Geography, slavery, and income in Brazilian municipalities in the 1870s: a spatial equilibrium approach

Guilherme Lambais and Eustáquio Reis *

September 27, 2019

Abstract

This paper applies a spatial equilibrium approach to understanding Brazil's initial heterogeneity of prices, working population, and income across municipalities in the late XIX century. Understanding the spatial equilibrium in the past is important to generate new insights on the development path taken subsequently into the XX century that would show a very unequal spatial development and a high degree of economic backwardness in an otherwise natural resource abundant country with a high land-labor ratio. Preliminary results show that the initial spatial equilibrium when shifted by exogenous geographic variables, such as terrain ruggedness, climate and soil suitability for commodities and staple foods generates perverse results for productivity and welfare that sometimes gets amplified by the factor share of slavery in a municipality.

1 Introduction

What determines the large degree of regional inequalities and economic retardation in late XIX century Brazil? Some argue that Brazil has a geography problem. Others argue that initial factor endowments, slavery, and colonial institutions are to blame. Others go even further, arguing, more in general, that a commodity-based economy with external dependency are the culprits. This paper argues that it is unlikely that one of these factors alone are the sole cause of regional inequality and economic backwardness, but that they all interact in spatial equilibrium to explain heterogeneity in prices, income and population levels across municipalities.

To understand Brazil's place in the "wealth of nations" it is first necessary to understand the "wealth of cities", as put by Glaeser and Gottlieb 2009. While a pure macroeconomic view tells us about business cycles, aggregate income and growth levels across the world, taking a look at how wealth is generated in cities will tell us about the large differences in income and population

*U. Brasília (gbrlambais@gmail.com) and IPEA (e.reis@ipea.gov.br). Work in progress. We thank Ed Glaeser for very helpful comments and suggestions, Márcia Pimentel for research assistance and Faperj/CNPq (Proc. E52 168.171/2006/Pronex) for financial assistance.

density observed within countries[1]. Brazil is not the only place where these large differences are observed, but it was unexpected since it has land abundance and high land-labor ratios and could have perfectly followed a similar path with respect to productivity and distribution of income as the United State, a pertinent observation made in Leff 1972b.

The secular roots of spatial inequality is explored in detail in Reis 2014 and, moreover, inequality in general is a paramount theme in Brazil's economic history (see, for example, Cano 1977, Leff 1972a, Buescu 1979, Bértola et al. 2009, Bértola et al. 2010, Monasterio 2010, and Zamberlan Pereira 2019). Some authors such as Summerhill 2005 attribute some source of regional inequalities to high transportation costs, an issue also discussed in Leff 1972b, and, consequently, to a geography problem. Other current of literature, arising from seminal papers such as Sachs and Warner 1995 blame natural resource abundance and geography in general, including climate. On the other hand, other current of literature that arose from seminal contributions such as Engerman and Sokoloff 1994 attribute it to initial factor endowments, which likewise is connected to geography but only in relation to soil suitability for sugar and other commodities that has economics of production in the use of slaves and, consequently, to the colonial institutions put in place to explore these commodities. Colonial institutions play a prime role in the literature derived from Acemoglu, Johnson, and Robinson 2001 to explain why some countries are less developed than others. Slavery, specifically, plays the main role in the view of Nunn 2008. Finally, the structuralists account much of the blame on the country's export orientation, where commodities production would constitute export-oriented enclaves in the territory that would dictate the development path of the economy (Furtado 1968). Leff 1972b argues that the problem is not commodities for export in itself, but that the rate in which export growth materialized between the different cultures of the Southeast (coffee) and the Northeast (sugar and cotton) caused by a shift in the country's comparative advantage (along with redirection of capital but with imperfect labor mobility) that generated the regional differentials in economic development.

In contributing to this literature, we add one more explanation for the puzzle of why a country with such abundance of natural resources and high land-labor ratios would generate such unequal development path could rest on the configuration of the initial spatial equilibrium across municipalities. Combining a novel source of income data in 1876 from archival research with the first nation-wide census in 1872 and GIS-constructed data on geography, climate, and soil suitability for commodities and staple foods we are able to shed some light on this topic.

We show that ruggedness, temperature, and soil suitability are correlated with income, workforce, and our proxy for land and housing prices. But that, ruggedness, for example, which can be considered "bad geography" is actually correlated with higher levels of productivity but only in places with lower levels of factor share of slavery and that it depresses welfare[2] in general. Further-

[1]This literature deals in terms of cities, which we use interchangeably with municipalities that is the official level of local aggregation in Brazil. The difference is that the municipality refers to the legal territory in its entirety and not just the urban center. This distinction might be more appropriate considering the urban dynamics of the XIX century and also ensure comparability with the bulk of the literature on historical Brazil that uses the municipality as the locus of political and economic activity (see, for example, Leal 1977).

[2]In this version of the paper we consider welfare to be interchangeable with the income of the consumers. Future

more, we show that people do seem to pay a welfare premium on higher average temperatures, which is strange considering that the country is warm in general, and that this higher temperatures also take a toll on productivity. In addition, we show that soil suitability for a commodity, sugarcane in the case, depresses welfare in the form of higher prices, but that is correlated with higher levels of productivity, which decreases as the factor share of slaves increases. Finally, we show that soil suitability for staple foods increases welfare but that it ceases to depresses productivity only with high levels of slavery factor share.

This paper is organized as follows, in section 2 we provide a brief historical background of Brazil in the XIX century and, more specifically, of the 1870s decade, in section 3 we outline our theoretical approach, in section 4 we detail the data sources and how we construct our variables and estimating equations to apply the spatial equilibrium framework and also show a direct test of the model, section 5 details the marginal impacts of our variables on municipality-specific welfare and productivity and what is the role of slavery and section 6 concludes the paper.

2 Historical background

Brazil in the 1870s was in the midst of a structural transformation. The gold cycle was long gone and the cycles of sugar and cotton in Northeast were rapidly giving place to the coffee cycle in the Southeast[3]. The recent Paraguay War (1864-1870), involving Paraguay against Brazil, Uruguay and Argentina, and the American Civil War (1861-1865) would intensify abolitionist ideals, and the origins of mass immigration to Brazil was in the making to substitute slave labor. According to the 1872 Census, Brazil had almost 10 million people, of which 1.5 million were slaves, but already 2.5 million were free men of working age (16-60 years old). In the ten years between 1870 an 1879, an average of 20,780 immigrants would enter the workforce annually, a number that would only grow; in the last decade of the XIX century an average of 118,170 immigrants would enter the country annually (Leff 1972b).[4]

In 1871 the "Lei do Ventre Livre" would declare that all children born to female slaves would be born free. During this decade the number of voluntary manumissions would increase. Between 1873 and 1887, the year before the abolition, the number of slaves would decrease from 1.5 million to 723 thousand and the concentration of slaves in the coffee provinces would increase from 57% to 67% of all slaves. Beyond the 1871 law, during the first half of the 1870s the country would see reforms in the police, the judiciary, the national guard, the first nation-wide census, the connection of Brazil to Europe through Lisbon by telegraph, the adoption of the metric system, and construction of railroads (Carvalho 2012). In 1874 the country had almost 1300 kilometers of railroad (800 miles),

versions should improve this definition.

[3]Coffee would grow from 18.4% of the total exports considering the eight main commodities (coffee, sugar, cocoa, mate herb, tobacco, cotton, rugger, and leather) in the decade of 1821-1830 to 56.6% in the 1871-1880 decade. In contrast, sugar would fall from 30.1% in the 1821-1830 decade to 11.8% in the 1871-1880 decade and cotton from 20.6% to 9.5% considering the same periods (Carvalho 2012).

[4]To put these numbers in perspective, the population in the United States in 1870 was of almost 40 million people, having received in the 1870s 3 million immigrants, against 527 thousand that went to Brazil (Carvalho 2012)

a figure that would grow to 2100 kilometers (1300 miles) in 1876 and 6200 kilometers (3900 miles) in 1884 (Leff 1972b).[5]

This profusion of new ideas, new people, and technological innovations makes the 1870s a good starting place to analyze the initial spatial equilibrium across municipalities in relation to prices, income, and working population, taking into account also the degree of slavery, that is, the decision of consumer and producers in deciding where to locate across the territory. It would be interesting, of course, to analyze previous spatial equilibria, but the lack of previous nation-wide census makes this task almost impossible, or at least incomplete.

3 The spatial equilibrium framework

We frame the empirical exercise employing a basic model of spatial equilibrium across cities, in what is known as the Rosen-Roback framework (Rosen 1979, Roback 1982), following the exposition in Glaeser 2008 and Harari 2018. In this most basic version of the model we use cities and municipalities interchangeably, but in the next iterations of the paper we would like to bring the model closer to the reality of the municipality in the late XIX century, that is, to better model the role of slavery and agricultural production. For now this framework features consumers, or the free population, and production and construction markets. Consumer households choose optimally in which municipality to live. The spatial equilibrium condition is given at the point where consumers are indifferent across cities with different amenities, that is, the indirect utility value of a location choice must equal a reservation utility level. Production and construction of housing is done competitively in each municipality over a fixed supply of land. The production sector considers production in general, making no distinction between agriculture and industry, and slaves are regarded as a form of tradable capital.

Consumer households have a Cobb-Douglas[6] utility function $\theta C^{1-\alpha} H^{\alpha}$ defined over tradable goods, the numéraire good C, non-traded housing H, and municipality-specific amenities, captured by an index θ. The supply of labor is inelastic, for which they receive a municipality-specific wage W. Optimizing behavior gives the following indirect utility

$$log(W) - \alpha log(p_H) + log(\theta) = log(\bar{\vartheta}) \tag{1}$$

where p_H is the rental price of housing. This shows that the reservation utility ($\bar{\vartheta}$) is equalized across cities, otherwise households could move around to exploit differences in utility. This shows the key intuition behind the spatial equilibrium model: that consumers, in equilibrium, implicitly pay for amenities θ by having lower wages (W) or higher housing prices (p_H) in a municipality. Note that θ could include amenities and disamenities. Thus empirically looking at these compensating differentials reveals the value placed by households on certain amenities (or disamenities).

[5] By comparison, the US had 85 thousand kilometers in 1870 and 135 thousand kilometers in 1880 (Carvalho 2012).

[6] The intuition of the Rosen-Roback framework does not require functional forms assumptions, but in bringing the model to the data it is usual to turn to particular functional forms, where the Cobb-Douglas form is a fairly standard assumption to derive a set of estimable equations.

Turning to the production sector, firms also choose where to locate and competitively produce a good Y, using free labor N, traded numéraire slave capital K and a fixed supply of non-traded capital $\bar{Z}$. Assume that every municipality is characterized by a location-specific productivity level A, the production function is then given by $AN^{\beta}K^{\gamma}\bar{Z}^{1-\beta-\gamma}$. The zero-profit condition for the firms yields the following demand curve

$$(1-\gamma)log(W) = (1-\beta-\gamma)(log(\bar{Z} - log(N)) + log(A) + \kappa_1. \tag{2}$$

To close the model, housing is also produced competitively in a combination of height (h) and land (L). Thus the total quantity of housing supplied (H) is equal to hL. There is a fixed quantity of land available in each municipality, denoted by $\bar{L}$. In the short run this variable can be taken as given, either by regulators or, as we develop here, by geography. Using a measure of quantity of land available to be developed given by geography constraints has the advantage of being totally exogenous in the short run, as opposed to available land given by regulators, which can be thought of as an endogenous process. This fixed supply $\bar{L}$, in turn, determines an endogenous price for land (p_L) and housing (p_H).

The cost of producing H unites of housing is $c_0 h^{\delta} L - p_L L$ where $\delta > 1$, thus the profit function is $p_H H - c_0 h^{\delta} L - p_L L$. Free entry of construction gives us the equilibrium condition of housing prices as a function of population and income

$$log(p_H) = \frac{1}{\delta}log(\delta c_0) + \frac{\delta - 1}{\delta}(log(\alpha N W) - log(\bar{L})). \tag{3}$$

Now using the three optimality conditions 1, 2, 3, we have three equations with three unknowns, namely population, income and housing prices. Solving for these endogenous variables as functions of municipality-specific productivity A, amenities θ, and fixed supply of land $\bar{L}$, we have the following system of equations

$$log(N) = F_N log(A) + E_N log(\theta) + D_N log(\bar{L}) + I_N \tag{4}$$

$$log(W) = F_W log(A) + E_W log(\theta) + D_W log(\bar{L}) + I_W \tag{5}$$

$$log(p_H) = F_p log(A) + E_p log(\theta) + D_p log(\bar{L}) + I_p \tag{6}$$

where E, F, G, and I denote constant functions of the model's deep parameters and $F_N, F_W, F_p > 0$; $E_N, E_p > 0$; and $E_W < 0$. This shows another key intuition of the Rosen-Roback spatial equilibrium framework: that it is impossible to analyze income, population or prices separately or alone because there is a set of markets in which they all interact. "Population, wages, and rents are all increasing functions of the municipality-specific productivity parameter A. Intuitively, higher A allows firms to pay higher wages, which attracts households and bids up rents. Similarly, population and rents are increasing in the amenity parameter θ: better amenities attract households and bid up rents. Wages are decreasing in θ because firms prefer cities with higher production amenities, whereas consumers prefer cities with higher consumption amenities, and factor prices - W and p_H - strike the balance between these conflicting location preferences." (Harari 2018)

Now consider a geography-based exogenous shifter of income, population, and housing prices. Assume that for this geographic exogenous shifter, denoted G,

$$log(A) = I_A + \lambda_A G + \mu_A \tag{7}$$

$$log(\theta) = I_\theta + \lambda_\theta G + \mu_\theta \tag{8}$$

$$log(\bar{L}) = I_L + \lambda_L G + \mu_L \tag{9}$$

where I_i are constants, λ_i are coefficients, and μ_i are error terms $\forall i \in \{A, \theta, L\}$. Since A and θ are unobservables, we substitute these back into equations 4, 5, and 6, which then imply

$$log(N) = \kappa_N + \frac{(\alpha + \gamma - \alpha\gamma)\lambda_A + (1-\gamma)(\delta\lambda_\theta + \alpha(\delta-1)\lambda_L)}{\delta(1-\beta-\gamma) + \alpha\beta(\delta-1)} G + \mu_N \tag{10}$$

$$log(W) = \kappa_W + \frac{(\delta-1)\alpha\lambda_A - (1-\beta-\gamma)(\delta\lambda_\theta + \alpha(\delta-1)\lambda_L)}{\delta(1-\beta-\gamma) + \alpha\beta(\delta-1)} G + \mu_W \tag{11}$$

$$log(p_H) = \kappa_p + \frac{(\delta-1)(\lambda_A + \beta\lambda_\theta - (1-\beta-\gamma)\lambda_L)}{\delta(1-\beta-\gamma) + \alpha\beta(\delta-1)} G + \mu_p \tag{12}$$

where κ_i are the constants independent of G and μ_i error terms independent of G, $\forall i \in \{N, W, p\}$.

Now we could use some geography-based shifter G that is connected with prices, population, and income, given plausible values for α, β, γ, and δ, to provide estimates of λ_A, the marginal productivity impact of G, λ_θ, the marginal willingness to pay for G, and λ_L, the welfare impact of G. Denote by $\hat{B}_N$, $\hat{B}_W$, and $\hat{B}_p$ the estimated reduced-form coefficients of the geography variable G impact on the population, price, and income regressions.

Totally differentiating the indirect utility of consumers, equation 1, with respect to G yields

$$\frac{\partial log(\theta)}{\partial G} = \alpha \frac{\partial log(p_H)}{\partial G} - \frac{\partial log(W)}{\partial G} \tag{13}$$

thus λ_θ can be estimated as

$$\hat{\lambda}_\theta = \alpha \hat{B}_p - \hat{B}_W. \tag{14}$$

Totally differentiating the zero-profit condition of the production sector, equation 2, with respect to G yields

$$\frac{\partial(A)}{\partial G} = (1-\beta-\gamma)\frac{\partial log(N)}{\partial G} + (1-\gamma)\frac{\partial log(W)}{\partial G} \tag{15}$$

thus λ_A can be as estimated as

$$\hat{\lambda}_A = (1-\beta-\gamma)\hat{B}_N + (1-\gamma)\hat{B}_W. \tag{16}$$

Finally, totally differentiating the zero-profit condition of the construction sector, equation 3, with respect to G yields

$$\frac{\partial log(\bar{L})}{\partial G} = \frac{\partial log(N)}{\partial G} + \frac{\partial log(W)}{\partial G} - \frac{\delta}{\delta - 1}\frac{\partial log(p_H)}{\partial G} \tag{17}$$

which suggests that the estimation of λ_L is connected to the other variables in the following manner

$$\hat{\lambda}_L = \hat{B}_N + \hat{B}_W - \frac{\delta}{\delta - 1}\hat{B}_p. \tag{18}$$

Since θ and A is unobserved, we can back out $\hat{\lambda}_\theta$ and $\hat{\lambda}_A$ by using $\hat{B}_N$, $\hat{B}_W$, and $\hat{B}_p$. In our application, $log(p_H)$ is also unobserved, thus we can directly estimate λ_L since the variable $\hat{L}$ is constructed based on geography and we can then back out $\hat{B}_p$ using $\hat{\lambda}_L$, $\hat{B}_N$, and $\hat{B}_W$. The next sections implement the model.

4 Empirical strategy

4.1 Data sources and variable construction

Working with data referent to the XIX century in Brazil has certain limitations. The first nation-wide representative census was only in 1872 and did not contain any mention of income or wages, being mostly about demography and occupations. The demographic data on working population and slaves is thus from the 1872 Census (Brasil 1876), in a version worked out by Cedeplar. Data on incomes is collected from the wages of municipal civil servants in 1876 published by the Statistical Report of the Empire in 1878 (Brasil 1878).

The data on terrain ruggedness is constructed using GIS and the SRTM digital elevation data produced by NASA and revised and distributed by CGIAR (Jarvis et al. 2008). Specifically, we use the DEM with 250m resolution at the equator. Climate data is calculated using the Bioclim dataset from WorldClim (Fick and Hijmans 2017) with a 1km resolution at the equator and soil suitability is calculated using the FAO GAEZ database (IIASA/FAO 2012) and has a resolution of 10km at the equator.

The construction of the variable for $\bar{L}$, the short-run exogenous amount of land available for development, is done using GIS with the following procedure. First, 10km[7] geodesic buffers are created around the coordinates for the municipality seat according to IBGE (note that this is not the centroid of the municipality shape, but the actual municipality downtown center), then all water bodies, made available in shapefiles by IBGE, are removed. Then these buffers are intersected with municipality shapes so the buffer only covers the actual municipality territory, avoiding its extension into neighboring municipalities. Next a slope raster is calculated using the STRM digital elevation model. Finally, the total area within the buffer that is not covered by water bodies or does not have slope greater than 15% can be calculated. This is illustrated in Figure 1 for the region of the Empire

[7] A radius of 10km is a realistic distance for the time where most of transportation was done by foot or horses and mules.

capital. Figure 2 shows the whole country. Thus we have data for W, N and $\bar{L}$, which leads us to the following strategy.

4.2 One direct test of the model

One fact that could support the relevance of the spatial equilibrium approach in this setting is to see if there is correlation between area prices and area income. Since we still do not have data for local prices at this time in Brazil, we use our constructed variable of land available for development as proxy for area prices. The reasoning is simple: places with less area available for development tend to have higher housing and land prices because of the short supply in relation to places with less constraints.

Figure 3 displays the correlation between the variables. There is no sense of what is causing what because the model treat both variables as endogenous, still the link is valid to see the fit of the model. The correlation shown is

$$log(Income) = 13.24 - 0.23 \text{ x } log(\bar{L})$$

with a r-squared of 0.011 and a robust standard error of 0.101. This relationship appears to support the assumption that high wages are compensated by higher land prices. Glaeser 2008 shows based on current surveys in the US that the average family spends about 30 percent of its income on housing, which would equal a coefficient of 0.3 in the regression. Our coefficient being 0.23 shows that we could consider the housing share between 0.13 and 0.33 to be very plausible values for Brazil in the late XIX century. This also shows a remarkable well fit for the Cobb-Douglas assumption of this model.

4.3 Estimating the impact of exogenous shifters

Now we turn to properly estimating the model with exogenous geographic shifters. We estimate a version of equations 9, 10, and 11

$$log(\bar{L}) = K_L + \hat{\lambda}_L geography_i^k + \mu_L \tag{19}$$

$$log(N) = \kappa_N + \hat{B}_N geography_i^k + \mu_N \tag{20}$$

$$log(W) = \kappa_W + \hat{B}_W geography_i^k + \mu_W \tag{21}$$

where $geography_i^k$ is a different variable k for each municipality i. The models are estimated for k equal to terrain ruggedness, average annual temperature, and soil suitability for sugarcane, maize and cassava.

5 The marginal impact of geography on municipality-specific welfare and productivity and the role of slavery

5.1 The case of "bad geography"

Terrain ruggedness can be considered "bad geography", due to its detrimental effect in the long run (Nunn and Puga 2012). It has negative effects because it most likely hurts productivity due to difficulties created for market access and in transporting goods[8]. Note, however, that market access is more conditioned by distance to market centers or by ruggedness in the roads to market centers than by local ruggedness itself. Most importantly, in agrarian contexts, ruggedness creates difficulties for the cultivation of agricultural crops caused, in particular, by leaching. In urban contexts, the difficulties are related to the construction of housing but, in the ancient city, ruggedness have substantial advantages in what concerns sanitary and security conditions. There is also difficulties in commuting within the city, problems that are aggravated in peripheral countries that do not have the resources to invest in proper transportation systems.

The reduced-form results of the regressions of ruggedness on income, land available for development, and population are presented in Table 1. Using plausible values for α, β, γ, and δ, we can analyze the reduced-form results through the lens of the spatial equilibrium framework. Suppose the value of $\delta = 1.5$, which is a standard value for the elasticity of housing supply in the literature. In combination with the reduced-form estimates $\hat{\lambda}_L = -0.004973$, $\hat{B}_N = 0.005374$, and $\hat{B}_W = 0.004927$, the estimated value of the impact of ruggedness on $logp_H$ is $\hat{B}_p = 0.005091$. Assuming $\alpha = 0.2$, which we saw is a plausible value for the share of housing expenditure of consumers income, we have that $\hat{\lambda}_\theta = -0.003399$. This means that there is a negative welfare impact imposed by bad geography, a 10 point increase in ruggedness is worth the same as a 0.034 log point decline in income. This is not insignificant as the ruggedness scale in Brazil goes from around 2 (level terrain surface) to a maximum of 170 (intermediary rugged) and the standard deviation of log income is 0.8, thus a 30 point increase in ruggedness would be the same as if income goes down by a full standard deviation.

Now suppose that the value for the share of fixed capital in production is $(1 - \beta - \gamma) = 0.1$, then we can analyze the role of slavery in this framework. Assume that slavery is a form of tradable capital and that only free labor constitutes the labor share. Looking at the 1872 Census there are only five municipalities with less than 50 slaves, which would mean a slave share less than 0.1. Overall the mean slave share is 0.3 with a standard deviation of 0.2 and a maximum of 0.9, meaning that there are municipalities with very little slavery, at least in the census, up to municipalities where slaves constitute 90 per cent of the factor share with only 10 per cent of free labor. Thus, for the empirical exercises we can look at the variation for γ, the slave share, from 0.1 to 0.9, with 0.3 as the most plausible value.

For the slave share equal to 0.1 we have $\hat{\lambda}_A = 0.0049$, for the slave share equal to 0.3, $\hat{\lambda}_A = 0.0039$,

[8]As explored by Nunn and Puga 2012, ruggedness has positive effects on long-run development only in Africa because it facilitated for populations to hide from the slave trade.

and for the slave share equal to 0.3, $\hat{\lambda}_A = 0.00103$. For example, take the mean value for the slave share, this suggests that a 10 point increase in ruggedness is worth the same as 0.039 increase in city-specific productivity. This is an unexpected result. Of course this is not the causal impact of ruggedness on productivity, as there might be uncontrolled for factors associated with productivity ruggedness that are positively correlated with ruggedness. What this could be pointing to is that most of economic activity at this time in Brazil was concentrated along the coast line and states such as Minas Gerais which is naturally more rugged because of a very large relief difference in the Southwest region, where going 100km inland often means going up 1000m, which can be noted in Figure 2, thus ruggedness could be correlated with some economies of density that are driving the result. Other interesting fact is that the greater the free labor share in the model, the greater is the exchange between bad geography and productivity. Together this shows the contradictory nature of Brazil's initial development, where places with less slavery bad geography is connected to more productivity.

5.2 The case of climate

Other issue of debate is the role of climate in Brazil's development. Table 2 shows the regression of annual average temperature on land available for development, population of workers and income, where we see that higher average temperatures throughout the year are correlated with both lower levels of working population and income. Assuming the same values as above for housing share and the elasticity of housing supply, we have that $\hat{\lambda}_\theta = 0.0293$. This means that an increase in 10 degrees Celsius is worth the same as a 0.293 log point increase in income. Thus it appears that warm weathers offset lower wages.

Turning to productivity, the predicted value of $\hat{\lambda}_A$ lies between -0.099 and -0.0489 for the lower and higher bonds of the slave share. Taking the mean slave share, then $\hat{\lambda}_A = -0.0865$, which means that higher temperatures are depressing productivity. Again, this is surely not the causal impact, however there is some correlated factor with temperature that is driving productivity down.

5.3 The case of soil suitability

At last, Tables 3, 4, and 5 show the regression results for soil suitability for sugarcane, cassava, and maize, respectively. Sugarcane is a very important commodity in Brazil's history and cassava and maize are two staple foods widely consumed throughout the country.

First, in relation to sugarcane, we see in Table 3 that a municipality having more soil suitable for sugarcane is correlated with higher levels of income but there is no effect on working population. Again, assuming the same parameter values for the elasticity of housing supply and the housing share, the results for sugarcane show that $\hat{\lambda}_\theta = -0.0724$. This is also an unexpected result, a 10 percent increase in sugarcane suitability depresses welfare in 0.7 percent, but consumers appear to be paying in the form of higher prices ($\hat{B}_p = 0.01163$) rather than lower wages. In relation to productivity, the estimate value of $\hat{\lambda}_A$ varies between 0.0673 for the lower bound of the slave share to 0.00748 to the higher bound. This shows that sugarcane suitability is positively correlated with productivity and that this correlation decreases with higher levels of the slave share.

Second, in relation to the staple foods, we see in Tables 4 and 5 that both staples are positively correlated with working population and negatively correlated with income. Since the results for both cassava and maize run in the same direction it suffices to further discuss only one of them. Take, for instance, cassava, the estimated value for $\hat{\lambda}_\theta$ is 0.0655. This means that a 10 percent increase in the soil suitability for cassava is worth a 6.5 percent increase in welfare. However, this increase in welfare does not seem to come from higher wages, but from lower prices ($\hat{B}_p = -0.0178$). In relation to productivity, the predicted value $\hat{\lambda}_\theta$ lies within -0.055 for the lower bound and 0.00026 for the higher bound of the slave share. Looking at the mean value of the slave share, $\hat{\lambda}_\theta = -0.0412$, thus higher soil suitability for cassava ceases to depresses productivity only with high factor shares of slavery.

6 Conclusion

This paper has shown that the spatial equilibrium framework has a good fit for municipalities in late XIX century Brazil. In taking a step to understand some puzzles related to Brazil's position in the wealth of nations we first turn to understand heterogeneity in prices, income and population levels in municipalities. Combining archival resources to build income data for the late 1870s with the first census in history and GIS-derived data it is possible to generate great insight driven by theory and facts in a otherwise context of poor data that usually led only to subjective interpretations. Specifically, we have shown that the initial spatial equilibrium in Brazil worked in perverse ways that is sometimes amplified by the factor share of slavery. Future research could more exogenous variables and predict how income behaves at the municipality and more aggregate levels.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation". *The American Economic Review* 91 (5): 1369–1401.

Bértola, Luis, et al. 2010. "Between the Colonial Heritage and the First Globalization Boom: On Income Inequality in the Southern Cone". *Revista de Historia Económica / Journal of Iberian and Latin American Economic History* 28 (02): 307–341.

— . 2009. "Income Distribution in the Latin American Southern Cone during the First Globalization Boom and Beyond". *International Journal of Comparative Sociology* 50 (5-6): 452–485.

Brasil, D. G. d. E. 1876. *Recenseamento Geral do Imperio de 1872*. Rio de Janeiro: Typ. Leuzinger, Typ. Commercial.

— . 1878. *Relatorio e trabalhos estatisticos apresentados ao Illm. e Exm. Sr. Conselheiro Dr. Carlos Leoncio de Carvalho Ministro e Secretario de Estados dos Negocios do Imperio pelo Director Geral Conselheiro Manoel Francisco Correia em 20 de Novembro de 1878.* Rio de Janeiro: Typographia Nacional.

Buescu, Mircea. 1979. *Brasil: Disparidades de Renda no Passado: Subsídios Para o Estudo dos Problemas Brasileiros.* Rio de Janeiro: ANPEC.

Cano, Wilson. 1977. *Raízes da concentração industrial em São Paulo.* São Paulo: Difel.

Carvalho, José Murilo de. 2012. *História do Brasil Nação: 1808-2010. Vol. 2. A construção nacional 1830-1889.* Rio de Janeiro: Objetiva.

Engerman, Stanley L, and Kenneth L Sokoloff. 1994. *Factor Endowments: Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States.* Working Paper, Historical Working Paper Series 66. National Bureau of Economic Research.

Fick, Stephen E., and Robert J. Hijmans. 2017. "WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas". *International Journal of Climatology* 37 (12): 4302–4315.

Furtado, Celso. 1968. *Formação Econômica do Brasil.* 8ed. São Paulo: Companhia Editora Nacional.

Glaeser, Edward. 2008. *Cities, Agglomeration, and Spatial Equilibrium.* Oxford University Press.

Glaeser, Edward L., and Joshua D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States". *Journal of Economic Literature* 47 (4): 983–1028.

Harari, Mariaflavia. 2018. *Cities in Bad Shape: Urban Geometry in India.* Working Paper. University of Pennsylvania.

IIASA/FAO. 2012. *Global Agro-ecological Zones (GAEZ v3.0).* Tech. rep. IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Jarvis, A., et al. 2008. *Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org).* Tech. rep. CGIAR - Consortium for Spatial Information.

Leal, Vitor Nunes. 1977. *Coronelismo: The Municipality and Representative Government in Brazil.* Cambridge: Cambridge University Press.

Leff, Nathaniel H. 1972a. "Economic Development and Regional Inequality: Origins of the Brazilian Case". *The Quarterly Journal of Economics* 86 (2): 243.

— . 1972b. "Economic Retardation in Nineteenth-Century Brazil". *The Economic History Review* 25 (3): 489.

Monasterio, Leonardo Monteiro. 2010. "Brazilian spatial dynamics in the long term (1872–2000): "path dependency" or "reversal of fortune"?" *Journal of Geographical Systems* 12 (1): 51–67.

Nunn, Nathan. 2008. "The Long-Term Effects of Africa's Slave Trades". *Quarterly Journal of Economics* 123 (1): 139–176.

Nunn, Nathan, and Diego Puga. 2012. “Ruggedness: The Blessing of Bad Geography in Africa”. *The Review of Economics and Statistics* 94 (1): 17.

Reis, Eustáquio. 2014. “Spatial Income Inequality in Brazil, 1872–2000”. *EconomiA* 15 (2): 119–140.

Roback, Jennifer. 1982. “Wages, Rents, and the Quality of Life”. *Journal of Political Economy* 90 (6): 1257–1278.

Rosen, Sherwin. 1979. “Wage-Based Indexes of Urban Quality of Life”. *Current Issues in Urban Economics*: 74–104.

Sachs, Jeffrey D, and Andrew M Warner. 1995. *Natural Resource Abundance and Economic Growth.* Working Paper, Working Paper Series 5398. National Bureau of Economic Research.

Summerhill, William R. 2005. “Big Social Savings in a Small Laggard Economy: Railroad-Led Growth in Brazil”. *The Journal of Economic History* 65 (01).

Zamberlan Pereira, Thales Augusto. 2019. “The North-South Divide: Real Wages and Welfare in Brazil during the Early 20th Century”. *Revista de Historia Económica / Journal of Iberian and Latin American Economic History*: 1–30.

7 Tables and Figures

Figure 1: Cities of the Empire in the Rio de Janeiro region and land available for development. The dots represent the downtown center. The light green dot represents Rio de Janeiro, the Empire Capital. The dark green lines represents a maximum of 10km buffers around the downtown center which do not contain water bodies and that do not transverse into territories of other municipalities. The yellow area represent land that is above 15% of slope. The grey area represents all other land. Thus, the area available for development in the short run ($\bar{L}$) is the grey area within the green buffer.

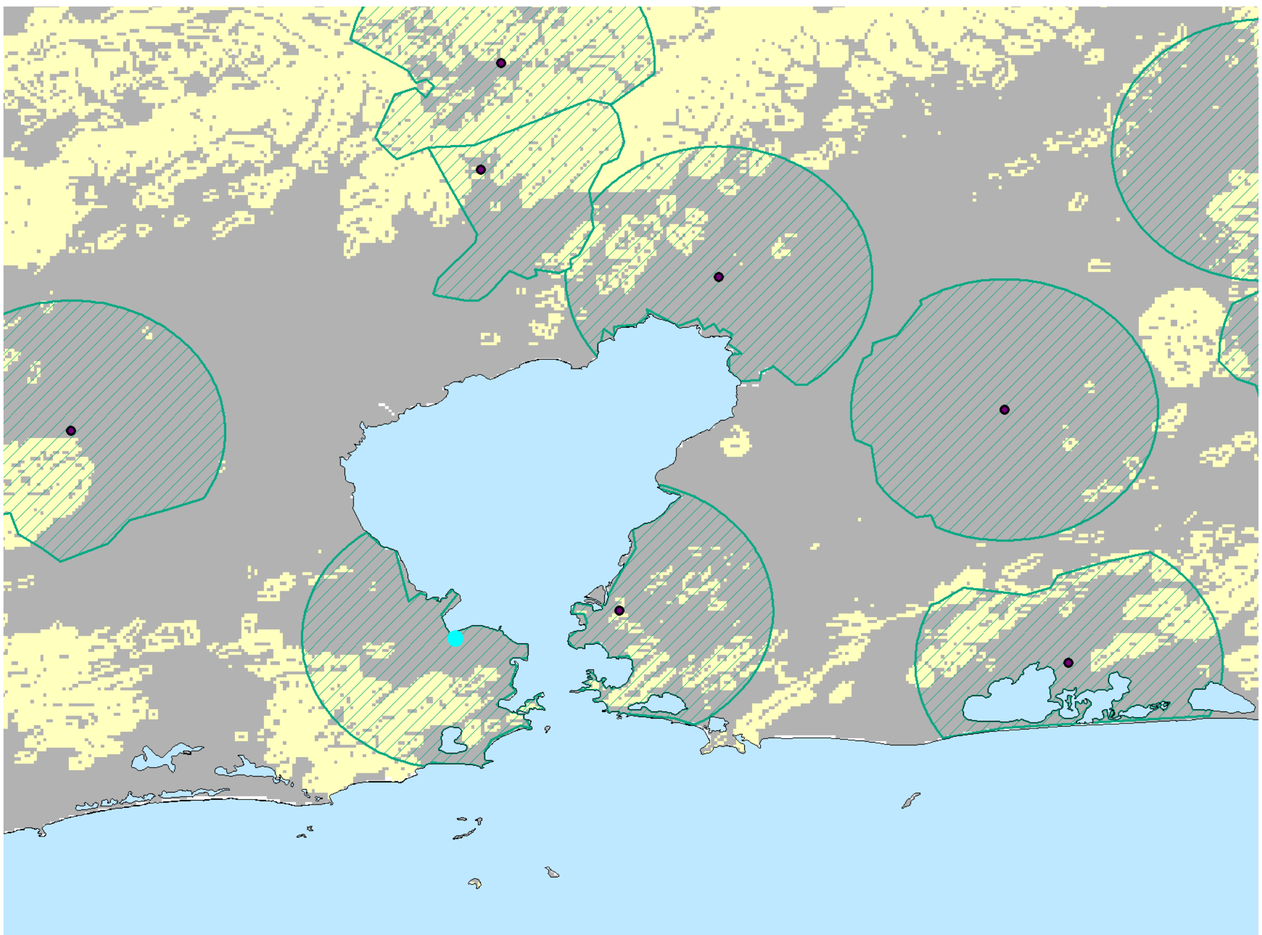

Figure 2: Buffers zones and 15% slope raster for Brazil. The dark green circles represents a maximum of 10km buffers around the downtown center which do not contain water bodies and that do not transverse into territories of other municipalities. The yellow area represent land that is above 15% of slope. The grey area represents all other land.

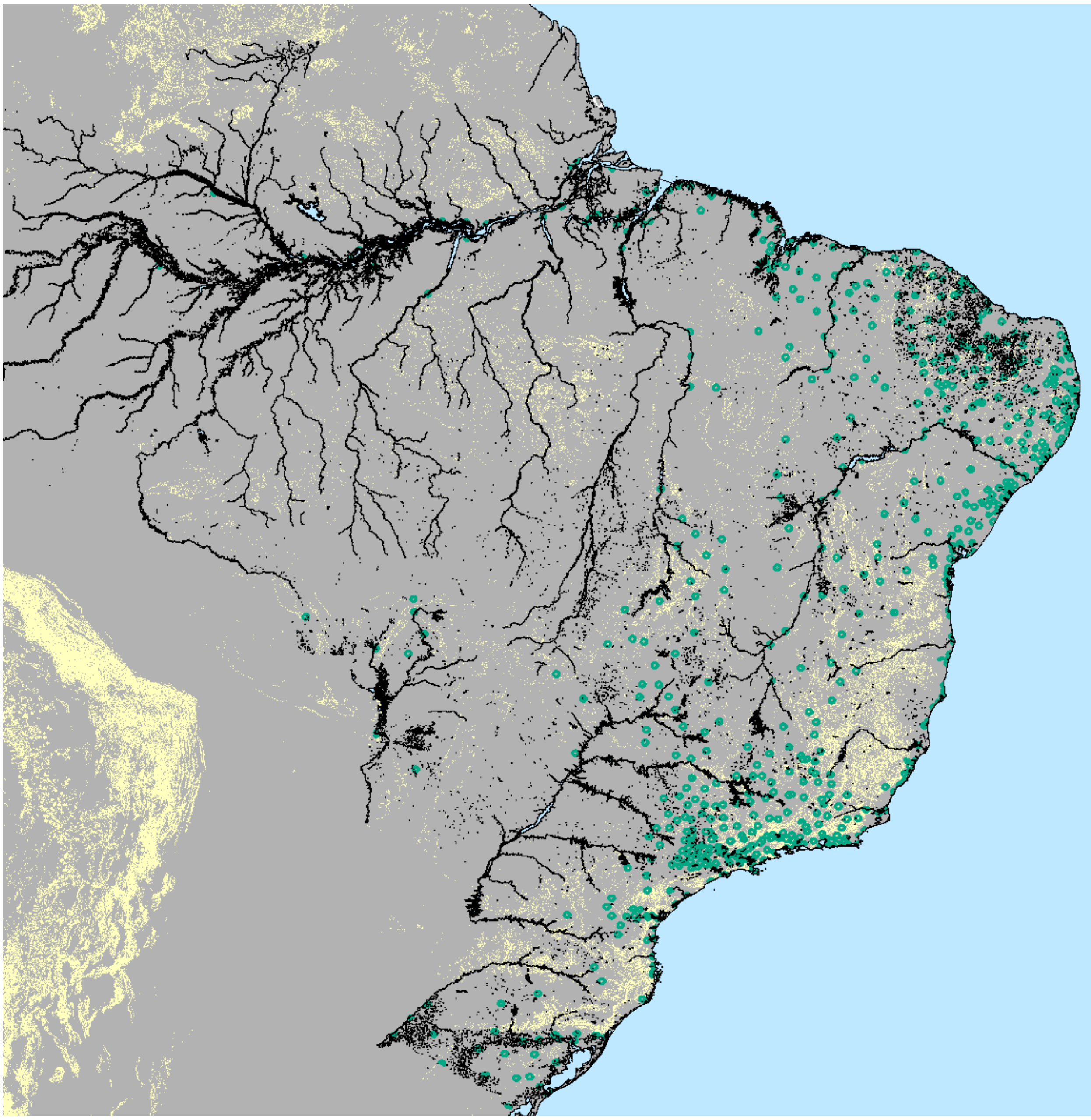

Figure 3: The correlation between log income and log land available for development.

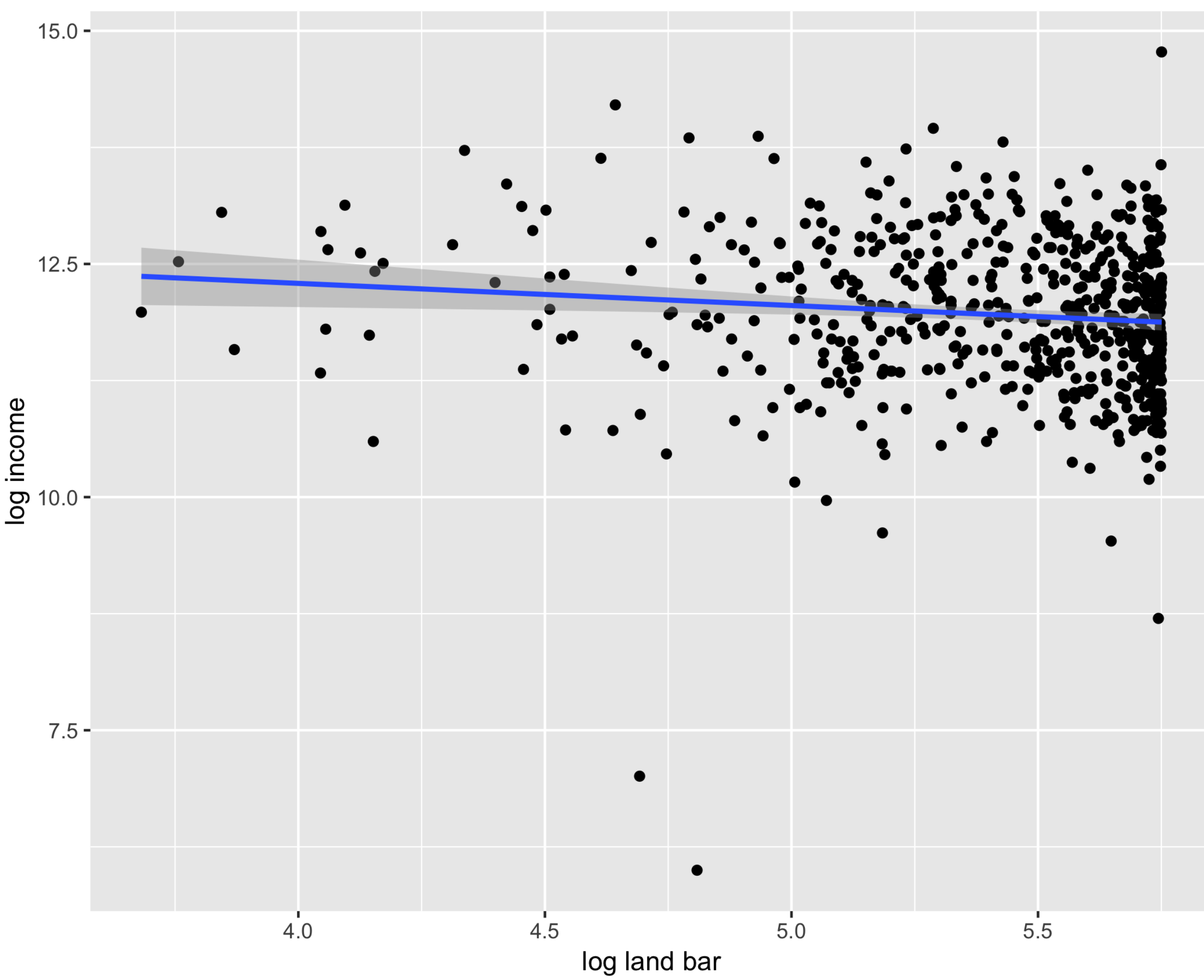

Table 1: The impact of ruggedness (the geographic exogenous shifter G) on land available for development, population of workers, and wages

	Dependent variable:		
	$log(\bar{L})$	$log(N)$	$log(W)$
	(1)	(2)	(3)
ruggedness	−0.004973***	0.005374***	0.004927***
	(0.0005)	(0.001)	(0.001)
Constant	5.638***	7.585***	11.732***
	(0.024)	(0.081)	(0.059)
Observations	641	641	619
R^2	0.160	0.020	0.032
Adjusted R^2	0.159	0.019	0.031
Residual Std. Error	0.346 (df = 639)	1.140 (df = 639)	0.826 (df = 617)
F Statistic	121.574*** (df = 1; 639)	13.085*** (df = 1; 639)	20.630*** (df = 1; 617)
Note:	Robust standard error in parenthesis. *p<0.1; **p<0.05; ***p<0.01		

Table 2: The impact of annual average temperature (the geographic exogenous shifter G) on land available for development, population of workers, and wages

	Dependent variable:		
	$log(\bar{L})$	$log(N)$	$log(W)$
	(1)	(2)	(3)
annual avg temperature	0.011357**	−0.042634***	−0.062721***
	(0.005339)	(0.010370)	(0.011299)
Constant	5.153812***	8.870669***	13.385320***
	(0.126037)	(0.239453)	(0.262699)
Observations	635	635	614
R^2	0.007753	0.020783	0.049136
Adjusted R^2	0.006185	0.019236	0.047582
Residual Std. Error	0.377315 (df = 633)	0.859434 (df = 633)	0.818632 (df = 612)
F Statistic	4.945740** (df = 1; 633)	13.434810*** (df = 1; 633)	31.625060*** (df = 1; 612)
Note:	Robust standard error in parenthesis. *p<0.1; **p<0.05; ***p<0.01		

Table 3: The impact of sugarcane suitability (the geographic exogenous shifter G) on land available for development, population of workers, and wages

	Dependent variable:		
	$log(\bar{L})$	$log(N)$	$log(W)$
	(1)	(2)	(3)
log(sugarcane suitability)	0.039978**	0.026945	0.074869***
	(0.016233)	(0.022103)	(0.025776)
Constant	5.109590***	7.695602***	11.390940***
	(0.127208)	(0.161270)	(0.195399)
Observations	635	635	614
R^2	0.011051	0.000955	0.008094
Adjusted R^2	0.009489	−0.000623	0.006474
Residual Std. Error	0.376688 (df = 633)	0.868092 (df = 633)	0.836113 (df = 612)
F Statistic	7.073367*** (df = 1; 633)	0.605056 (df = 1; 633)	4.994156** (df = 1; 612)
Note:	Robust standard error in parenthesis. *p<0.1; **p<0.05; ***p<0.01		

Table 4: The impact of cassava suitability (the geographic exogenous shifter G) on land available for development, population of workers, and wages

	Dependent variable:		
	$log(\bar{L})$	$log(N)$	$log(W)$
	(1)	(2)	(3)
log(cassava suitability)	0.056160***	0.071774***	−0.069195***
	(0.018696)	(0.020218)	(0.021615)
Constant	4.972079***	7.336931***	12.498510***
	(0.151835)	(0.158293)	(0.174011)
Observations	635	635	614
R^2	0.038195	0.011867	0.012118
Adjusted R^2	0.036676	0.010306	0.010504
Residual Std. Error	0.371482 (df = 633)	0.863338 (df = 633)	0.834415 (df = 612)
F Statistic	25.137850*** (df = 1; 633)	7.601861*** (df = 1; 633)	7.507380*** (df = 1; 612)
Note:	Robust standard error in parenthesis. *p<0.1; **p<0.05; ***p<0.01		

Table 5: The impact of maize suitability (the geographic exogenous shifter G) on land available for development, population of workers, and wages

	Dependent variable:		
	$log(\bar{L})$	$log(N)$	$log(W)$
	(1)	(2)	(3)
log(maize suitability)	0.219559***	0.238882***	−0.182458***
	(0.059050)	(0.052487)	(0.063360)
Constant	3.695164***	6.031317***	13.383090***
	(0.468018)	(0.416890)	(0.502607)
Observations	635	635	614
R^2	0.093789	0.021118	0.013373
Adjusted R^2	0.092358	0.019572	0.011761
Residual Std. Error	0.360586 (df = 633)	0.859287 (df = 633)	0.833885 (df = 612)
F Statistic	65.512990*** (df = 1; 633)	13.656260*** (df = 1; 633)	8.295237*** (df = 1; 612)
Note:	Robust standard error in parenthesis. *p<0.1; **p<0.05; ***p<0.01		