitation: 1900-2010 Gridded Monthly Time Series (1900 - 2010) (V 3.01 added 6/14/12). Technical report, University of Delaware. University of Delaware. http://climate.geog.udel.edu/ climate/.

Oldekop, J. A., Sims, K. R., Karna, B. K., Whittingham, M. J., and Agrawal, A. (2019). Reductions in deforestation and poverty from decentralized forest management in nepal. *Nature sustainability*, 2(5):421–428.

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.

Phalan, B., Green, R. E., Dicks, L. V., Dotta, G., Feniuk, C., Lamb, A., Strassburg, B. B. N., Williams, D. R., Zu Ermgassen, E. K. H. J., and Balmford, A. (2016). How can higher-yield farming help to spare nature? *Science*, 351(6272):450–451.

Pimentel, D. (2009). Energy inputs in food crop production in developing and developed nations. *Energies*, 2(1):1–24.

Porras, I., Miranda, M., Barton, D., Chacon-Cascante, A., et al. (2012). Payments for environmental services in Costa Rica: from Rio to Rio and beyond. *IIED Briefing Paper-International Institute for Environment and Development*.

Reis, E., Pimentel, M., Alvarenga, A., and Santos, M. C. H. (2011). Áreas mínimas comparáveis para os períodos intercensitários de 1872 a 2000. *I Simpósio Brasileiro de Cartografia histórica*.

Rosenzweig, M. and Wolpin, K. I. (1993). Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income

Countries: Investment in Bullocks in India. *Journal of Political Economy*, 101(2):223–44.

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

Sanches, F., Silva Junior, D., and Srisuma, S. (2018). Banking privatization and market structure in Brazil: a dynamic structural analysis. *The RAND Journal of Economics*, 49(4):936–963.

Sekhri, S. (2011). Public Provision and Protection of Natural Resources: Groundwater Irrigation in Rural India. *Amercian Economic Journal: Applied Economics*, 3(4):29–55.

Smith, P., Bustamante, M., Ahammad, H., Clark, H., Dong, H., Elsiddig, E., Haberl, H., Harper, R., House, J., Jafari, M., Masera, O., Mbow, C., Ravindranath, N., Rice, C., Abad, C., Romanovskaya, A., Sperling, F., Tubiello, F., and Bolwig, S. (2014). *Agriculture, Forestry and Other Land Use (AFOLU)*, pages 811–922. Cambridge University Press, United Kingdom.

Souza-Rodrigues, E. (2019). Deforestation in the amazon: A unified framework for estimation and policy analysis. *The Review of Economic Studies*, 86(6):2713–2744.

Staiger, D. and Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3):557–586.

Ulyssea, G. (2010). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics*, 91(1):87 – 99.

Usmani, F. and Fetter, T. R. (2020). Fracking, Farmers, and Rural Electrification in India. Technical report, Ruhr Economic Paper No. 864.

World Bank (1990). Brazil - Irrigation subsector review. Sector Report 7797, World Bank.

World Bank (2007). *World Development Report 2008: Agriculture for Development*. The World Bank.

Zomer, R. J., Trabucco, A., Coe, R., and Place, F. (2014). Trees on farms: an update and reanalysis of agroforestry's global extent and socio-ecological characteristics. *World Agroforestry Center Working Paper 179*.

Appendices

A Derivations

A.1 Solution for model in section 4

Proposition 1. *The optimal farmer's choices, $H_c^*(\Omega)$, $H_g^*(\Omega)$, $K^*(\Omega)$, satisfy equations (2a)–(2d).*

Proof. The solution to the farmer's problem is given by the set of first-order conditions

$$\begin{aligned}
&\text{wrt } H_c: && \Omega K^* F_H(H_c^*) = (1+\lambda)p && (10)\\
&\text{wrt } H_g: && F_H(H_g^*) = (1+\lambda)p && (11)\\
&\text{wrt } K: && \Omega F(H_c^*) = (1+\lambda)r && (12)\\
&\text{constraint} && rK^* + p(H_c^* + H_g^*) - M = 0. && (13)
\end{aligned}$$

Result 1. $\frac{\partial K^*}{\partial \Omega}$ *and* $\frac{\partial H_c^*}{\partial \Omega}$ *have the same sign.*

Proof. It suffices to show that $\frac{\partial K^*}{\partial H_c^*} > 0$. To see this, note that (10) and (12) imply that $K^* = \frac{p}{r}\frac{F(H_c^*)}{F_H(H_c^*)}$; Since we assume that $F_{HH} < 0$ and $F_H > 0$, we have that

$$\frac{\partial K^*}{\partial H_c^*} = \frac{p}{r}\frac{\left(F_H(H_c^*)\right)^2 - F(H_c^*)F_{HH}(H_c^*)}{\left(F_H(H_c^*)\right)^2} > 0 \tag{14}$$

□

Result 2. *Let $\Phi(\Omega) \equiv \Pi(\Omega, K^*(\Omega), H_c^*(\Omega), H_g^*(\Omega))$. $\Phi(\Omega)$ is strictly convex.*

Proof. We must show that, for all Ω and h, $\Phi(\Omega+h) > \Phi(\Omega) + h\Phi'(\Omega)$.

$$\begin{aligned}
\Phi(\Omega+h) &= \Pi(\Omega+h, K^*(\Omega+h), H_c^*(\Omega+h), H_g^*(\Omega+h))\\
&> \Pi(\Omega+h, K^*(\Omega), H_c^*(\Omega), H_g^*(\Omega)) \quad \text{(by definition of } K^*, H_c^*, H_g^*\text{)}\\
&= (\Omega+h)K^*F(H_c^*) + F(H_g^*) - rK^* - p(H_c^* + H_g^*)\\
&= \Pi(\Omega, K^*(\Omega), H_c^*(\Omega), H_g^*(\Omega)) + hK^*F(H_c^*)\\
&= \Phi(\Omega) + h\Phi'(\Omega) \quad \text{(by the envelope theorem)}
\end{aligned}$$

□

Now, since $\Phi(\Omega)$ is convex, $\frac{\partial^2\Phi}{\partial\Omega^2} > 0$. But

$$\frac{\partial^2\Phi}{\partial^2\Omega} > 0 \iff \frac{\partial K^*}{\partial\Omega} F(H_c^*) + K^* \left(F_H(H_c^*) \frac{\partial H_c^*}{\partial\Omega} \right) > 0$$

Since $F(), F_H()$, and $K^* > 0$, either $\frac{\partial K^*}{\partial\Omega}$ or $\frac{\partial H_c^*}{\partial\Omega}$, or both, are positive. By result 1, we conclude that both must be positive. That proves equations 2a and 2b.

Once we have established equations 2a and 2b, it is easy to show equations 2d and 2d. From equation 13:

$$\begin{aligned}
\frac{\partial K^*}{\partial\Omega} &= -\frac{p}{r}\frac{\partial(H_c^* + H_g^*)}{\partial\Omega} > 0 \\
&\iff \frac{\partial(H_c^* + H_g^*)}{\partial\Omega} < 0 \quad \text{(which proves equation 2d)} \\
&\Rightarrow \frac{\partial H_g^*}{\partial\Omega} < 0 \quad \text{(which proves equation 2c)}
\end{aligned}$$

□

A.2 Derivation of equation 8

Let V_T denote total native vegetation, V_I native vegetation inside farms, and $k(C - F)$ native vegetation outside of farms, expressed as a fraction $k \in [0, 1]$ of the difference between county area C and farmland F. This notation gives us an accounting identity for native vegetation: $V_T = V_I + k(C - F)$. Dividing both sides of this identity by C and multiplying and dividing the first term by F yields

$$\begin{aligned}
\frac{V_T}{C} &= \frac{V_I}{C} + k\frac{C - F}{C} \\
\iff \frac{V_T}{C} &= \frac{V_I}{F}\frac{F}{C} + k\frac{C - F}{C} \\
\iff \frac{\partial V_T/C}{\partial\Omega} &= \frac{\partial V_I/F}{\partial\Omega}\frac{F}{C} + \frac{\partial F/C}{\partial\Omega}\frac{V_I}{F} - k\frac{\partial F/C}{\partial\Omega} \\
\iff \frac{\partial V_T/C}{\partial\Omega} &= \frac{\partial F/C}{\partial\Omega}(\frac{V_I}{F} - k) + \frac{\partial V_I/F}{\partial\Omega}\frac{F}{C}
\end{aligned}$$

B Instrument construction: Probit

Table B.1 shows the probit regression used to rank the grid points in terms of their hydropower suitability (see section 6).

Table B.1: Probit Regression for Hydropower Geographic Cost Parameters

Indicator for location has a river	0.012
	[0.067]
Log of maximum flow accumulation	0.018
	[0.013]
Average river slope	0.039
	[0.031]
Maximum river slope	0.059***
	[0.012]
Amazon indicator	−0.630***
	[0.106]
Observations	32,578

Notes: The dependent variable is an indicator for location has a hydropower plant. Standard errors clustered by county brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Rainfall Data

In section 7.5, we use past rainfall realizations as exogenous shocks to farmer income. We measure rainfall using historical precipitation data from Matsuura and Willmott (2012). This is a gridded dataset with monthly precipitation measures from 1900 to 2012 with a 0.5-degree spatial resolution (approximately 50 kilometers at the equator).

To construct our measure of rainfall shock, we take a procedure similar to other to other studies in the literature (see, e.g., Hidalgo et al., 2010). First, we assign each county to the grid point that is closest to the county's centroid. We then standardize every county-month observation using the county's historical data. Next, we sum up these standardized monthly measure to the yearly level using county-specific month-of-the-year weights. These monthly weights are calculated from the 1985 Census of Agriculture, and capture the importance of each month of the year for that county's planting season. The weights are equal to the share of each crop on the country's agricultural production (in Reais) times the share of farmers that report planting that crop on that month. Finally, we standardize the annual observations by county-year. Our measure of rainfall shock therefore is the deviation from a county-specific historical means.

C.1 Instrument Validity

we check whether placement of power plants and transmission lines simulated by the forecasting model can be predicted by farm production indicators in earlier years by estimating equations of the form $instrument_{it} = \beta outcome_{i,t-1} + \lambda_t + \theta_i + \xi_{it}$.. Under the null that our instrument is as good as random the estimates of β should be statistically equal to zero. Results are shown in Table C.1. Almost all coefficients are small in magnitude, suggesting that pre-existing production factors are not a major factor in determining electricity placement.

The largest coefficient is the one associated with the share of cropland, which implies that converting 100% of farm area into crops in one decade increases our instrument 0.2 a decade later. Four other coefficients are statistically significant, but their magnitudes are small. Overall, past agricultural outcomes do not appear to be strong predictors of the instrument.

D Robustness Checks

D.1 Lee, Moreira, McCrary, and Porter (2020) Confidence Intervals

Table C.1: Instrument Validty

	Dependent Variable: Instrument at t	
Outcomes at $t-1$	coefficient	std. error
Land Use – Census of Agriculture		
Farmland/County Area	-0.042^{***}	0.011
Pastures/Farmland	-0.028	0.020
Cropland/Farmland	0.198^{***}	0.036
Native Vegetation/Farmland	-0.005	0.010
Land Use – Satellite Data		
Pastures/County Area	0.025	0.031
Cropland/County Area	0.084	0.061
Forest/County Area	-0.038	0.040
Farm Production and Capital Usage		
Grain Yields (log)	0.017^{***}	0.005
Heads of Cattle Per Hectare	0.006^{**}	0.002
Share of Cattle Production Value	-0.037^{***}	0.011
Grain Storage Facilities	-0.000	0.000
Percentage of Farms with Irrigation	0.002^{***}	0.001
Planting and Harvesting Machines	-0.000	0.000
Plows	-0.000	0.000
Tractors	-0.000	0.000

This table shows the extent to which our instrument—the forecast expansion of electricity infrastructure—can be predicted by agricultural indicators in earlier years. Each row shows results from a regression of the form $instrument_{it} = \beta outcome_{i,t-1} + \theta_i + \lambda_t + \varepsilon_{it}$. Sample size of each regression is 6,516. Standard errors clustered by county.

Table D.1: Lee, Moreira, McCrary, and Porter 95% Confidence Intervals

Dependent Variable	$\widehat{\beta}^{IV}$	95% CI
Grain Yields (log)	0.948	[0.49 , 1.40] (0.41 , 1.47)
Heads of Cattle Per Hectare	−0.399	[−0.91 , 0.11] (−0.99 , 0.187)
Share of Cattle Production Value	−0.242	[−0.41 , −0.07] (−0.43 , −0.05)
Percent of County Area in Farmland	−0.472	[−0.73 , −0.21] (−0.78 , −0.16)
Percent of Farmland in Pastures	−0.371	[−0.62 , −0.13] (−0.66 , −0.08)
Percent of Farmland in Grains	0.146	[0.04 , 0.25] (0.03 , 0.26)
Percent of Farmland in Native Vegetation	0.385	[0.13 , 0.64] (0.09 , 0.68)
Percentage of Farms with Irrigation	12.738	[7.21 , 18.26] (12.09 , 13.39)
Grain Storage Facilities	102.976	[68.62 , 137.33] (62.66 , 143.30)
Planting and Harvesting Machines	7.955	[0.84 , 15.07] (−0.20 , 16.11)
Plows	21.476	[10.56 , 32.39] (8.67 , 34.29)
Tractors	6.561	[−2.31 , 15.44] (−3.86 , 16.98)

Notes: The table shows 95% confidence intervals based on clustered standard errors reported in the main text in brackets, and the Lee et al (2020a) 95% confidence intervals in parentheses. As described in Lee et al. (2023) the tF procedure in most cases leads to confidence intervals that are too large. Each row shows the results for one outcome.

D.2 Controlling for Geographical Factors

The presence of county-fixed effects in our analysis ensures that identification of treatment effects uses only within-county variation. However, one possible concern with our empirical exercise relates to the variation introduced by the geographic factors used in the instrument: water flow, river gradient and the Amazon. To mitigate these concerns, our preferred specification adds a control for a polynomial of the hydropower suitability index interacted with time fixed-effects. The idea is to remove possible trends that are specific to places with high, or low, suitability for hydropower both in our hypothetical grid expansion and in the outcomes that we analyze. In this Appendix, we provide a more detailed analysis of what happens to our instrument when we control for different combinations of geographical and time-varying factors.

Table D.2 reports the first-stage of our 2SLS estimation when we control for all possible combinations of the geographical factors, time-varying factors, and their interactions. [25] Row 0 reports the first stage results when we control only for county fixed-effects and time fixed-effects, reproducing table 5, column 1 of the main text. Next, row 1 directly controls for our measure of water flow interacted with the time-varying decade budget. The coefficient and partial F-statistic barely change. Row 2 controls for interactions of river gradient with decade budget, and so on.

Table D.2 reveals one weakness of our instrument. Controlling for interactions between the Amazon dummy and time-varying factor renders our instrument weak. This can be seen in rows 3–5, 8, 10, 11, and 13. The Amazon-specific time trends simply absorb too much variation from our instrument. That can be explained by the fact that the Amazon plays an important role in the probit that determines the hydropower suitability of grid points, because it is expensive to build hydropower plants in these counties due to dense forests. Our instrument, therefore, displays little variation in those counties, and actual electricity infrastructure is also scarce.

One may wonder, therefore, if Amazonian counties could be removed from the sample. Although there are few Amazonian counties in the sample (130 out of 2,172), these counties are large in area. Since our regressions use the number of grid points in the county as weights, these counties effectively represent 59 percent of our weighted sample. In this Appendix, we show how sensitive our results are to the presence of Amazonian counties in the sample.

To assess our result's sensitivity to the exclusion of the Amazon from the sample, we group the Amazonian counties into three groups – one for each of the two largest Amazonian states, Amazonas and Para, and one group for all other states. Each of

[25] (Lipscomb et al., 2013) perform a similar exercise, and table 1 of (Lipscomb et al., 2021) presents the table for various versions of Amazon controls.

Table D.2: Sensitivity Analysis by Directly Controlling for Geographic Factors in the First Stage.

Specification (description of control set added to RHS)	(1) Electricity Infrastructure	(2) Partial F-stat
0. No direct controls	0.456*** [0.056]	66.99
1. Water flow × decade budget	0.457*** [0.056]	67.20
2. River gradient × decade budget	0.446*** [0.055]	65.20
3. Amazon dummy × decade budget	0.139*** [0.049]	8.12
4. Water flow × decade budget and Amazon dummy × decade budget	0.140*** [0.049]	8.02
5. River gradient × decade budget and Amazon dummy × decade budget, water flow ×budget	0.139*** [0.049]	7.93
6. River gradient × decade budget and water flow × decade budget	0.445*** [0.055]	65.52
7. Water flow × year dummies	0.457*** [0.056]	67.16
8. Amazon dummy × year dummies	0.136*** [0.049]	7.61
9. River gradient × year dummies	0.444*** [0.055]	64.32
10. Water flow × year dummies and Amazon dummy × year dummies	0.138*** [0.049]	7.76
11. River gradient × year dummies and Amazon dummy × year dummies	0.135*** [0.50]	7.39
12. Water flow × year dummies and river gradient × year dummies	0.445*** [0.055]	64.73
13. River gradient × year dummies, water flow × year dummies, and Amazon dummy × year dummies	0.135*** [0.050]	7.35
14. Quartic suitability rank × year dummies	0.385*** [0.064]	36.29

Notes: This table reports the first-stage for various specifications that directly control for geographic factors and their interactions with time fixed-effects. The dependent variable is the prevalence of electricity infrastructure in the county (fraction of grid points within 50 kilometers of a transmission substation) measured from infrastructure inventories. Column 1 reports the coefficient and standard errors clustered at the county level in brackets. Column 2 reports the partial F-statistic. In table 5 in the text, we report row 0 in column 1 , and row 16 in column 2. All specifications include decade fixed effects and AMC fixed effects * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

these groups roughly represent one-third of the Amazon land area. We then run the first-stage regression excluding from the sample one subset of these groups at a time.

Table D.3 reports the results. Column 1 reproduces our baseline result (shown in Table 5, column 3). The next three columns remove one group of counties at a time. In column 2 we remove from the sample all counties in Para, which account for 15 percent of the weighted sample. The first-stage coefficient barely changes, and the F-statistic drops from 36.3 to 26.8. We observe similar patterns in the next two columns, where we remove from the sample counties in Amazonas (column 3), and on the remaining states (column 4). In columns 5–7, we remove combinations of two groups of counties. Removing Para and Amazonas, which together account for 58 percent of the Amazon, still leaves us with an F-statistic of 18.6 (column 5). Removing any pair of groups still yields F-stats above 10. In column 8 we remove all Amazonian counties, which account for 59 percent of the weighted sample. Perhaps unsurprisingly, our instrument becomes weak (F-statistic 6.1). To conclude, the Amazonian counties help the first-stage regression, as removing them makes the first-stage weaker. Nevertheless, we are able to exclude up to 44 percent of our sample – or 75 percent of the amazonian sample – and still get meaningful first-stage results.

We proceed by checking the sensitivity of the second-stage results to excluding Amazonian counties. Table D.4 replicates the main results, originally displayed in Tables 7 and 6 in the paper. In Panel A, we drop counties in Para and Amazonas (the corresponding first-stage regression is displayed in Table D.3, column 5). In Panel B, we keep only counties in Amazonas (the corresponding first-stage regression is in column 6 of Table D.3). In Panel C we keep only counties in Para (the corresponding first-stage regression is in column 7 of Table D.3), and in Panel D we exclude all counties in the Amazon (note that the F-Stat is only 6.1 in panel D, so the instrument is weak. We retain these results to show that the coefficient estimates remain stable).

Inspecting Table D.4, we see that both the IV estimates are quantitatively close to the original estimates. Most specifications remain significant, with the exception of those excluding all amazonian counties when the instrument has become weak. Comparing results across samples should be done cautiously, because the complier subpopulation may not be evenly distributed in across the territory. By changing the complier subpopulation, our LATE estimates will naturally change, even if the variation used to identify the effects is as good as random. Still, it is reassuring to see that our conclusions do not change even when we exclude the entire amazonian sample.

Table D.3: First-Stage Results: Sensitivity to Exclusion of Amazonian States

	Dependent Variable: Electricity Infrastructure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Instrument	0.385***	0.346***	0.359***	0.315***	0.291***	0.255***	0.280***	0.118**
	[0.064]	[0.067]	[0.065]	[0.063]	[0.067]	[0.058]	[0.068]	[0.048]
Amazonian States Included in the Sample								
Para	yes	no	yes	yes	no	no	yes	no
Amazonas	yes	yes	no	yes	no	yes	no	no
Others	yes	yes	yes	no	yes	no	no	no
Grid Points	32,578	27,711	26,407	24,476	21,540	19,609	18,305	13,438
Counties	2,172	2,136	2,145	2,105	2,109	2,069	2,078	2,042
County-years	10,860	10,680	10,725	10,525	10,545	10,345	10,390	10,210
Partial F-stat	36.3	26.8	30.4	24.8	18.6	19.4	16.9	6.1

Notes: This table address the sensitivity to the exclusion of amazonian counties from our sample. To perform this exercise, we group amazonian counties in three groups. Column 1 includes all counties, which reproduces the results of table 5, column 2 in the text. Columns 2–4 remove one group of counties at a time. Columns 5–7 remove two groups of counties at a time. Finally, column 8 removes all amazonian counties in the sample. Although one observation in the estimating sample is a county-year, all specifications are weighted by the number of grid points contained in the county. All specifications include county fixed effects, time fixed-effects, and control for a quartic polynomial on the suitability index. Standard errors clustered at the county level in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.4: Sensitivity to Exclusion of Amazonian Counties

		Land Use results							Production results					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Sample excludes Amazonian counties in Para and Amazonas states (F-statistic 18.6)														
IV	−0.327**	0.178**	0.262**	−0.624***	0.885***	−0.756**	−0.352**	−0.363	0.565**	0.381	−0.009	1.712**	−1.155**	−0.358
	[0.161]	[0.084]	[0.123]	[0.169]	[0.324]	[0.294]	[0.143]	[0.350]	[0.243]	[0.244]	[0.194]	[0.874]	[0.463]	[0.269]
Panel B: sample excludes all Amazonian counties except those in Amazonas State (F-statistic 19.4)														
IV	−0.307**	0.261***	0.655***	−0.153*	1.620***	−0.221	−0.224**	−0.363	0.565**	0.381	−0.009	1.712**	−1.155**	−0.358
	[0.146]	[0.069]	[0.223]	[0.084]	[0.429]	[0.366]	[0.113]	[0.350]	[0.243]	[0.244]	[0.194]	[0.874]	[0.463]	[0.269]
Panel C: Sample excludes all Amazonian counties except those in Para State (F-statistic 16.9)														
IV	−0.574***	0.266***	0.344***	−0.274***	1.297***	−0.246	−0.296**	−0.363	0.565**	0.381	−0.009	1.712**	−1.155**	−0.358
	[0.177]	[0.067]	[0.133]	[0.101]	[0.307]	[0.283]	[0.118]	[0.350]	[0.243]	[0.244]	[0.194]	[0.874]	[0.463]	[0.269]
Panel D: Sample Excludes all Amazonian Counties (F-statistic 6.1)														
IV	−0.363	0.565**	0.381	−0.009	1.712**	−1.155**	−0.358	−0.363	0.565**	0.381	−0.009	1.712**	−1.155**	−0.358
	[0.350]	[0.243]	[0.244]	[0.194]	[0.874]	[0.463]	[0.269]	[0.350]	[0.243]	[0.244]	[0.194]	[0.874]	[0.463]	[0.269]

Each panel removes groups of amazonian counties from the sample. Panel A excludes amazonian counties in Para and Amazonas states. The corresponding first-stages are displayed in Table D.3. All regressions are weighted by the number of grid points in the county, and include county fixed effects, time fixed effects, and directly control for a quartic polynomial on the suitability index. Standard errors clustered at the county level in brackets.

Table D.5: Sensitivity to Exclusion of Center-West

	Land Use results				Production results		
	(1) $\frac{\text{Pasture}}{\text{Farmland}}$	(2) $\frac{\text{Cropland}}{\text{Farmland}}$	(3) $\frac{\text{Native Vegetation}}{\text{Farmland}}$	(4) $\frac{\text{Farmland}}{\text{County Area}}$	(5) Grain Yields (log)	(6) Heads of Cattle Per Hectare of Pastureland	(7) Cattle Production Value/Total Production Value
IV	-0.514^{***}	0.244^{***}	0.582^{***}	-0.303^{***}	1.383^{***}	-0.642^{**}	-0.349^{***}
	[0.179]	[0.055]	[0.178]	[0.101]	[0.311]	[0.327]	[0.126]
Observations	10,040	10,040	10,040	10,040	9,886	9,835	10,040
Mean dep. var.	0.45	0.11	0.18	0.70	0.04	1.09	0.27

Sample excludes counties from the Center-West region. The first-stage coefficient is 0.33 (clustered standard error 0.068), and partial F-statistic 24.08. All regressions are weighted by the number of grid points in the county, and include county fixed effects, time fixed effects, and directly control for a quartic polynomial on the suitability index. Standard errors clustered at the county level in brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.3 Exclusion of the Center West

In table 4, we perform an exercise where we check, region-by-region, decade-by-decade, the rank correlation between hydropower suitability and agricultural outcomes. In that exercise, we found that those correlations were higher for the Center-West region, which could be a reason for concern. In table D.5 we exclude the Center-West region from the sample and redo the analysis for the main outcomes. We find that the results are qualitatively not affected by this exercise.

E Crop Productivity

Our main specifications provide us with an estimate of the impact of electrification on land use. We show that electrification impacts land use through increased productivity and switching from cattle herding toward crop production. We may instead be interested in the direct impact of crop productivity on agricultural land use and native vegetation. We can estimate this relationship by substituting agricultural productivity (log grain yields per hectare) for our endogenous electricity measure, and instrumenting for agricultural productivity using our engineering modeled electricity access variable. We use the following system of equations similar to our main specification in equations 6 and 7, estimated using Two Stage Least Squares, with $A_{c,t}$ as log grain production per hectare (productivity) within a county, decade, and $Z_{c,t}$ is our modeled

Table E.1: First-Stage Results

Dependent Variable	Log Grain Yields	
	(1)	(2)
Instrument	0.571***	0.359***
	[0.098]	[0.066]
Year dummies	Yes	Yes
Quartic suitability rank $\times$ year dummies	No	Yes
Observations	10,703	10,703
Number of Counties	2,171	2,171
Mean of Outcome	0.066	0.066
Partial F-stat	33.8	30.0

Notes: The dependent variable is the log of grain yields per hectare in a county. Column 2 adds controls that soak up variation from our instrument–a quartic polynomial of the suitability index interacted with year dummies. We keep the specification in column 2 which matches column 2 from table 6 (our preferred specification throughout the paper). All specifications include county fixed effects and are weighted by the number of grid points contained in the county. Standard errors clustered at the county level in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

electricity instrument, and $Y_{c,t+1}$ is land use in the next decade:

$$A_{ct} = \alpha_c^1 + \gamma_t^1 + \delta Z_{c,t} + \theta^1 X_{c,t} + \eta_{c,t} \tag{15}$$

$$Y_{c,t+1} = \alpha_c^2 + \gamma_t^2 + \beta \widehat{A}_{c,t} + \theta^2 X_{c,t} + \varepsilon_{c,t} \quad , \tag{16}$$

$\widehat{A}_{c,t}$ are the fitted values for productivity from the first stage regression (15), and by Z_{ct} is our modeled electricity instrument. Note that both $Z_{c,t}$ and $E_{c,t}$ are county-level averages of grid point values. As in the main specifications, we therefore weight regressions by the county's number of grid points and cluster standard errors by county. As in our preferred specification, we include interactions between the hydropower suitability index and time fixed effects in both estimating equations, which flexibly control for geography-specific trends.

The first stage in Table E1 shows that the modeled electricity instrument remains a strong instrument for productivity with an F-statistic of 30. The second stage results in Table E2 show that a 10% increase in grain yields leads to a 4 percentage point reduction in pasture land as a portion of farmland. Similarly, a 10% increase in productivity increases cropland as a portion of farmland by 1.6 percentage points and native vegetation by 3.5 percentage points. Farmland decreases as a proportion of county area by 5.1 percentage points.

Table E.2: The Effects of Agricultural Productivity on Land Use

Dependent Variable	(1) Pastures / County Area	(2) Cropland / Farmland	(3) Native Vegetation / Farmland	(4) Farmland / County Area
log Grain Production	-0.407^{***}	0.155^{***}	0.346^{***}	-0.511^{**}
	[0.124]	[0.037]	[0.097]	[0.228]
OLS	-0.045^{***}	0.016^{***}	0.039^{**}	0.004
	[0.011]	[0.004]	[0.020]	[0.010]
Observations	10,703	10,703	10,703	10,703
Mean dep. var.	0.47	0.11	0.18	0.71

Notes: The table shows how land use responds to a productivity shock (changes in log grain production per hectare). We use the engineering modeled electricity variable as the instrument for productivity. The intensive margin–land use within farms–is analyzed in columns 1-3. The dependent variable in column 1 is the farm area in pastures divided by the county's total farm area. The dependent variable in column 2 is the grains harvested area divided by the total farm area. Grains include maize, soybeans, cotton, wheat, beans, and rice. The dependent variable in column 3 is farm area in native vegetation divided by the total farm area. Column 4 analyzes the extensive margin by using county's farm area divided by the county's total area as the dependent variable. Standard errors clustered at the county level in brackets.
All regressions are weighted by the number of grid points in the county, and include county fixed effects, year fixed effects, and directly control for a quartic polynomial on the suitability index (see Table 5)
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$