Legacies of Inequality: The Case of Brazil*

Evan Wigton-Jones

October 22, 2017

Abstract

This research explores the long-run social and economic effects of inequality in Brazil. I exploit within-country variation in climate to instrument for historical land inequality in a two-stage least squares instrumental variables framework. To this end, I create an index quantifying the relative suitability of the climate for plantation versus smallholder agriculture to instrument for the municipal-level distribution of land holdings in 1920. I consider the effects of this inequality on the local institutional structure, as measured by local public welfare spending in 1923 and over the 1995-2005 time period, as well as on contemporary human development, quantified by the local human development index (HDI) for the year 2000. IV estimates show that historical inequality is associated with lower public welfare spending, both historically and in contemporary times, as well as with significant reductions in local HDI. I argue that such results are suggestive of well-entrenched local institutions, which may act to perpetuate inequality and underdevelopment.

1 Introduction

This paper analyzes the long-run impact of historical land inequality on institutional and economic development within Brazil. Brazil provides a useful context for the study of inequality, as the country's long agrarian history has made land a uniquely valuable asset which well-reflects broader inequality in welfare. Furthermore, the substantial intra-country variation in the distribution of landholdings, along with a high degree of local government

*I would like to thank Steven Helfand, Joe Cummins, and David Brady for their insight and helpful comments. I am also grateful to Eustaquio Reis at IPEA for comments on this paper and for generously sharing data on 1920 land inequality. All errors are my own.

autonomy over public spending decisions, allows for a more comprehensive analysis of inequality on within-country development. To this end, I instrument for the municipal-level distribution of landholdings in 1920 using a new index that captures the suitability of the climate for plantation over smallholder agriculture in a two stage least squares instrumental variables framework. I then show that this historical distribution of land has influenced the structure of local institutions, the persistence of economic inequality over time, as well as contemporary economic development.

A key theme of this paper is that inequality is in part determined by those environmental conditions which govern the suitability of the climate for plantation versus smallholder agricultural production. I argue that these conditions determined the local distribution of land, which in turn influenced the extent to which a small group of landowners could expend public resources on their own interests instead of those of the general public. I argue that this power was attenuated considerably in municipalities with a more egalitarian distribution of land, where elites held relatively less sway. Consistent with this hypothesis, I find that historically unequal areas devote significantly less public spending to welfare and public goods, both historically as well as in contemporary times. I also find that local inequality has persisted throughout the 20th century in Brazil, and that historically unequal areas display lower levels of development, as measured by the local human development index (HDI).

The contribution of this paper is twofold. First, I provide evidence that historical inequality can exert a strong detrimental effect on institutional and economic development. This research is thus similar to other studies of the historical determinants of institutions, such as those by Acemoglu, Johnson, and Robinson (2001, 2002). Here I argue that inequality is a key determinant of the local institutional structure as well as economic development in the Brazilian context: Hence, these results contribute to an increasingly important discussion within economics over the broader consequences of disparities in economic welfare. The wealth of subnational data in Brazil also allows for a more comprehensive analysis of inequality, enabling an examination of its effects on both human and institutional development. Additionally, the within-country context of this study permits an examination of inequality which is unobscured by heterogeneity in national-level institutions.

This paper also contributes to a growing literature on geographic endowments and long-run development. This is a theme that has received a renewed interest in recent years (for example, by Nunn and Qian (2011), Alsan (2014), and Easterly (2007)). An important finding of this research is that such endowments can create a path dependency even within countries. When local governments have a degree of autonomy over the allocation of public resources, they may be just as susceptible to elite capture as those at a national level. In such cases, the local governing norms and distribution of resources may persist for generations, even when, as in the case of Brazil, there are significant changes in national-level institutions.

This remainder of this paper is organized as follows. Section II summarizes the relevant theoretical and empirical literature on inequality. In Section III I discuss inequality and development in the context of Brazil, while in Section IV I describe the plantation suitability index (PSI). The main empirical specifications are introduced in Section V. In this section I also discuss alternative determinants of Brazilian development which may serve as confounding factors in the empirical analysis: My strategy for accounting for such factors is described here as well. Section VI describes the main empirical results, while in section VII I discuss a number of robustness checks. Section VIII concludes.

2 Literature Review

Twentieth century economic literature on inequality arose in tandem with the availability of country-level data on income shares. The first of these studies can be traced to Kuznets (1955) whose analysis of the relationship between the distribution of incomes in the United States, the United Kingdom, and Germany motivated subsequent research on the economic effects of inequality. His famed hypothesis held that inequality rose as a country developed, but later flattened or decreased as the nation became wealthier. Yet Kuznets was acutely aware that rising inequality would have ramifications beyond economics; in his paper, he questioned whether the existing political framework of many developing countries could withstand such widespread disparity in well-being. In later research (Kuznets, 1973) he further hypothesized that economic growth would lead to a decline in the relative power of farmers, small-holders and landowners; such a loss of influence would not likely be accepted without conflict or significant political change.

Yet subsequent theoretical research has been divided on the potential effects of inequality. Lewis (1954) argued that wage inequality between the agriculture and non-agriculture sectors of an economy was a necessary condition for investment and long-run development. More recently, some economists have suggested positive consequences could accrue from widespread disparity of income. Such studies rarely assume benefits flow from inequality directly, but rather from the public policies and taxation that could yield Pareto improvements in social welfare. For example, Verdier and Saint Paul (1993) construct a model in which higher inequality leads to redistribution in the form of increased funding for public education. This results in greater human capital accumulation with benign effects for economic growth. A similar result is obtained by Galor and Tsiddon (1997) who posit inequality in human capital formation is a necessary condition for economic advancement. In this case, public authorities should tolerate a degree of inequality in income and education in order to spur innovation; a country that chooses redistributive policies at an early point in its development may hinder economic growth. Benabou (2000) constructs a heterogeneous agent model to show that in an economy marked by imperfect credit and insurance markets, public policies aimed at redistributing income may be Pareto improvements. Importantly, these models assume that public authorities themselves are not corrupted or otherwise affected by inequality.

Other theoretical studies have obtained opposite conclusions; by incorporating distributional conflict and market imperfections into their models, these theorists posit corrosive effects of economic inequality. Alesina and Rodrik (1994) construct a model of endogenous growth with distributive conflict among agents with different factor endowments. They argue this initial inequality can spur strife among agents, resulting in a variety of adverse economic consequences. A similar approach is used by Persson and Tabellini (1994). In their model, such conflict may cause political authorities to tax investment and other growth-promoting economic activities. Galor and Zeira (1993) employ an over-lapping generations model to argue that initial inequality, combined with credit market imperfections, may constrain long-run investment spending and output. Their results suggest a large middle class may be important for long-run macroeconomic performance.

Empirical research on inequality developed with the emergence of cross-country data on income distributions. Gini coefficients were compiled from national surveys in OECD countries by Jain (1975) and Lecaillon et al. (1984). Fields (1989) analyzed existing datasets and calculated national inequality measures for a set of thirty-five developing countries. Taylor and Hudson (1972) and Deininger and Squire (1996) constructed Gini coefficients for land inequality at the country-level. While the land inequality measures of the latter datasets are qualitatively very different from the former measures for income, they are still used in many empirical papers as inequality in land is likely to be correlated with the accumulation of assets (Alesina & Roderik 1994). Inequalities in land may therefore closely track broader inequality in welfare; yet as will be discussed below, the connection between the two will vary depending on the country and time period under consideration. Nevertheless, the advent of such data resulted in a flurry of research employing cross-country regressions to ascertain the long-run effects of inequality. Different conclusions, however, have been drawn from such studies.

The majority of the cross-country empirical literature has found inequality to have an adverse effect on investment and economic performance. Alesina and Rodrik (1994) claimed that cross-country evidence supported their model of inequality and distributional conflict, since countries with higher levels of inequality exhibit slower growth. Alesina and Perotti (1996) suggested that income inequality, by fueling social discontent, would increase political instability. This instability would increase uncertainty, and inhibit investment. Their results seem to hold empirically; in their seventy-country sample, nations with higher inequality see less investment and growth. Perotti (1995) used data on quantile shares of income to find that countries characterized by greater equality in the distribution of income have lower fertility rates and greater investment in education. He argued that this result could be due to greater political stability in more egalitarian societies. Similarly, Deininger and Squire (1998) employed country-level data on income and asset distributions to show that nations with higher initial asset inequality face lower long-run growth.

Not all cross-country empirical studies, however, have obtained this negative relationship between inequality and socio-economic welfare. Barro (2000) found only conditional support for a negative relationship between inequality and growth; while developing countries see adverse economic effects from distributional inequalities, developed countries do not. Using data from Deininger and Squire (1996) to construct a panel of countries by inequality level, Li and Zou (1998) find that higher inequality appears to increase economic growth. This positive relationship was also obtained by Forbes (2000), who found that increases in inequality are associated with positive and statistically significant increases in GDP. Barro (2000) suggested that the idiosyncratic results of the previous two studies could be due to their use of fixed effect estimators with small sample sizes; such a methodology yields few observations, and may exacerbate measurement error. Furthermore, Banerjee and Duflo (2003) find evidence that the relationship between inequality and growth is non-linear; the authors note such a phenomenon may explain why empirical studies have reached such different conclusions.

The literature above sought to identify a relationship between inequality and welfare from a macro-economic perspective. That is, theoretical studies often utilize over-lapping generations or representative agent models, and the empirical research employs cross-country growth regressions. In the latter case, such methodology could be subject to criticisms of small sample sizes or endogeneity problems. A full review of the literature would therefore be incomplete without addressing the micro-economic or within-country studies of inequality. Although these studies are less numerous than the macro-economic analyses of inequality, many still provide a rich set of insights.

The micro-economic theoretical literature has often sought to model how initial asset inequality may influence the composition of the economy, or relationships between groups within it. For example, Banerjee and Newman (1993) model how occupational choice can influence the structure of employment within an economy. Their main insight is that initial inequality combined with credit market imperfections determines whether an agent becomes an "entrepreneur" or a "worker" (who works for the entrepreneur). The authors use this framework to explain why economies may develop small-scale cottage industries or large-scale manufacturing. Similarly, Banerjee et al. (2001) explore how inequality in asset ownership affects rent-seeking. Specifically, they model how inequality in land-ownership among members of sugar cooperatives in Maharashtra may lead to rent-seeking by wealthier farmers. They find the implications of their model are empirically consistent with the behavior of such farmers.

Empirical micro-economic or within-country studies have also yielded insights into the relationship between inequality and welfare. Galor et al. (2009) posit that inequality in the distribution of land-ownership adversely affected the emergence of human capital promoting institutions, such as public schooling or child-labor regulations; that is, landed elites had little interest in industrialization or human capital development. Using data from the United States census, they show that this hypothesis holds empirically. Rodrik (1995) compares the economic development of South Korea and Taiwan, and argues that the relatively equal asset distributions in both countries enabled effective investment decisions, efficient political coordination, and a lack of opportunities for rent-seeking. A similar conclusion is reached by Gregorio and Lee (2004), who compare differences in institutional equality among East Asian and Latin American countries. The authors argue differing levels of initial inequality may explain the disparity in institutional and socio-economic development between the two regions. Differences in initial inequality seem to have an effect within states as well: For example, Feirrera et al. (2010) find that economic growth is less effective in reducing poverty among states with higher inequality in Brazil.

3 Inequality and Development within Brazil

Agriculture has long been intertwined in the social and economic fabric of Brazil. Soon after it was claimed for the Portuguese empire in 1500, sugar became the foundation of the colonial economy. The unique suitability of agro-climatic conditions for sugarcane production in the northeast led to the development of large-scale plantation agriculture and slave labor. This system at first exploited the forced labor of Indians, before moving to imported African slaves. Sugar production depended on forced labor to such an extent that by the eighteenth century slaves constituted half the population of the northeast; in sugar growing-regions, the proportion rose to nearly three-quarters (Smith 2002). Overall, the economy that developed in the 16th and 17th centuries was one based on extractive wealth from the land; industrial development was never pursued by the landed elite or colonial administrators – a legacy that had ramifications for post-independence development (Levine 1999).

The landed elite continued to exert a heavy influence in the political and economic spheres of Brazil after independence in 1822. Wealthy land-owners were largely responsible for the forced abdication of both rulers in Brazil's short-lived monarchy; in each case, they were overthrown for their liberal sentiments, particularly in regard to abolition. Faced with a threat to their economic interests, the landed oligarchs proved united in their effort to preserve the status quo. While the sugar economy began a slow decline in the late 17th century, this did not herald the end of powerful land-holder interests: A new elite arose with the advent of coffee production around São Paulo and Minas Gerais. Following the fall of the monarchy in 1888, the era of the Old Republic (1889-1930) witnessed the rise of coronelismo, or a system of machine politics dominated by agrarian oligarchs, mostly producers of coffee and milk.

For the purposes of this analysis, the most important feature of the post-independence period was the impact of the agricultural elite on the economic and social development of areas under their influence. In general, large-scale landowners favored development only if it served their own economic interests and home locales. They were also in constant dispute with one another over patronage and government revenues. Nevertheless, nearly all were united in opposing trade unions and other forms of civil society deemed hostile to the status quo. During the Old Republic, the presidential cabinet had no ministry for labor, education, health or social welfare, and unlike Mexico and Argentina, the elite displayed no interest in widespread education (Levine 1999)[1]. Prior to 1930, the overall economic situation of the majority of Brazilians was characterized by rural poverty and a subsistence-level existence from a network of laws designed to keep them from owning and cultivating land.

Nevertheless, the prevalence of plantation agriculture and the distribution of land holdings were not uniform throughout Brazil, as heterogeneity in factor endowments led to differences in the magnitude of inequality and the rigidity of the social structure (Assuncao et al. 2012). This hypothesis is consistent with data from early agricultural censuses: Areas less conducive to sugar production, such as the inland north- and south-east, were on average less unequal in their distribution of land. Large-scale landowners may have wielded power regardless of their location, yet the extent of their holdings likely varied according to a variety factors, among them the suitability of agro-climatic conditions for plantation agriculture.

An important aspect of Brazilian history post-1930 is a general lack of a structural break that would a priori erode the legacy of agricultural interests and inequality. A series of populist governments from 1930 to 1964 put industrialization at the forefront of economic policy. Yet the power of large-scale landed interests was still very much relevant; for example, it was with their support that a military-led coup displaced a democratically elected leftist government in 1964. The return to democracy in 1985, followed by the writing of a new Constitution in 1988, witnessed a renewed emphasis on poverty reduction and socio-

[1] Until 1920, Brazil had one of the lowest rates of primary school enrollment in Latin America (Frankema, 2009)

economic development; but a large-scale redistribution of land or wealth has evidently not been successful[2]. Thus there is reason to suspect that the effects of land inequality, both historical and contemporary, may still exert an impact on certain socio-economic aspects of Brazilian society.

4 Plantation Suitability Index

To calculate an index for the relative suitability of land for plantation versus small-holder agriculture, I use a methodology that is based on Brazilian history as well as on the taxonomies of certain plants. While the former can yield insights into the specific plants on which such an index should be based, the latter will be utilized to give the index an actual structure. Specifically, I argue that using scientific taxonomies to demarcate different classes of crops offers a useful way to analytically develop an index based on a synthetic smallholder ("SS") and a synthetic plantation ("SP") crop. Below I describe the key features of this index, but a detailed derivation is left to the appendix. First, however, I discuss the most relevant smallholder and plantation crops in Brazilian agricultural history.

4.1 Prominent Plantation Crops in Brazilian Agricultural History

Manioc and maize constituted the primary food crops in colonial Brazil, with rice, beans and wheat generally of secondary importance. Manioc was grown throughout the colony, but was slightly more prevalent along the coast. In areas further inland, maize tended to predominate. Rice was grown on a more limited scale in certain regions of the north- and southeast, while wheat production was generally concentrated in the far south. Beans were grown throughout Brazil, but they seem to have been produced on a far smaller scale than

[2] With the exception of a few states in the North, state-level land inequality has remained fairly constant from the mid-1970s until 2006 (Hoffman and Ney, 2010).

either Manioc or Maize (Prado 1967). A more detailed overview of these food crops and their historical importance is given in the appendix (in subsection a).

Sugar and Coffee proved to be the most important plantation crops in Brazilian history. Sugar production began almost immediately after colonization in 1500, and served as the most valuable export crop for over three hundred years. Production was generally concentrated in large plantations along the coast. Coffee too was grown in coastal areas in the southeast, but later spread inland. It served as the dominant export crop throughout the nineteenth and twentieth centuries. While coffee was generally grown on large plantations, it was also produced on a smaller scale in the more mountainous interior. A more detailed discussion of sugar and coffee production, as well as a review of other plantation crops in Brazilian history, is given in the appendix (in subsection b).

4.2 Aspects of an Optimal Suitability Index

For the purposes of this analysis, the construction of a suitability index for plantation versus smallholder agriculture should be based on the crops listed above. Other crops were grown in the early history of Brazil, but none were as significant for either the economy or the survival of the early colonists. Furthermore, this index should only take into account those environmental conditions actually faced by the settlers and early farmers: That is, the use of modern irrigation systems and fertilizers should not be incorporated into the index. This specification would be similar to the "low input level" crop suitability indices constructed by the FAO. These indices utilize both climatic and soil conditions, but it is assumed that there is only a rain-fed water supply, with no application of nutrients or pesticides. For example, an FAO-constructed low input level potato suitability index was recently used by Nunn and Qian (2011) to analyze the effect of potatoes on historical European urbanization.

Other recent institutional-economic history literature (e.g., Michalopoulos (2012) and Alsan (2014)), utilizes similar suitability measures based on "natural" growing conditions. The previous two authors use estimates of the suitability of land for agriculture from Ramankutty et al. (2002). Ramankutty's data is well suited for such historical research since it does not take into account the use of fertilizers or modern irrigation techniques: They only account for those soil and climatic conditions which are generally unaffected by human activity. Such a methodology would be appropriate here since fertilizers and irrigation were absent from early Brazilian agriculture, with slash and burn farming being a far more common practice (Prado 1967).

As the latter fact would suggest, soil conditions are generally more vulnerable to changes by man-made activity than climate. The climate has not changed appreciably over the past several hundred years until very recently; soil conditions such as pH, in contrast, might have. For these reasons, soil characteristics are omitted from the index.

4.3 Calculating a Spectrum for the Synthetic Plantation and Smallholder Crop

The underlying rationale for constructing an index that captures the relative suitability between a synthetic smallholder and plantation crop is that these two would require fundamentally different methods of production to be profitable. This is related to the hypothesis of Engerman and Sokoloff (2000) who argued that plantation crops have much greater economies of scale than do cereals; this historically led to the use of slaves in the production of sugar and other plantation crops. This observation, however, was not novel for the time, and goes back to at least Fogel (1989), who hypothesized that the grains grown in the northern United States may have had an optimal capital-to-labor ratio that favored a family farm model, in contrast to the cotton plantations of the South, where far more labor could be utilized for profitable production.

However, cereal crops are too broad a category to base an index on, since some of them share substantial similarities with plantation crops. Perhaps for this reason, some cereals can be produced on quite a large scale. For example, maize was often grown in the same plantation fields as cotton. The most common cereal crops in Brazil, such as wheat, rice, and maize, have a wide variation in their precipitation and temperature requirements. Some, such as rice, require conditions very similar to those of plantation crops like sugar. Others, such as wheat, are better suited to colder climates with less precipitation, which differentiates them from most plantation plants.

I therefore use the scientific taxonomies of different plants to formally identify the crops on which to base the index. Importantly, such taxonomies are themselves based on the biological similarities and differences between plants. In this way, it is possible to differentiate between those cereal crops which grow in conditions only suitable for smallholder agriculture versus those which can be grown in plantation-like conditions. For example, all cereal crops belong to the family *Poaceae* ("true grasses"), which encompass an enormous range of grass-like plants. But these cereals can be differentiated by their sub-families or genera; importantly, plants of different subgroups are fairly different in their temperature and precipitation requirements, while plants of the same subgroup are generally quite similar.

Specifically, I argue that those cereals belonging to the subfamily *Pooideae* constitute the "ideal" small-holder crop; these cereals all require very different environmental conditions from plantation plants. *Pooideae* food crops include wheat, rye, oats, and barley (Hands et al., 2012); the precipitation and temperature requirements of these crops are discussed in the appendix. What is striking about the members of the *Pooideae* subfamily is how similar these requirements are across plants. But more importantly, they all thrive in far colder temperatures and with less precipitation than plantation crops such as sugar or coffee. These conditions may well be considered ideal for a synthetic small-holder crop.

One other important group of plants warrants inclusion into the index; the *Phaseolus* genus. This genus includes several bean species and hundreds of sub-species, all of them native to the Americas (Delgado-Salinas et al. 2011). Importantly, a myriad of different types of beans were grown in the New World, and, as explained above, Brazil was no exception.

The growing conditions for different bean sub-species are not significantly different, and nearly all are grouped under the name "common bean" (*Phaseolus Vulgaris*). Like Pooideae, I would argue that the genus Phaseolus should correspond to an "ideal" smallholder crop, due to the prevalence of beans in smallholder subsistence farming in much of colonial Brazil (Prado 1967).

Perhaps the most important plant on which the synthetic plantation crop should be based is that of sugar. Since the synthetic smallholder crop was not based on one plant species in particular, I argue that sugar should enter the index through its genus *Saccharum*. There are many species of sugarcane: Most can be cross-bred, and all share affinity for high temperatures and precipitation. However, two species of the *Saccharum* genus do not produce edible sugar (specifically, *S. Edule* and *S. Spontaneum*). Since the economic value of these two species is negligible, I omit them from the analysis.

Coffee is another crop which warrants inclusion into the index. Generally, coffee required a plantation-style method of production in Brazil, although it was occasionally produced on a smaller scale in areas that were more mountainous. Nevertheless, the large size of coffee plantations in states such as Sao Paulo, Rio de Janeiro, and Minas Gerais, along with the crop's economic importance in Brazilian history, should provide enough justification to include it in the index. For analytical consistency, I incorporate coffee into the index through its genus *Coffea*. This genus includes the well-known species *Coffea Arabica* and *Coffea Canephora (Robusta)*, as well as the more economically-minor species *Coffea Liberica*.

Thus in the index specification presented below, these four groups of crops are used to construct the index: That is, the *Pooideae* subfamily and *Phaseolus* genus are used to construct the synthetic smallholder crop, and the *Saccharum* and *Caffea* genera are used for the synthetic plantation crop. This index, however, omits two of the most important food crops in Brazilian history: Manioc (from the subfamily *Crotonoideae*) and maize (from the subfamily *Panicoideae*). It also omits some of the lesser plantation crops such as cotton and tobacco. I now briefly discuss the rationale for their omission.

The key reason for creating an index for smallholder/plantation agriculture is based on the principle of economic efficiency. That is, in colder climates with less precipitation, it is more efficient to produce crops on a smaller scale; the opposite would hold true in warmer areas with greater rainfall. This idea is important when considering those crops that can thrive in a wide variety of climatic conditions. In essence, I assume that such "ambiguous" crops will be produced on a scale consistent with the prevailing climatic conditions because it is economically efficient to do so: In colder and drier climates, these crops would be produced according to a small-scale production model; the opposite would hold true in more tropical areas. I thus consider the scale of production for these ambiguous crops to be at least partly endogenous to the climate. And while this relationship may have been attenuated with advent of modern agronomic techniques, it was probably quite strong historically.

An example may help illuminate this idea. The optimal conditions for growing maize include a temperature between 18 -33C, and an annual precipitation range of 600 – 1200mm. These optimal conditions overlap with both wheat (15-23C, 750-900mm) and the common sugarcane species *S. Sinense* (20-32C, 1000-1500mm). Consider a hypothetical climate where the temperature varies between 20-23C with a rainfall of 800mm. Such conditions would be ideal for smallholder production of wheat, but poor for producing sugarcane; in this case, one could reasonably hypothesize maize production would take a smallholder form, since the prevailing climatic conditions are ideal for crops that are best produced on a small-scale. Alternatively, if the hypothetical climate has an average temperature between 27-30C, with 1100mm of rainfall, then the prevailing conditions would be suitable for plantation sugarcane production, and maize would presumably be produced on a much larger scale. Note that the above example would also work with cassava (ideal temperature between 22-32C, precipitation between 800-2200mm).

There is historical evidence to support this hypothesis as well. Maize appears to have been grown on a relatively small scale in the cooler areas of Minas Gerais; but it was regularly grown on plantations in areas closer to the coast. Cash crops such as cotton and tobacco seem to have been grown on plantations in areas close to the coast, but produced on a much smaller scale in the drier inland areas (Prado 1967). The key idea is that for those plants which can survive in both smallholder and plantation conditions, the profitable production method will depend on which of these conditions predominate.

Importantly, the scale of production of crops from *Pooideae*, *Phaseolus*, *Caffea*, and *Saccharum* are themselves less endogenous to the climate. This point will again illustrate the usefulness of classifying crops based on their taxonomies. Crops of the *Pooideae* family are notorious for their tolerance to cold weather, due to the presence of a unique type of cellular anti-freeze protein. This class of ice-binding proteins inhibits ice crystallization of cellular fluid, and is unique to the *Pooideae* subfamily lineage (Sandve et al 2008). Food crops from the plant family *Fabaceae*, of which *Phaseolus* is a member, also contain anti-freeze proteins to help insulate the plants from cold shocks (Pierce, Vianelli, and Cerabolini 2005). Importantly, however, plants of tropical and sub-tropical origin (such as those from the *Saccharum* and *Caffea* genera) lack any such mechanism for survival in cold temperatures (Sanghera et al 2011).

A similar argument applies to precipitation as well. Crops of the *Pooideae* subfamily evolved to their current form during the Miocene era, when a kind of evolutionary advantage lay in adapting to temperate and arid conditions; that is, in areas wet enough to support plant growth, but dry enough to inhibit forestation (Stromberg 2011). Certain crops of the *Phaseolus* family, such as the Tepary bean, may have developed a tolerance for lower rainfall in a similar way (Gepts et al. 2005). Interestingly, other bean plants of the *Phaseolus* genus may have developed a tolerance for droughts due to selective breeding over thousands of years by early humans (Cortes et al. 2012). Similarly, *Caffea* plants are native to Africa, where frequent rainfall leaves the soil moist, nearly to the point of saturation; thus *Caffea* crops developed an affinity for high precipitation, but a weakness for arid conditions (Cheserek and Gichimu 2012). *Saccharum* plants likely appeared in a manner similar to that of *Pooideae* – at a point in history where conditions favored adaptation to wet climates (Stromberg 2011).

4.4 Index Description

Here I briefly describe the key features of the suitability index: A more detailed discussion of its derivation is given in the appendix (in subsections c through g). The overall structure of the index is similar to that of Jarvis, Ramirez-Villegas and Laderach (2013). These authors use the monthly temperature and precipitation requirements of individual crops to construct a plant-specific suitability index. That is, observed temperatures and precipitations for a given area are compared with the amounts necessary for the production of a given crop. Based on these observed conditions, two suitability indices are calculated: One for temperature, another for precipitation. Each index is calculated on a scale from 0 (completely unsuitable) to 100 (perfectly suitable). The overall suitability index is the product of the temperature and precipitation indices. Here I alter their methodology so that the index refers to the relative suitability for a synthetic plantation (SP) and a synthetic smallholder (SS) crop. In essence, I combine temperature and precipitation requirements of individual plants within each subfamily or genus to yield requirements for the SP and SS crops: This allows for a calculation of SP and SS suitability indices. Since the general hypothesis of this research is that the distribution of land is governed by the relative suitability of these two crops, the final suitability index is calculated as the ratio of the SP and SS suitability indices. A map of the municipal-level PSI and 1920 Gini coefficients are given in Figure 1 below.

As might be expected, the Amazon region displays an exceptionally high suitability for plantation agriculture, as well as extremely high inequality. In some cases, regions with the highest inequality were also those with the lowest population, and were of more marginal significance to the economy of Brazil. As a robustness check, I omit those areas which historically contained a very low population. The results described below are not sensitive to the exclusion of these underpopulated areas.

5 Empirical Analysis

5.1 Measures of Human Development, Institutions, and Contemporary Inequality

An empirical analysis of the above framework requires sub-national measures of human development, institutions, and inequality. Since the municipal boundaries of Brazil have changed substantially over the 20th century, the units of analysis in this study are "AMCs", or Minimum Comparable Areas. AMCs consist of one or more municipalities whose borders are the same in the year 2000 as they are in 1920. Thus to quantify local development, I utilize the AMC-level Human Development Index devised by the UNDP. Like its country-level counterpart, this index accounts for schooling, health, and income (with these dimensions measured by adult literacy, life expectancy, and household income, respectively). This indicator is published every ten years, coinciding with the decadal censuses. Here I use data on HDI in the year 2000 – the most recent year for which this index has been calculated at an AMC-level. In the empirical analysis below, I omit those AMCs which did not contain a state capital in 1920, but did in the year 2000 (this affects a total of five AMCs in the states of Roraima, Rondônia, Amapá, Tocantins and Mato Grosso do Sul). Excluding these AMCs lends a better degree of continuity to the analysis, since the inauguration of a state capital often coincided with the arrival of non-local elites. These elites often exerted an influence on local politics (Dietz and Myers 2002), but were likely less affected by the local distribution of resources.

Calculating an appropriate metric for institutions is more difficult. Ideally, such a measure would capture the degree to which the local institutional structure served the interests of the elite instead of the broader population. I therefore use municipal public welfare spending to quantify the extent to which local institutions are elite dominated. The reasoning is straightforward. A municipality characterized by a small agricultural elite may yield a social system built around its interests instead of the interests of the general public. This system might be very different from one with a more egalitarian distribution of land, where elites have relatively less power. It would therefore seem reasonable to suppose that elite dominated institutions would manifest in lower levels of welfare spending, since such spending reflects the degree to which ruling authorities are willing to expend resources for the benefit of the general public.

Data on municipal finances for the year 1923 are available from the Estatistica das Finanças do Brazil, published in 1926. This document classifies local spending according to eight categories: Spending on the municipal government and legislature, public education, health and sanitation, police and judiciary, public works, public electricity, payments on existing municipal debts, and "other". I define welfare expenditure as the sum of spending on education, health and sanitation, public works, and public electricity. I use the sum of these categories since the state and federal governments also provided public goods, and there is likely a degree of substitutability between these types of welfare spending. For example, if the state government funds a local school, authorities may reduce spending on education and allocate it towards health or public works. Since payments on existing municipal debts are contractual obligations (and the purpose of this debt-financed spending is not clear), I take welfare spending as a percentage of the primary budget (i.e., net of payments on municipal debts) as the final metric for the historical degree of elite-domination.

Contemporary data on municipal public spending is available from Brazil's ministry of finance. However, there are significantly more categories for municipal expenditure than there were in 1923. Such categories include: Administration and Planning, Agriculture, Assistance and Welfare, Communications, Education and Culture, Energy and Mineral Resources, Health and Sanitation, Housing, Industry, Commerce and Services, Judiciary, Legislative, Regional Development, Transportation, Security and National Defense, and Other Spending. Here I define welfare spending as the sum of expenditures on Assistance and Welfare, Health and Sanitation, Housing, Regional Development, and (for some years, as explained below) Education and Culture. My final measure of contemporary institutions is the average AMC-level share of the budget spent on the above categories for the years 1995-2005 (although there is no data on municipal spending for the year 1996).

I omit some types of spending (for example, agriculture and transportation) that might conceivably benefit the broader the public. Yet I do not include expenditure for these categories since there is a large degree of ambiguity as to whether it is targeted to the public or to the elites. For example, spending on agriculture may aid smallholders or farmers of modest means: Yet it could also serve the interests of larger agro-businesses. Similarly, transportation spending may yield benefits for all social strata, or be tailored toward the interests of the wealthy. Additionally, I only classify expenditure on education as welfare spending for the years 1995 and 1997. Beginning in 1998, the federal government became closely involved in regulating municipal education expenditure. Federal laws mandated a certain minimum in per-pupil education spending, and funds were redistributed among municipalities to ensure that those with historically lower levels of education spending or poorer education outcomes received sufficient funds. Since local authorities have considerably less allocative discretion with respect to education expenditure, I do not classify it as welfare spending post-1997.

Finally, my choice of 1995 to 2005 time period is based on three considerations. First, this timeframe roughly coincides with available data on the AMC-level human development index (in the year 2000), as well as on income and land inequality (which are available for the years 2000 and 2006, as described below). Second, data on municipal spending prior to the year 1995 is generally considered unreliable due to hyperinflation and missing observations (Dix-Carneiro, Soares, and Ulyssea 2016). Lastly, I average AMC-level expenditure over ten years in order to create a better measure of the general, more long-term patterns of public spending across AMCs.

For contemporary measures of inequality I draw from two sources: The agricultural census of 2006 and the population census of 2000. The former can be used to calculate Gini indices for the distribution of land, while the latter can be used to construct Ginis for income (using per capita household income). These measures can be regressed on the 1920 land Gini index to evaluate the persistence of inequality over time.

5.2 Necessary Conditions for an Instrumental Variable

The value of an IV strategy in this context is that it can obviate two key problems inherent to the application of a standard ordinary least squares specification. The first problem is that of simultaneous causality. That is, while municipal inequality may inhibit local development, this same lack of development may work to perpetuate inequality. A more general problem, however, is that inequality may be associated with certain factors (for example, "local politics") that cannot be controlled for. This may result in omitted variable bias. Both issues suggest an instrumental variables research design.

For the IV strategy to be convincing, the instrument must be both relevant and exogenous. Relevancy implies that the index is positively correlated with inequality. As I show in Section VI, the index is a strong predictor of inequality. The exogeneity requirement (or exclusion restriction) requires that, conditional on included control variables, the effect of the instrument on development is solely through land inequality.

I will argue here that the exclusion restriction is plausible. Inequality in the distribution of land may be correlated with a variety of social, political, and economic factors which would manifest in the error term of a simple OLS regression. However, it is unlikely that the climate itself is correlated with any of these extraneous factors. And unlike instruments based on "policy shocks" or other human-devised institutional changes, the climate is a cause of inequality rather than a result of it. That is, lagged values of land inequality will not constitute omitted variables with a climate-based instrument. Nevertheless, as I describe in greater detail below, the most likely threats to exogeneity include the disease environment, transportation costs, the suitability of the climate for agriculture in general, and immigration. In what follows, I present the baseline empirical model and discuss each of the factors identified above.

5.3 Baseline Empirical Specifications

In the most basic specification, the first stage estimation equation will regress land inequality from 1920 on the plantation suitability index. Controls for the factors described above can then be added. The general structure of the first stage can therefore be described as follows:

$$Historical\ Land\ Inequality_i = \psi + \varphi SUIT_i + \eta Z + \varepsilon_i \tag{1}$$

Where i indexes the AMC; ψ is the constant term, SUIT refers to the value of the plantation suitability index, Z is a vector of controls, and ε is the error term.

The hypothesis presented here necessitates three second-stage specifications. First I show the effects of 1920 inequality on the extent of elite domination of historical and contemporary institutions, with this domination proxied by the share of the AMC-level budget devoted to public welfare spending. I then analyze the causal effect of this historical inequality on contemporary development outcomes, as measured by the HDI. The second-stage empirical specifications for these hypotheses can be summarized as follows:

$$Historical\ Institutions_i = \alpha_1 + \beta_1 Historical\ \widehat{Land}\ Inequality_i + \delta_1 \boldsymbol{Z} + \epsilon_i \tag{2}$$

$$Contemporary\ Institutions_i = \alpha_2 + \beta_2 Historical\ \widehat{Land}\ Inequality_i + \delta_2 \boldsymbol{Z} + \epsilon_i \tag{3}$$

$$Contemporary\ Outcomes_i = \alpha_3 + \beta_3 Historical\ \widehat{Land}\ Inequality_i + \delta_3 \boldsymbol{Z} + \epsilon_i \tag{4}$$

Where Land Inequality denotes the predicted values of inequality obtained from the first stage regression, and Z denotes a vector of controls.

Lastly, I examine the degree of persistence in local inequality and institutions through a simple correlation of contemporary measures of these variables on their historical measures.

That is, metrics of contemporary inequality can be regressed on 1920 land inequality, while public welfare spending over the 1995-2005 time period can be regressed on such spending from 1923. These OLS specifications can be described as follows:

$$Contemporary\ Inequality_i = \alpha_4 + \beta_4 Historical\ Inequality_i + \epsilon_i \tag{5}$$

$$Contemporary\ Institutions_i = \alpha_4 + \beta_4 Historical\ Institutions_i + \epsilon_i \tag{6}$$

Where α and ϵ are the intercept and error terms. The coefficients of interest, β4 and β5, will thus quantify the unconditional correlation between the historical and contemporary distribution of resources and public welfare spending, respectively.

As described above, there are other determinants of development within Brazil, and some of them may be correlated with both inequality and the climate. First, the prevalence of tropical diseases could influence development through institutional channels (Acemoglu, Johnson, and Robinson, 2001), or through contemporary public health (Gallup and Sachs, 2001). To account for both of these effects, I control for the disease environment using the malaria ecology index constructed by Kizsewski et al (2004). Second, municipalities located further from the coast likely faced greater isolation and higher transportation costs. Since climatic conditions can vary depending on this distance as well, I control for the distance of an AMC to the Atlantic coast. Additionally, the PSI might be correlated with a greater amenability for agriculture in general. I therefore utilize data from Ramankutty et al. (2002), which quantifies the suitability of land for agricultural production using climatic and soil quality variables. Lastly, immigrants likely brought different forms of human and social capital to Brazil, and also clustered in the more temperate regions of the South. To account for the effects immigration, I include a control for the log number of immigrants per 1000 people in 1920.

5.4 Summary Statistics

Table 1 provides summary statistics for all outcome and control variables. To ease the interpretation of the PSI in the regressions that follow, I normalize the index to have a mean of zero and a standard deviation of one. Although mean AMC-level land inequality in 1920 is less than that in 2006, this is almost certainly due to a downward bias resulting from a fewer number of farm size bins in the 1920 census. Interestingly, the historical measures of inequality and public welfare spending have a much larger variance than the contemporary measures do. It should also be noted that while the malaria severity index reaches a maximum value of 7, this is not particularly high by international standards: In equatorial Africa, the index approaches 30. To provide a frame of reference for the measure of agricultural suitability, note that this metric is scaled between 0 and 1, with a higher number indicating a greater suitability of land for agricultural production.

6 Empirical Results

6.1 First Stage Estimation

Table 2 displays the first stage estimation results (eq. 1). Column one presents a simple regression of 1920 land inequality on the PSI. The F-statistic of 31 is well above the critical value of 10 that is generally considered indicative of weak instruments (Stock and Yogo, 2005). This coefficient would imply that a one standard deviation increase in the PSI would yield a 2.5 unit increase in the Gini index (or approximately a .19 standard-deviation increase in inequality). The remaining columns add controls for malaria, distance to the coast, suitability for agriculture, and the prevalence of the foreign born population. The coefficient and significance of the PSI remain stable as these controls are added.

6.2 Historical Inequality and Institutions

Table 3 displays OLS and IV estimates from regressions of historical public welfare spending on 1920 land inequality (eq. 2). IV specifications are shown in columns one through five, while the complete OLS specification is shown in column six. The OLS estimates suggest a small, although statistically significant, impact of historical inequality on historical institutions. Conversely, the IV estimates suggest the impact is substantial: In the most complete IV specification (column four), a one unit increase in the 1920 land Gini is associated with a .83 percentage point decline in welfare spending. Nevertheless, as I show in the next section, this result is primarily driven by localities in the North of Brazil (in the modern-day states of Acre, Amazonas, Para, Tocantins, Roraima, Amapá, and Rondônia). With region controls, these IV estimates are insignificant, and the entire effect of inequality is absorbed by the region indicators.

6.3 The Persistence of Local Inequality and Institutions

To assess the persistence of inequality, I examine the correlation between contemporary measures of the distribution of resources (income and land inequality in 2000 and 2006, respectively) and the distribution of land in 1920. These correlations are shown in the first two columns of Table 4. Column one suggests that a one unit increase in the 1920 land Gini is associated with a .11 unit increase in this metric for in 2006 (a 15% increase over the mean land inequality in 2006). Similarly, the second column suggests that a one unit increase in 1920 inequality is associated with a .04 unit increase in the 2000 income Gini, or a 7% increase over the mean income inequality. To assess the persistence of local institutions, I regress the current (1995-2005) percentage of the local budget spent on public welfare on this percentage from 1923. The third column shows historical welfare spending to be a significant predictor of such spending in contemporary times: The coefficient suggests that a one percentage point increase in welfare spending in 1923 is associated with a .07 percentage point increase in such spending over the 1995-2005 timeframe, or a 7% increase over average contemporary spending. These relatively strong correlations are suggestive of a persistence in inequality and local institutional norms over the 20th century.

6.4 Historical Inequality and Contemporary Institutions

Table five reports the OLS and IV specifications for the effect of 1920 inequality on contemporary institutions (eq. 3). Over the 1995-2005 time period, AMCs allocated an average of 43% of their budget to welfare spending as defined above (with a minimum of 6% and a maximum of 64%). While OLS estimates in column six suggest a negligible impact of historical inequality on contemporary institutions, the IV estimates suggest the impact is large: For example, the most complete IV specification in column five would suggest a one unit increase in predicted inequality is associated with a .42 percentage point decrease in contemporary welfare spending. Furthermore, as I show below, this relationship is not sensitive to the inclusion of region controls.

6.5 Historical Inequality and Contemporary Development

Table six presents the OLS and IV specifications for the impact of historical inequality on human development in the year 2000 (eq. 4). Mean HDI (scaled from 0 to 100) is 70 (with a minimum of 52 and a maximum of 89). As in previous tables, columns one through five display the IV estimates, while column six reports the complete OLS specification. Again, the OLS specification shows a negligible effect of historical inequality on contemporary development. Conversely, the IV estimates suggest the effect is considerable: In the most complete specification (column 4), a predicted one unit increase in 1920 inequality is associated with a .33 unit decrease in HDI. As I show below, the estimated coefficients are nearly unchanged with the inclusion of region controls, suggesting that this effect is due to differences in development within regions, and not only to differences between them.

6.6 Accounting for Differences in the OLS and IV Specifications

The OLS and IV estimates differ significantly when assessing the effect of historical inequality on contemporary institutions and development. Yet this difference may not be unusual in the context of Brazilian history. There was almost certainly an elite presence in areas that were of commercial, military, or otherwise geo-strategic importance. It is therefore likely that such areas were more unequal. But the importance of such regions would suggest that they were settled for longer periods of time, and may have contained more economic opportunities for the people that inhabited them: Hence a greater degree of human development.

What is important for the analysis presented here is that such extraneous factors are orthogonal to the climatic suitability for plantation agriculture (conditional on the included controls). As I have argued previously, there is strong evidence to suggest that they are. Thus the differences between these two specifications may illustrate the importance of the IV approach: An exogenous source of variation in historical inequality is critical for accurately understanding these relationships when they are otherwise obscured by so many confounding variables.

7 Robustness Checks

In this section I present five robustness checks. First, I include region controls and omit low-population AMCs from the analysis. Second, I use per-capita public welfare spending as an alternative measure of institutions. Next, I repeat the analysis above with crop controls.

I then include a specification where I account for the historical effects of slavery. Finally, I control for the possible differential effects of land inequality among municipalities that are relatively more urbanized. To save space, I only include the IV specifications (all OLS regressions are qualitatively similar to those presented above). In what follows, I discuss the rationale for each check, and discuss how the IV results differ from those presented above.

7.1 Region Controls and the Exclusion of Low-Population AMCs

Region controls can be added to the baseline specifications above to analyze the extent to which the impacts of inequality are due to differences between regions as opposed to within them. Importantly, the distinctions between the five regions of Brazil are based largely on climate and geography, and hence there are unlikely to be any sources of omitted variable bias at this geographic level. Table 7 displays the first-stage and IV regressions with these region-fixed effects. Region controls have little effect on the PSI as a predictor of land inequality in 1920; they also have a negligible impact on the relationship between historical land inequality and contemporary welfare spending and development. Surprisingly, however, the effect of land inequality on historical welfare spending is completely attenuated with the addition of region controls: The predicted 1920 land Gini index is insignificant, and the sign changes from negative to positive. The coefficient on the indicator variable for the North is statistically significant and large in magnitude, indicating that the effects of inequality on historical institutions are being driven by AMCs from this region of Brazil.

One potential explanation for this result may be the low population of the North[3] (which is predominantly covered by rainforest). While it constitutes the largest region of Brazil, the North also contains many municipalities which were historically sparsely populated and of marginal economic significance (which to some extent remains true today). The municipal

[3]Here I use the official definition of Northern Brazil, which includes the state of Tocantins (which was separated from Goiás, a state in the Center-West, in 1988).

distribution of the population in 1920 is shown in Figure 2. These low-population areas may have devoted a lower proportion of their resources to public welfare spending, if, for example, they contained more isolated and self-sufficient households.

To examine whether the relationship between historical inequality and the outcomes of interest (including historical welfare spending) are being driven by low-population areas, I rank all AMCs by population size in 1920, and exclude those in the bottom quintile. In practice, this means excluding all AMCs with a population less than 10,017 (this affects 191 observations). Roughly 6% of the excluded AMCs come from the North, 34% from the Northeast, 44% from the Southeast, 9% from the South, and 7% from the Center-West. Results are presented in Table 8. Interestingly, the exclusion of the low-population areas does not affect the relationship between historical inequality and historical welfare spending, nor does it affect the relationship between historical inequality and the contemporary outcomes.

In Table 9 I present additional analyses of 1920 inequality on historical public welfare spending, with and without the North of Brazil. Importantly, AMCs of this region all share climatic conditions which are remarkably suitable for plantation agriculture, and nearly all of the North's municipalities lie in the top quintile of the PSI distribution. As might be expected, the concentration of land is extreme, with an average AMC-level Gini of nearly 72. The effect of this inequality is shown in Columns one and two of Table 9. The first column displays the coefficient on 1920 inequality (instrumented with the PSI) in the full sample of AMCs, while column two omits AMCs in the North (all controls are omitted from both of these specifications). When the North is excluded (column two), the coefficient on inequality decreases by 29% (although it is still significant at the 5% level). When the malaria control is added in column three, however, the magnitude of historical inequality dissipates completely. The same is true when region controls are added (column four). Thus while the unconditional relationship between the instrumented values of 1920 inequality and historical public welfare spending is robust to the exclusion of the North, this robustness is due to variation between regions; when region fixed effects are added, the negative relationship disappears.

That the relationship between 1920 land inequality and historical welfare spending is less robust than that of 1920 inequality and contemporary welfare spending may be due to the limited availability of historical data on municipal expenditure. Data on the latter exists for only one year, and there are over 100 missing observations. Furthermore, municipalities in the early 20th century were able to issue debt, and deficit spending among local governments was quite common; currently, however, such spending is restricted (Mettenheim, 2016). This historical deficit spending may have weakened the trade-off between expenditure on public welfare and on those items favored by the elite. In such circumstances, historical public welfare spending may not as accurately reflect authorities' decision-making processes and the ability of the elite to divert resources to their own interests.

7.2 Per-Capita Spending as a Measure of Institutions

The main analysis has used the share of public welfare spending as a metric for the elite domination of local institutions. However, welfare spending in per-capita terms could constitute an alternative measure. A potential drawback is that per capita spending is sensitive to the relative wealth of the municipality. Municipalities in more prosperous regions have generally received greater tax revenue from exports, property, and industry, both historically as well as in contemporary times. It is conceivable, and in fact likely, that municipalities can be poor and contain an egalitarian distribution of resources (the relative equality of the poorer inland Northeast, as seen in Figure 1, would seem to be evidence for this). In such cases, local governments may devote a larger share of spending to public welfare, but this may not translate into a higher per-capita amount since the overall budget is limited.

Nevertheless, Table 10 presents results with this alternative institutional metric. Columns one and two present IV regressions of log per-capita public welfare spending in 1923 on 1920 inequality (with and without region controls). The results in column one are marginally weaker than those obtained in the baseline specification (historical inequality is only significant at the 10% level), and region controls again attenuate the significance of inequality completely. Columns three and four present these IV regressions with contemporary public welfare spending. Here the results are similar to those obtained in the main specification: Historical inequality exerts a strong detrimental impact on per-capita welfare spending, and this effect is robust to the inclusion of region controls.

7.3 Crop Controls

As the PSI is based on growing conditions of certain crops which are held to be uniquely smallholder or plantation in their method of production, a potential concern is that the estimated effects of inequality are in fact due to the different types of crops produced. To account for this, I control for the AMC-level production of sugarcane, coffee, wheat and beans in both the first stage regression and IV regressions. For each crop, the unit of analysis is tons per-capita. To account for the effect of ranching (historically, an important economic activity in Brazil), I include a control for cattle production per capita. All variables are obtained from the 1920 Census of Agriculture, although the data is available electronically through IPEA.

Table 11 presents the results with these crop-specific controls. While the variables for Sugar, Coffee, Wheat, and Cattle production tend to be significant, they do not affect the significance of inequality nor do they greatly alter the magnitude of the coefficient. This would suggest that the results presented above are due to inequality, and not to the institutional legacies of specific crops.

7.4 Legacies of Slavery

A key hypothesis of this research holds that certain climatic conditions may lead to increasing returns to scale in the agricultural production process. Historically, this likely manifested in large plantations maintained by a multitude of slaves and a small cabal of landowners. Slavery has played an important role in the social and economic development of Brazil, and it is likely a key channel through which climatic conditions determined the historical distribution of wealth and income. An important question, however, may be whether the effects of inequality are completely driven by slavery. In this case, the results described above may not be due to land inequality per se, but rather to the institution of slavery.

In general, there is no overwhelming reason to believe that the estimated effects of land inequality are solely driven by this. For example, immigrant and native Brazilian labor were often utilized on plantations, especially after abolition. Nevertheless, to examine its impact I control for the enslaved fraction of the population by AMC using data from the 1872 census (slavery was abolished in Brazil in 1888, and data on race was not collected in the 1920 census). Here I control for the proportion of slaves as a percentage of the 1872 population in each AMC.

As hown in Table 12, the effects of slavery are relatively minor. The coefficient on inequality is only slightly attenuated in regressions of historical spending on inequality, while the effect of slavery on contemporary institutions (1995-2005 welfare spending) appears negligible. Similarly, the inclusion of a slavery control only marginally weakens the relationship between historical inequality and HDI. While it may appear odd that slavery has such a marginal effect on these outcomes, this result may make sense given the state of the Brazilian economy in the latter 19th century. The 1872 census coincided with a coffee boom in which the slave population was largely concentrated on coffee plantations in the Southeast. The states with the largest share of slaves included Rio de Janeiro (38%), Espirito Santo (27%), Minas Gerais (19%), and São Paulo (16%) (only Maranhão, with a slave population of 19%, approaches the levels of the Southeast). Coffee production, however, tended to be a transient phenomenon, with production moving further inland every 25 to 30 years after the soil was depleted of nutrients (Topik 1999). If the prevalence of slavery followed a similar pattern, this may have inhibited any longer-term legacies of this institution.

7.5 Urbanization and Historical Inequality

The effects of historical inequality may differ depending on a municipality's level of urbanization. A priori one might expect these effects to be mitigated in more urbanized areas, since the power of the agrarian elite was likely circumscribed by a more industrial economy. Nevertheless, the extent of urbanization is likely endogenous to historical land inequality. In the 19th and early 20th centuries the agrarian oligarchs were in competition with industry over scarce labor, and industrialization was often viewed with hostility by the landed elite. For example, these large-scale landowners lobbied for controls on internal migration to stem the flow of agricultural workers to the cities (Wagner and Ward 1980). Thus the extent of urbanization is likely a channel through which historical inequality affects institutional and development outcomes.

My empirical strategy for examining the effects of urbanization is as follows. I obtain data on the percentage of the urban population by AMC from IBGE (for the year 1940 – the first year for which urbanization data is available). The median AMC-level share of the population living in urban areas is .18. I create an indicator variable which takes on value 1 if an AMC has more than 18% of its population residing in urban areas, and 0 if it is less. Since urbanization may affect the coefficient on historical inequality, I include an interaction term where 1920 inequality is multiplied by the indicator for urbanization. (More specifically, I instrument for this interaction term by interacting the PSI with the urban indicator). Results are presented in Table 13.

The effect of inequality on contemporary public welfare spending (column one) is slightly attenuated with the addition of the urban indicator (although both variables are significant).

The interaction term in this specification is insignificant. The effects of urbanization are similar in regressions of HDI on historical inequality, although in this regression both the urban indicator and its interaction with 1920 inequality are significant predictors of contemporary HDI. Note that an analysis of the effect of urbanization on 1920 welfare spending is not possible since statistics on the former have only been collected from 1940 onwards. Yet the median AMC in 1940 had only 18% of its population in an urban area. Given this, it is likely that all but a few municipalities would have an overwhelmingly large rural population in 1920. Furthermore, most municipalities at this time consisted of a city (or cities) and the surrounding countryside. Therefore, the landed elite may have had an influence even in majority-urban areas since there would be an abundant supply of agricultural land from they could derive their wealth and income. Thus even if urbanization controls for 1920 could be found, the results are unlikely to be significantly different from those obtained in the main specifications presented above.

8 Conclusion

This study has used plausibly exogenous variation in climatic suitability for plantation versus smallholder agriculture to analyze the causal effect of inequality on long-run development. Instrumenting for the local distribution of landholdings in 1920 allowed for an analysis of historical inequality on institutions (as proxied by local public welfare spending) and contemporary development (measured by AMC-level HDI). This IV strategy revealed the adverse effects of inequality on such outcomes, a finding which would remain obscured by confounding factors in an OLS framework. More specifically, the IV estimates suggest that a predicted one unit increase in the 1920 land Gini index is associated with a .83 percentage point reduction in welfare spending in 1923, and a .42 percentage point reduction in welfare spending averaged over the 1995-2005 time period. While the effect of inequality on historical welfare spending is primarily driven by localities in Brazil's North, the reduction in contemporary welfare spending is not driven by inequality in any one region. Furthermore, historical inequality also exerts a detrimental effect on local development, with a (predicted) one unit increase in the 1920 land Gini index associated with a .33 unit decrease in the municipal-level HDI for the year 2000.

This research is fundamentally a study of long-run development, but it contains lessons for both contemporary public policy and future research on inequality. First, these findings suggest that it is important to account for within-country differences in long-run development patterns when assessing the desirability of certain social or economic policies. For example, investing local authorities with significant discretion over public spending may not be advisable in areas marked by great disparities in economic welfare. If local institutions are vulnerable to elite capture, then political decision making may not adequately take into account, or be responsive to, the interests of the more marginalized social strata. In practical terms, this suggests an oversight role for state or federal authorities, who may be less susceptible to the different norms of local governance.

That a given distribution of resources can persist for generations inherently raises normative questions over the proper role of the state in ensuring an equitable distribution of resources. These questions are especially pertinent given the pernicious effect of inequality on development. Here institutions may allow the existing elite to capitalize on structural economic changes, while inhibiting the ability of others to do so. This would seem to be the case in Brazil, which is no longer an agrarian country but retains a distribution of economic welfare that reflects its agricultural past. When institutions preclude the benefits of economic modernization to the broader populace, crucial questions of economic justice and redistribution may need to be addressed.

These results also contain methodological implications for empirical studies of inequality. In particular, a standard ordinary least squares approach may not be suitable for this type of research. In this study, confounding factors led to substantially biased estimates, which when viewed in isolation, would suggest a negligible effect of inequality on institutional and contemporary development. Yet the IV specification revealed the OLS results to be spurious. Given the importance of inequality for issues of economic development, growth, and equity, it is imperative that researchers exercise vigilance in choosing an appropriate empirical strategy for evaluating its consequences.

References

[1] Acemoglu, D., Johnson, S. & Robinson, J. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. American Economic Review 91(5), pp. 1369-1401.

[2] Acemoglu, D., Johnson, S. & Robinson, J. (2002). Reversal of Fortune: Geography and Institutions in the Making of the World Income Distribution. Quarterly Journal of Economics 117(4), pp.1231-1294.

[3] Alesina, A. & Perotti, R. (1996). Income Distribution, Political Instability, and Investment. European Economic Review 40(6), pp. 1203-1228.

[4] Alesina, A., & Rodrik, D. (1994) Distributive Politics and Economic Growth. Quarterly Journal of Economics 109(2), pp.465-490.

[5] Alsan, M. (2014). The Effect of the TseTse fly on African Development. The American Economic Review 105(1), pp. 382-410.

[6] Assuncao, J., Naritomi, J., & Soares, R. (2012). Institutional Development and Colonial Heritage Within Brazil. Journal of Economic History 72(2), pp. 393-422.

[7] Barro, R. (2000). Inequality and Growth in a Panel of Countries. Journal of Economic Growth 5(1), pp. 5-32.

[8] Banerjee, A. & Duflo, E. (2003). Inequality & Growth: What do the Data Say? Journal of Economic Growth 8(3), pp 267-299.

[9] Banerjee, A., Mookherjee, D., Munshi, K, & Ray, D. (2001). Inequality, Control Rights, and Rent Seeking: Sugar Cooperatives in Maharashtra. Journal of Political Economy 109(1), pp. 138-190.

[10] Banerjee, A. & Newman, A. (1993). Occupational Choice and the Process of Development. Journal of Political Economy 101(2), pp. 274-298.

[11] Benabou, R. (2000). Unequal Societies: Income Distribution and the Social Contract. American Economic Review 90(1), pp.96-129.

[12] Bertola, L., Castelnovo, C. & Willebald, H. (2012). Income Distribution in Brazil, 1870-1920. XVII Jornadas Anuales de Economia, Banco Central de Uruguay

[13] Bethell, L. (1987). Colonial Brazil. Cambridge: Cambridge University Press

[14] Bethell, L. (1989). Brazil: Empire and Republic 1822-1930. Cambridge: Cambridge University Press

[15] Bergad, L. W. (1999). Slavery and the Demographic and Economic History of Minas Gerais, Brazil 1720-1888. Cambridge: Cambridge University Press

[16] Carvalho De Mello, P. (1978). Aspectos economicos da organizaçao do trabalho da economia cafeeira do Rio de Janeiro, 1850 – 88. Revista Brasileira de Economia 32 (1), pp 19-67

[17] Clarence-Smith, W. & Topik, S. The Global Coffee Economy in Africa, Asia, and Latin America, 1500-1989. Cambridge: Cambridge University Press.

[18] Cheserek, J., & Gichimu, B. (2012). Drought and Tolerance in Coffee: A Review. International Journal of Agricultural Science and Soil Science 2(12), pp. 498-501.

[19] Cortes, A., et al. (2012). Molecular Ecology and Selection in Drought Related ASR Gene Polymorphisms in Wild and Cultivated Common Bean (Phaseolus Vulgaris L.). BMC Genetics 13(58).

[20] De Ferranti, D., Guillermo, E., Ferriera, G. & Walton, M. (2004). Inequality in Latin America and the Caribbean: Breaking with the Past? Washington, DC: World Bank.

[21] Delgado-Salinas, A., Thulin, M., Pasquet, R., Weeden, N. & Lavin, M. (2011). "Vigna (Leguminosae) sensulato: the names and identities of the American segregate genera". American Journal of Botany 98(10), pp. 1694- 1715.

[22] Deininger, K. & Squire, L. (1998). New Ways of Looking at Old Issues: Inequality and Growth. Journal of Development Economics 57(2), pp. 259-287.

[23] Deininger, K. & Squire, L. (1996). A New Dataset Measuring Income Inequality. World Bank Economic Review 10(3), pp. 565-591.

[24] Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. Econometrica 78(6), pp 1863- 1903.

[25] Dell, M., Jones, B. & Olken, B. (2009) Temperature and Income: Reconciling New Cross- Sectional and Panel Estimates. American Economic Review Papers and Proceedings 99(2), pp. 198-204.

[26] DeWitt, J. (2002). Early Globalization and Economic Development of the United States and Brazil. Westport, CT: Praeger Publishers.

[27] Dietz, H. & Henry, A. (2002). Capital City Politics in Latin America: Democratization and Empowerment. London, UK: Lynne Rienner Publishers.

[28] Dix-Carneiro, R., Soares, R. & Ulyssea, G. (2016). Economic Shocks and Crime: Evidence from the Brazilian Trade Liberalization. Columbia University Working Paper

[29] Easterly, W. (2007). Inequality Does Cause Underdevelopment: Insights From a New Instrument. Journal of Development Economics 85(2), pp.755-776.

[30] Engerman, S. & Sokoloff, K. (2000). History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World. Journal of Economic Perspectives 14(3), pp. 217-232.

[31] Estatistica das Finanças do Brazil: Receitas, Despesas e Dividas Publicas da União, dos Estados e dos Municipios. Ministerio da Agricultura, Industria e Commercio. Rio de Janeiro: 1926

[32] Ferreira, F., Leite, P. & Ravallion, M. (2010). Poverty Reduction Without Economic Growth? Explaining Brazil's Poverty Dynamics, 1985-2004. Journal of Development Economics 93(1), pp. 20-36.

[33] Fields, G. (1989). Changes in Poverty and Inequality in Developing Countries. World Bank Research Observer 4(2), pp 167-185.

[34] Fogel, R. (1989). Without Consent or Contract: The Rise and Fall of American Slavery. Norton Publishers: New York, NY.

[35] Forbes, K. (2000). A Re-assessment of the Relationship Between Inequality and Growth. American Economic Review 90(4), pp. 869-887.

[36] Frankema, E. (2006). Colonial Origins of Land Inequality: Causes and Consequences of Land Redistribution. Ibero-America Institute for Economic Research Discussion Paper No. 119

[37] Frankema, E. (2009) The Expansion of Mass Education in Twentieth Century Latin America: A Global Comparative Perspective, Revista de Historia Económica 27(3), pp. 359-396

[38] Gallup, J, & Sachs, J. (2001). The Economic Burden of Malaria. American Journal of Tropical Medicine and Hygiene 64(1), pp. 85-96.

[39] Galor, O., & Tsiddon, D. (1997). The Distribution of Human Capital and Economic Growth. Journal of Economic Growth 2(1), pp.93-124.

[40] Galor, O., Moan, O., & Vollrath, D. (2009). Inequality in Land-ownership, The Emergence of Human Capital Promoting Institutions, and the Great Divergence. Review of Economic Studies 76(2009), pp. 143-179.

[41] Galor, O. & Zeira, J. (1993). Income Distribution and Macroeconomics. Review of Economic Studies 60(1), pp. 35-52.

[42] Gennaioli, N., et al. (2013). Human Capital and Regional Development. Quarterly Journal of Economics 128(1), pp. 105-164.

[43] Gepts, et al. (2005). Legumes as a Model Plant Family: Genomics for Food and Feed report of Cross-Legume Advances through Genomics Conference. Plant Physiology 137(4).

[44] Gregorio, J., & Lee, J. (2003). Growth and Adjustment in East Asia and Latin America. Central Bank of Chile Working Papers no. 245, pp 1-66.

[45] Hoffman, R., & Ney, M.G. (2010). Agrarian structure and agricultural property in Brazil: large regions and states. Brasilia, Brazil: Ministry of Agrarian Development

[46] Hall, M. (1969). The Origins of Mass Immigration in Brazil (Unpublished doctoral dissertation). Columbia University, New York.

[47] Hands, P., Kourmpetli, S., Sharples, D., Harris, R., & Drea, S. (2012). Analysis of Grain Characters in Temperate Grasses Reveals Distinctive Patterns of Endosperm Organization Associated with Grain Shape. Journal of Experimental Botany 63(17).

[48] Hijmans, R., et al. (2005). Very High Resolution Interpolated Climate Surfaces for Global Land Areas. International Journal of Climatology 25(1965-1978).

[49] Jain, S. (1975). Size & Distribution of Income: A Comparison of Data. Unpublished manuscript, The World Bank.

[50] Juliano, S. & Lounibos, L. (2005). Ecology of Invasive Mosquitos: Effects on Resident Species and on Human Health. Ecology Letters 8(5), pp. 558-574.

[51] Kiszewski, A. et al (2004). A Global Index Representing the Stability of Malaria Transmission. American Journal of Tropical Medicine and Hygiene 70(5), pp. 486-498.

[52] Kuznets, S. (1955). Economic Growth and Income Inequality. American Economic Review 45(1), pp. 1-28.

[53] Kuznets, S. (1973). Modern Economic Growth: Findings and Reflections. American Economic Review 63(3), pp.247-258

[54] Lecaillon, J., et al. (1984). Income Distribution and Economic Development: An Analytical Survey. International Labor Office, Geneva.

[55] Levine, R. (1999). The History of Brazil. Westport: Greenwood Press.

[56] Lewis, W. (1954). Economic Development with Unlimited Supplies of Labor. The Manchester School 22, pp. 139-191

[57] Li, H. & Zou, H. (1998). Income Inequality is Not Harmful for Growth: Theory and Evidence. Review of Development Economics 2(3), pp. 318-344.

[58] Libby, D. (1984). Trabalho escravo e capital no Brasil – o caso de Morro Velho

[59] Metcalf, A. (2005). Go-Betweens and the Colonization of Brazil 1500-1600. Texas: University of Texas Press.

[60] Mettenheim, K. (2016). Monetary Statecraft in Brazil: 1808-2014. New York: Routeledge.

[61] Michalopoulos, S. (2012). The Origins of Ethnolinguistic Diversity. The American Economic Review, 102(4), pp. 1508-39.

[62] Michalopoulos, S. & Pappiaoanou, E. (2012). Pre-Colonial Ethnic Institutions and Contemporary African Development. Econometrica 81(1), pp.113-152.

[63] Michalopoulos, S. & Pappiaoanou, E. (2014). National Institutions and Subnational Development in Africa. Quarterly Journal of Economics 129(1), 151-213.

[64] Nunn, P. & Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. Review of Economics and Statistics 94(1), pp. 20-36.

[65] Nunn, N. & Qian, N. (2011). The Potato's contribution to Population and Urbanization: Evidence from a Historical Experiment. Quarterly Journal of Economics 126(2), pp.593-650.

[66] Prado, C. (1967). The Colonial Background of Modern Brazil. University of California Press: Berkeley

[67] Perotti, R. (1996). Growth, Income Distribution, & Democracy: What the Data Say. Journal of Economic Growth 1(2), pp.149-187.

[68] Persson, T. & Tabellini, G. (1994). Is Inequality Harmful for Growth? American Economic Review 84(3), pp. 600-621.

[69] Pierce, S., Vianelli, A. & Cerabolini, B. (2005). From Ancient Genes to Modern Communities: The Cellular Stress Response and the Evolution of Plant Strategies. Functional Ecology 19(5).

[70] Ramankutty, N., Foley, J., Norman, J., McSweeney, K. (2002). The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Climate Change. Global Ecology and Biogeography 11(5), pp. 377-392.

[71] Ramirez-Villegas, J., Jarvis, A., & Lederach, P. (2013). Empirical Approaches for assessing climate change on agriculture: The Eco-crop model and a case study with grain sorghum. Agriculture and Forest Meteorology 170 (2013), pp. 67-78.

[72] Rocha, R., Ferraz, C., & Soares, R. (2015). Human Capital Persistence and Development. IZA discussion paper No. 9101.

[73] Rodrik, D. (1995). Getting Interventions Right: How South Korea and Taiwan Grew Rich. Economic Policy 10(20), pp. 53-107.

[74] Sandve, S., et al. (2008). Tracking the Evolution of a Cold-Stress Associated Gene Family in Cold Tolerant Grasses. BMC Evolutionary Biology 8(245)

[75] Sanghera, G., et al. (2011). Engineering Stress Tolerance in Crop Plants. Current Genomics 12(1).

[76] Smith, J. (2001). A History of Brazil: 1500 – 2000. Edinburgh: Pearson Education Limited.

[77] St. Paul, G. & Verdier, T. (1993). Education, Democracy, and Growth. Journal of Development Economics 42(1993), pp. 399-407.

[78] Stock, J. & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. Andrews DWK Identification and Inference for Econometric Models. New York: Cambridge University Press

[79] Stromberg, C. (2011). Evolution of Grasses and Grassland Ecosystems. The Annual Review of Earth and Planetary Sciences 2011(39), pp. 517-544.

[80] Topik, S. (1999). “Where is the Coffee? Coffee and Brazilian Identity” Luso-Brazilian Review 36(2), pp. 87-92.

[81] Weinstein, B. (1983). The Amazon Rubber Boom. Stanford: Stanford University Press.

Table 1: Summary Statistics

Variable	**Mean**	**Max**	**Min**	**SD**	**Observations**
Land Gini Index (1920)	.624	.954	.137	.134	938
Plantation Suitability Index	0	2.538	-.456	1	953
Welfare Spending 1923	.472	.923	0	.181	846
Welfare Spending 1995-05	.431	.643	.058	.063	952
Income Gini Index (2000)	.554	.428	.840	.045	953
Land Gini Index (2006)	.750	.943	.112	.092	953
HDI	69.817	88.6	51.76	8.123	953
Foreign Population (per 1000 residents)	37.411	361.684	0	68.721	951
Distance to Coast (km)	159.008	1349.224	0	176.527	942
Malaria	2.356	6.786	0	1.999	944
Suitability for Agriculture	.331	.983	.001	.202	944

Table 2: First Stage Regression of 1920 Land Inequality on PSI

Dependent Variable is 1920 Land Gini					
	(1)	(2)	(3)	(4)	(5)
Plantation Suitability Index (PSI)	2.493***	2.684***	2.814***	3.069***	3.077***
	(0.445)	(0.490)	(0.486)	(0.495)	(0.487)
Malaria		-0.244	-0.399*	-0.443*	0.359
		(0.231)	(0.229)	(0.229)	(0.277)
Log Distance to Coast (km)			1.017***	1.038***	0.760***
			(0.283)	(0.289)	(0.289)
Suitability for Agriculture				5.462**	6.830***
				(2.138)	(2.141)
Foreign Population					1.383***
					(0.291)
Constant	62.43***	62.43***	51.02***	48.98***	48.78***
	(0.430)	(0.430)	(3.266)	(3.518)	(3.359)
Observations	938	938	938	938	937
R-squared	0.035	0.036	0.053	0.059	0.081
F-statistic	31.45	15.96	13.90	11.75	16.63

Table 3: Regressions of Historical Institutions on Historical Inequality

	Dependent Variable is share of Welfare Spending in 1923					
	IV Specifications					OLS Specification
	(1)	(2)	(3)	(4)	(5)	(6)
1920 Land Gini	-1.847***	-1.030***	-0.954***	-0.847***	-0.833***	-0.137***
	(0.429)	(0.332)	(0.301)	(0.274)	(0.269)	(0.0470)
Malaria		-0.0250***	-0.0267***	-0.0268***	-0.0169***	-0.0238***
		(0.00386)	(0.00369)	(0.00357)	(0.00518)	(0.00398)
Log Distance to Coast (km)			0.0140***	0.0132***	0.00966**	0.00590
			(0.00523)	(0.00511)	(0.00479)	(0.00398)
Suitability for Agriculture				0.0636*	0.0872**	0.0558*
				(0.0343)	(0.0368)	(0.0323)
Foreign Pop					0.0171***	0.00703*
					(0.00593)	(0.00407)
Constant	1.629***	1.175***	0.974***	0.896***	0.860***	0.514***
	(0.269)	(0.205)	(0.167)	(0.149)	(0.140)	(0.0511)
Observations	840	840	840	840	839	839

Table 4: The Persistence of Local Institutions and Inequality

	Land Gini 2006	Household Income Gini 2000	Welfare Spending (1995-05 Average)
	(1)	(2)	(3)
1920 Land Gini	0.110***	0.0407***	
	(0.0263)	(0.0113)	
Welfare Spending (1923)			0.0713***
			(0.0122)
Constant	68.09***	52.87***	0.399***
	(1.678)	(0.711)	(0.00632)
Observations	938	938	846
R-squared	0.026	0.015	0.043

Table 5: Regressions of Contemporary Institutions on Historical Inequality

Dependent Variable is 1995-2005 Average Welfare Spending						
	IV Specifications					OLS Specification
	(1)	(2)	(3)	(4)	(5)	(6)
1920 Land Gini	-0.511***	-0.491***	-0.453***	-0.429***	-0.424***	0.00718
	(0.123)	(0.123)	(0.112)	(0.103)	(0.101)	(0.0162)
Malaria		-0.000649	-0.00177	-0.00179	0.00479**	0.000635
		(0.00149)	(0.00141)	(0.00137)	(0.00195)	(0.00129)
Log Distance to Coast (km)			0.00796***	0.00777***	0.00547***	0.00268**
			(0.00228)	(0.00219)	(0.00203)	(0.00135)
Suitability for Agriculture				0.0147	0.0272*	0.0107
				(0.0138)	(0.0146)	(0.0112)
Foreign Pop					0.0114***	0.00556***
					(0.00243)	(0.00143)
Constant	0.750***	0.739***	0.629***	0.611***	0.591***	0.380***
	(0.0770)	(0.0759)	(0.0636)	(0.0592)	(0.0559)	(0.0183)

Table 6: Regressions of Contemporary Development on Historical Inequality

Dependent Variable is 2000 AMC-level HDI						
	IV Specifications					OLS Specification
	(1)	(2)	(3)	(4)	(5)	(6)
1920 Land Gini	-1.169***	-0.315**	-0.264**	-0.354***	-0.326***	-0.00939
	(0.240)	(0.128)	(0.105)	(0.106)	(0.0852)	(0.0114)
Malaria		-0.0274***	-0.0287***	-0.0286***	-0.0124***	-0.0153***
		(0.00120)	(0.00107)	(0.00120)	(0.00158)	(0.00106)
Log Distance to Coast (km)			0.00944***	0.0102***	0.00447***	0.00245**
			(0.00219)	(0.00225)	(0.00172)	(0.00123)
Suitability for Agriculture				-0.0563***	-0.0238**	-0.0359***
				(0.0122)	(0.00993)	(0.00696)
Foreign Pop					0.0284***	0.0242***
					(0.00182)	(0.00102)
Constant	1.428***	0.959***	0.825***	0.891***	0.830***	0.675***
	(0.150)	(0.0787)	(0.0595)	(0.0613)	(0.0470)	(0.0148)
Observations	934	934	934	934	933	933

Table 7: First-stage and IV Regressions Including Region Controls

	OLS (First Stage)	IV (Second Stage)		
	1920 Land Gini	Welfare Spending (1923)	Welfare Spending (1995-2005 Ave.)	HDI
	(1)	(2)	(3)	(4)
PSI	02.38***			
	(0.602)			
1920 Land Gini		0.322	-0.381***	-0.324***
		(0.399)	(0.140)	(0.114)
Malaria	-0.00146	-0.0174***	-0.00257	-0.00252
	(0.00387)	(0.00510)	(0.00228)	(0.00175)
Log Distance to Coast (km)	0.00806**	-0.000320	0.00548**	-0.000414
	(0.00329)	(0.00451)	(0.00217)	(0.00169)
Suitability for Agriculture	0.0586***	0.00803	0.00464	-0.00846
	(0.0214)	(0.0370)	(0.0137)	(0.0105)
Foreign Pop	0.0181***	0.000225	0.0185***	0.0205***
	(0.00328)	(0.00874)	(0.00325)	(0.00273)
North	-0.00307	-0.217***	-0.0239	-0.0832***
	(0.0282)	(0.0451)	(0.0191)	(0.0149)
Northeast	-0.0242	0.000163	0.0156	-0.112***
	(0.0235)	(0.0419)	(0.0155)	(0.0132)
South	-0.0694**	-0.0261	-0.121***	-0.0394**
	(0.0276)	(0.0565)	(0.0216)	(0.0174)
Southeast	-0.0657***	0.0532	-0.0250	-0.0364**
	(0.0213)	(0.0500)	(0.0183)	(0.0160)
Constant	0.524***	0.301	0.589***	0.941***
	(0.0455)	(0.229)	(0.0828)	(0.0668)
Observations	937	839	937	937

Table 8: First Stage and IV Regressions Excluding AMCs in Bottom Quintile of Population in 1920

	OLS (First Stage)	IV Specifications		
	1920 Land Gini	Welfare Spending (1923)	Welfare Spending (1995-2005 Average)	HDI
	(1)	(2)	(3)	(4)
PSI	3.26***			
	(0.549)			
1920 Land Gini		-0.749***	-0.477***	-0.272***
		(0.271)	(0.108)	(0.0683)
Malaria	0.00420	-0.0161***	0.00532**	-0.0126***
	(0.00313)	(0.00561)	(0.00233)	(0.00147)
Log Distance to Coast (km)	0.00693**	0.00732	0.00441**	0.00379***
	(0.00345)	(0.00468)	(0.00215)	(0.00136)
Suitability for Agriculture	0.0556**	0.0584	0.0110	-0.0280***
	(0.0241)	(0.0393)	(0.0166)	(0.0105)
Foreign Pop	0.0146***	0.0175***	0.0126***	0.0291***
	(0.00323)	(0.00644)	(0.00268)	(0.00169)
Constant	0.492***	0.850***	0.642***	0.805***
	(0.0408)	(0.148)	(0.0579)	(0.0366)
Observations	752	682	752	752

Table 9: Regression of Historical Institions on Historical Inequality with a Restricted Sample and Geographic Fixed Effects

	Dependent Variable is 1923 Welfare Spending			
	Full AMC Sample	Excluding AMCs in North	Excluding AMCs in North	Full AMC Sample with Region FE
	(1)	(2)	(3)	(4)
1920 Land Gini	-1.847***	-1.306**	0.125	0.0508
	(0.429)	(0.614)	(0.418)	(0.447)
Malaria			-0.0256***	
			(0.00329)	
North				-0.226***
				(0.0420)
Northeast				-0.0429
				(0.0492)
South				-0.00499
				(0.0517)
Southeast				0.0660
				(0.0475)
Constant	1.629***	1.295***	0.463*	0.441
	(0.269)	(0.381)	(0.260)	(0.312)
Observations	840	799	799	840

Table 10: Regressions of Per-capita Public Welfare Spending on Historical Land Inequality

	Log Per-capita Welfare Spending (1923)		Log Per-capita Welfare Spending (1995-2005 Average)	
	(1)	(2)	(3)	(4)
1920 Land Gini	-5.466*	-4.727	-4.635***	-4.523***
	(2.839)	(4.531)	(0.877)	(1.327)
Malaria	-0.0848	-0.183***	0.00960	0.0308
	(0.0585)	(0.0684)	(0.0169)	(0.0212)
Log Distance to Coast (km)	-0.0644	-0.0286	0.0329*	0.00688
	(0.0535)	(0.0643)	(0.0185)	(0.0202)
Suitability for Agriculture	0.643	0.420	-0.0823	-0.0860
	(0.445)	(0.462)	(0.119)	(0.126)
Foreign Pop.	0.563***	0.638***	0.171***	0.159***
	(0.0681)	(0.106)	(0.0206)	(0.0301)
North		0.528		-0.592***
		(0.723)		(0.164)
Northeast		0.978		-0.571***
		(0.654)		(0.147)
South		0.0778		-0.726***
		(0.782)		(0.199)
Southeast		0.278		-0.352**
		(0.738)		(0.175)
Constant	9.063***	7.801***	7.333***	8.001***
	(1.493)	(2.596)	(0.500)	(0.789)
Observations	937	937	937	937

Table 11: **First-Stage and IV Regressions Including Crop Controls**

	OLS (First-Stage)	IV (Second Stage)		
	1920 Land Gini	Welfare Spending (1923)	Welfare Spending (1995-2005 Average)	HDI
	(1)	(2)	(3)	(4)
PSI	2.72***			
	(0.493)			
1920 Land Gini		-0.665**	-0.405***	-0.350***
		(0.294)	(0.113)	(0.131)
Malaria	0.00385	-0.0183***	0.00463**	-0.00237
	(0.00273)	(0.00500)	(0.00190)	(0.00184)
Log Distance to Coast (km)	0.00563**	0.0106**	0.00600***	-0.000980
	(0.00273)	(0.00458)	(0.00201)	(0.00171)
Suitability for Agriculture	0.0746***	0.107***	0.0455***	-0.00695
	(0.0228)	(0.0356)	(0.0144)	(0.0121)
Foreign Pop.	0.0185***	0.0157**	0.0146***	0.0224***
	(0.00337)	(0.00711)	(0.00298)	(0.00358)
Sugar Production	-0.0266***	0.00600	-0.00772	-0.00762
	(0.00693)	(0.0122)	(0.00481)	(0.00472)
Coffee Production	-0.299***	0.163	-0.106	-0.116**
	(0.0969)	(0.159)	(0.0796)	(0.0566)
Wheat Production	-2.763***	-3.726***	-2.691***	-0.600
	(0.430)	(0.903)	(0.357)	(0.374)
Bean Production	-0.0688	0.0378	-0.0419	-0.00709
	(0.0746)	(0.102)	(0.0366)	(0.0320)
Ranching	0.0526	-0.246***	-0.119***	0.0354
	(0.0431)	(0.0666)	(0.0264)	(0.0263)
Constant	0.507***	0.749***	0.573***	0.959***
	(0.0329)	(0.161)	(0.0651)	(0.0777)
Observations	937	839	937	937

Table 12: Effect of Slavery on Historical Spending, Contemporary Spending, and Development

	Welfare Spending (1923)	Welfare Spending (1995-2005 Average)	HDI
	(1)	(2)	(3)
1920 Land Gini	-0.829***	-0.422***	-0.301***
	(0.264)	(0.102)	(0.0836)
Slave Share (1872)	0.185***	-0.0151	0.0101
	(0.0693)	(0.0296)	(0.0178)
Malaria	-0.0147***	0.00462**	-0.0121***
	(0.00511)	(0.00191)	(0.00148)
Log Distance to Coast (km)	0.0103**	0.00525***	0.00442***
	(0.00470)	(0.00202)	(0.00167)
Suitability for Agriculture	0.0841**	0.0285*	-0.0261***
	(0.0369)	(0.0147)	(0.00988)
Foreign Pop.	0.0149**	0.0113***	0.0282***
	(0.00593)	(0.00247)	(0.00180)
Constant	0.824***	0.594***	0.815***
	(0.144)	(0.0591)	(0.0483)
Observations	829	926	922

Table 13: Effect of Historical Inequality and Urbanization on Contemporary Outcomes

	1995-05 Welfare Spending	HDI
	(1)	(2)
1920 Land Gini	-0.364***	-0.383***
	(0.139)	(0.111)
Urban	0.0465***	-0.201***
	(0.00755)	(0.0670)
Urban*Inequality	0.00375	0.358***
	(0.0326)	(0.111)
Malaria	-0.00257	-0.0126***
	(0.00212)	(0.00145)
Log Distance to Coast (km)	0.00606***	0.00419***
	(0.00213)	(0.00148)
Suitability for Agriculture	0.0103	-0.0442***
	(0.0126)	(0.00900)
Foreign Pop.	0.0130***	0.0243***
	(0.00306)	(0.00135)
Constant	0.541***	0.872***
	(0.0785)	(0.0641)
Observations	937	933

Figure 1: Map of Plantation Suitability Index and 1920 Land Inequality

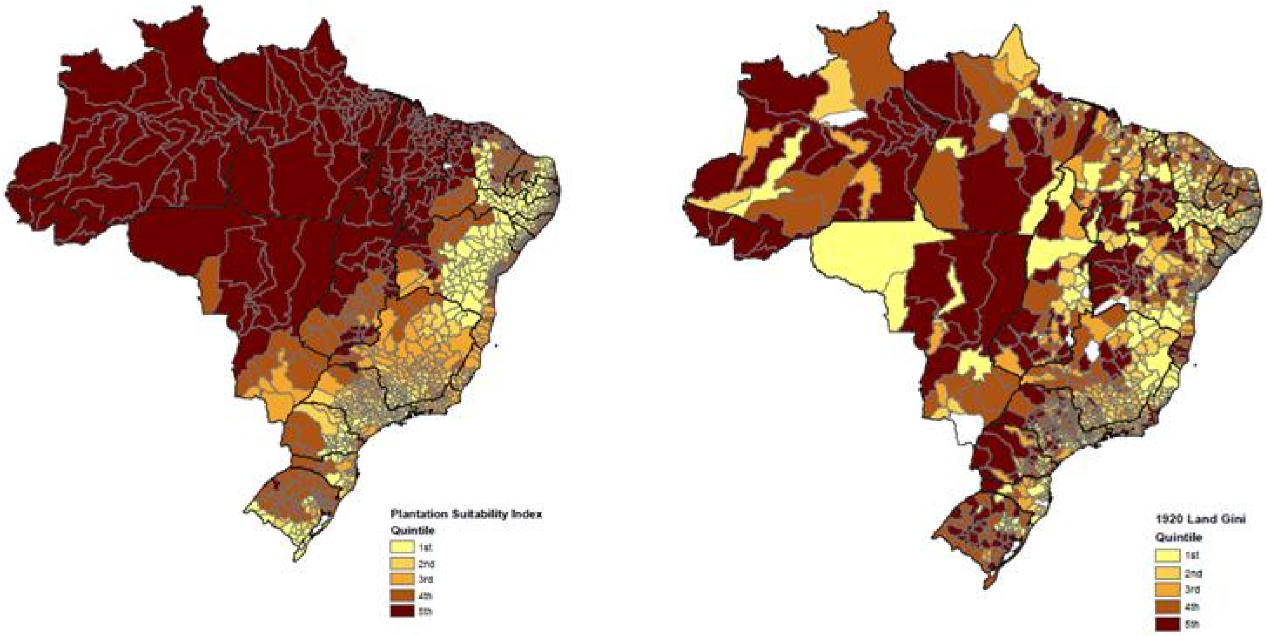
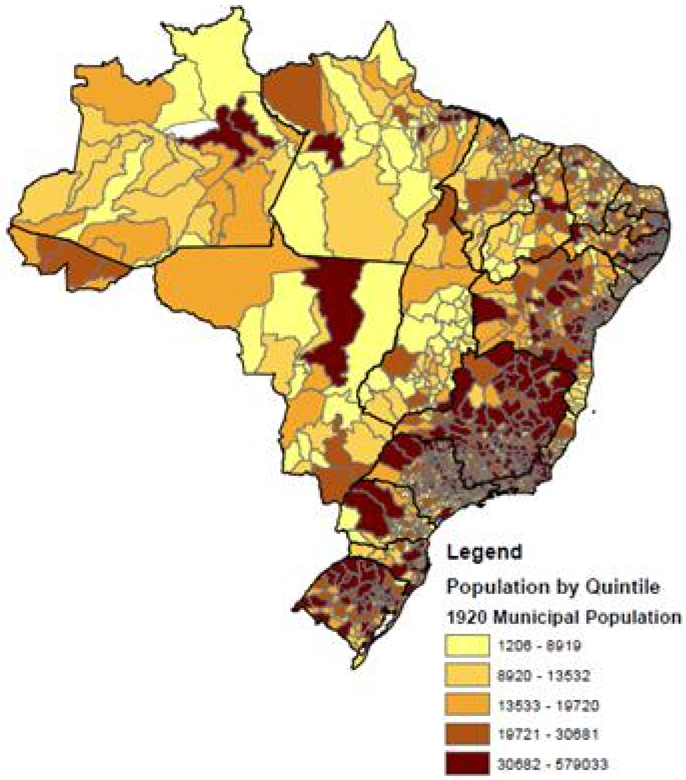


Figure 2: Municipal Population of Brazil by Quintile (1920)



9 Appendix

9.1 Prominent Smallholder Crops in Brazilian Agricultural History

Manioc (cassava) seems to have been the most common small-holder crop in colonial Brazil, in terms of both quantity produced and area under cultivation. As a result of its nutritional value and ability to grow in a variety of soils and climates, it quickly became the colony's primary food crop. It was particularly widespread in settlements throughout the northeast, and was popular along the coast of the southeast, in states such as Espirito Santo, Rio de Janeiro, São Paulo and even Santa Catarina. Other crops, such as maize, were more widespread further inland: The "maize-manioc" dividing line seems to have been on the frontier between Bahia and Minas Gerais. But the importance of manioc should not be underestimated; it was the main foodstuff of plantations, where it was consumed by wealthy land-owners and slaves alike. Many plantations grew manioc in between rows of sugarcane, and slaves grew it on their own personal plots of land, which they would manage on their limited days off (Prado 1967).

Maize proved to be another critical food crop in the colonial era. Unlike manioc, it was much more popular in the south and in settlements further inland. It was especially widespread in Minas Gerais, but was grown widely throughout Rio Grande do Sul, Santa Catarina, Paraná, as well as inland São Paulo. The climatic conditions of the south are generally more favorable to maize production, which may explain its historical regional popularity. Interestingly, maize was also important as feed for mules and donkeys, which were the primary mode of transportation in the hilly southern interior. The popularity of the crop may have partly been due to its convenience for consumption by both people and pack animals (Prado 1967).

Rice, beans, and wheat constituted the next most important food crops. Beans were grown throughout colonial Brazil, and were prevalent in the diet of colonists in Minas Gerais, Rio de Janeiro, São Paulo and Espirito Santo. Rice was both an export and domestically consumed crop. It was grown for export in Maranhão, but also produced in Rio de Janeiro, Pará, and Northern São Paulo. Since rice favors a climate with a higher humidity, it was absent from southern states such as Santa Catarina and Rio Grande do Sul. In these states wheat was the staple crop, with some wheat production present in Minas Gerais and Bahia, where the latter province provided wheat flour to the capital (Prado 1967).

These crops seem to have constituted the major food crops in colonial Brazil. Green vegetables and fruit were almost completely absent from the colonial diet. Orchards were present on only the largest estates, and fruits were considered a luxury item, available only to a small minority of the population (Prado 1967).

9.2 Prominent Plantation Crops in Brazilian History

Sugarcane was perhaps the most widely distributed plantation crop in colonial Brazil. It was grown along the coast from Pará to Santa Catarina. For nearly three hundred years, sugar proved to be the most valuable export crop, until it was surpassed by coffee in the 1830s. The organization of sugarcane production was remarkably similar throughout the colony. It was centered on the Engenho, a term originally referring to the processing plant which turned the harvested cane into sugar, but later came to encompass the entire plantation. At the pinnacle of the organization was the owner. But the land was also cultivated by lavradores (sharecroppers), who rented land from the owner and were obliged to send their cane to the owner's mill for processing. These lavradores typically paid the owner half of their cane crop for use of the mill; rent was further deducted from the lavradores' remaining half. There were also free copyholders, who owned the land they farmed, but still sent their cane to the owner's mill. Nearly all day-to-day labor was provided by slaves. A few free, paid workers served as technicians for the machinery of the mill, or as packers (those who packed the processed sugar into chests for export) (Prado 1967).

Aside from sugar, the most prominent cash crop in Brazilian agriculture was that of coffee. In 1830, coffee overtook sugar to become Brazil's leading export commodity; by 1850, it accounted for half of all Brazil's export earnings (Bethell 1989). Most coffee was produced according to a plantation system of production. The typical coffee fazenda included an owner's mansion, slave (or worker) quarters, and processing sheds (Bethell 1989). The first coffee plantations began in the highlands above the Paraiba river, from Rio de Janeiro, through São Paulo and to Southeast Minas Gerais. Production later spread to other parts of these states, as well as to Espirito Santo.

But unlike sugar, coffee was not always produced in a plantation-style manner. The coffee plantations in São Paulo were very large, and those in Rio de Janeiro, while smaller, still tended to be around 250 to 500 acres on average (Clarence-Smith and Topik, 2003). But production in Minas Gerais was more small scale, particularly in the southeast Mata

Mineira region (Clarence-Smith and Topik, 2003). The scale of production may have been even smaller in the more mountainous southern and central regions of the state, as coffee was primarily consumed domestically in these areas until at least the mid-nineteenth century (Bergad 1999).

Rubber proved to be another important, if relatively short-lived, cash crop that dominated the economy of Northern Brazil in the latter nineteenth and early twentieth centuries. After the discovery of the vulcanization process in 1839, world demand for rubber surged. As the natural habitat for the rubber tree Havea Brasiliensis is in the Amazon, Brazil was virtually the sole supplier of rubber until the 1880s (Weinstein 1983). In rubber growing regions, powerful landowners (patrões) would "lease" land to rubber-tappers (usually native Indians or caboclos, sometimes white Brazilians), who would collect latex from the trees. Unlike plantation production, rubber tappers lived in isolated settlements, and bought exorbitantly priced food, medicine, and equipment from the patrão. This often led to a quasi-feudal system of indebtedness and dependency, with abuse of rubber-tappers generally increasing in more isolated areas. Although demand for rubber slumped after WWI and the crop gradually decreased in its economic importance, it had a substantial impact on the economies of the remoter Brazilian states, such as Amazonas, Acre, Pará and northern Mato Grosso (Weinstein 1983).

Cotton was yet another important cash crop, with widespread production in the early to mid-eighteenth century. Cotton production was particularly common in Maranhão, but later spread to Pernambuco, Bahia, and Rio de Janeiro. At its peak, it was also produced in Paraná and Goiás (Prado 1967). The "cotton boom" was not nearly as long-lived as that of sugar; increased production from North America led to a fall in world prices, which damaged the profitability of the crop. Nevertheless, cotton provided 20% of Brazil's export earnings at the time of independence (second only to sugar), and was still the third most valuable export in 1850 (Bethell 1989). The method of cotton production was apparently quite varied. In some regions it was produced according to a plantation system similar to that of sugar

(Prado 1967), while in others in was more of a small-holder crop (Bethell 1989). Different methods of production may have arisen since the crop requires less capital investment than sugar or coffee, and was also grown farther inland, away from the established plantations (for example, it was a popular crop around the sertão region between Bahia and Minas Gerais) (Bethell 1989).

A relatively minor colonial-era cash crop was that of tobacco, accounting for only 3% of Brazil's export earnings between 1850 and 1870 (Bethell 1989). Its method of production varied even more than that of cotton: While it was produced on large-scale plantations, a significant amount was also grown on small-scale farms. The plant is relatively delicate and must be shielded from excessive heat and a multitude of pests; therefore, the care required may have reduced the advantages of large-scale production (Bethell 1989). Like cotton and sugarcane, it was produced widely throughout the colony, with the main tobacco growing areas concentrated in Bahia, Sergipe, and Rio de Janeiro (Prado 1967).

9.3 Index Foundations

In what follows, I will present a method for calculating the plantation suitability index (PSI). To construct the index, I draw heavily from information contained in FAO databases. This point is a relatively important one, since there is a certain degree of subjectivity in deciding the optimal growing conditions for various crops. FAO data is therefore useful since it draws from the opinions of agronomic experts to yield a consistent set of optimal and non-optimal growing conditions for a variety of economically important crops. The FAO uses these parameters to construct their own suitability indices, which have recently been employed in an economic context by Nunn and Qian (2011) as well as by Easterly (2007).

Some aspects of the PSI are similar to suitability indices developed by Ramirez-Villegas, Jarvis, and Laderach (2013). Their approach is quite useful in this setting, since their index is designed to be used with the FAO parameters described above. It is also based on precipitation and temperature conditions only, and does not account for soil properties. The functional form of the index will differ from theirs, however, since the index is not based on one particular plant, but is instead meant to quantify the relative suitability of the climate for plantation over smallholder agriculture. In what follows, I will describe the key elements of the PSI.

In particular, the index depends on a set of parameters, which are specific for each crop. The temperature parameters include $T_{\text{kill min}}$, the minimum temperature at which the crop is killed; T_{min} and T_{max}, the minimum and maximum temperatures at which the plant can grow; and T_{opmin} and T_{opmax}, the minimum and maximum optimum temperature for the plant. Similar parameters are needed for rainfall as well: R_{min} and R_{max}; rainfall below/above which the plant cannot grow, and $R_{\text{opmin}}/R_{\text{opmax}}$; the minimum/maximum optimum rainfall for the crop. It is also necessary to determine the length of a crop's growing season (G_{ave}). For each municipality p, twelve temperature suitability indices are calculated; That is, it is assumed that each of the 12 months of the year could be the start of a growing season. All these parameter values can be obtained from the FAO, which lists them for the most common crops. Below I will show how I aggregate these parameters to calculate the PSI.

9.4 Derivation of the Optimal Conditions for the Synthetic Crops

I will begin by deriving optimal conditions for the synthetic small-holder ("SS") crop. Recall the SS crop is based on two plant taxonomies *Pooideae* and *Phaseolus*, and that these groups in turn consist of multiple species of food crops. The minimum optimum ("opmin") and maximum optimum ("opmax") values for both temperature and precipitation are given below; these variables apply to individual species of crop (for example, wheat or barley). Formally, I choose only the optimum range of temperatures and precipitations common to all crops within the Pooideae subfamily or Phaseolus genus as the optimum range each group. This gives an optimum temperature and precipitation range for *Pooideae* and *Phaseolus*:

That is, a T^{po}_{opmin}, T^{po}_{opmax}, R^{po}_{opmin}, R^{op}_{max}, where superscript "po" refers to *Pooidea*, and T^{ph}_{opmin}, T^{ph}_{opmax}, R^{ph}_{opmin}, R^{ph}_{opmax}, where superscript "ph" refers to *Phaseolus*.

Intuitively, the above procedure simply entails choosing the maximum of the minimum optimum temperature (or precipitation) values for all crops in each subfamily; similarly, the minimum of the maximum optimum temperatures (or precipitation) values is chosen. For example, using the actual values for four individual crops, T^{po}_{opmin} and T^{po}_{opmax} would take the following form:

T^{po}_{opmin}= Max [15, 15, 15, 16] = 16

T^{po}_{opmax}= Min [23, 20, 20, 20] = 20

Thus, the optimal range of temperatures for the Pooideae subfamily is 16-20C. This same procedure would apply for deriving the optimum values for rainfall as well. Using this methodology, the following values are obtained for Pooideae:

T^{po}_{opmin}=16C

T^{po}_{opmax}=20C

R^{po}_{opmin}=247mm

R^{op}_{max}= 296mm

And these values are obtained for the Phaseolus subfamily:

T^{ph}_{opmin} = 20C

T^{ph}_{opmax} = 25C

R^{ph}_{opmin} = 263mm

R^{ph}_{opmax} = 329mm

These two sets of numbers must be combined in order to obtain the optimum conditions for the SS crop. To do this I make the assumption that both groups are "equally smallholder"; that is, crops in both the *Pooideae* and *Phaseolus* subfamilies are generally produced on a similar small scale. Such an assumption allows the index to be calculated with considerable ease. In particular, I choose the minimum optimum temperature for the SS crop (T^{SS}_{opmin}) as the minimum of T^{po}_{opmin} and T^{ph}_{opmin}; likewise T^{SS}_{opmax} is calculated as the maximum of T^{po}_{opmax} and T^{ph}_{opmax}. Thus, for example, if the temperature is slightly too warm to optimally grow a grain from *Pooideae*, it will be replaced with a crop from *Phaseolus*, and this *Phaseolus* crop will be grown on an equally small scale as the *Pooideae* crop would have been.

The exact same methodology is used to calculate the range of optimum precipitation for the SS crop. This methodology yields the following values:

$T_{SS_{opmin}} = 16$

$T^{SS}_{opmax} = 25$

$R^{SS}_{opmin} = 247mm$

$R^{SS}_{opmax} = 329mm$

The same procedure is then used to calculate the optimum conditions for the synthetic plantation ("SP") crop. Using the methodology above yields the following:

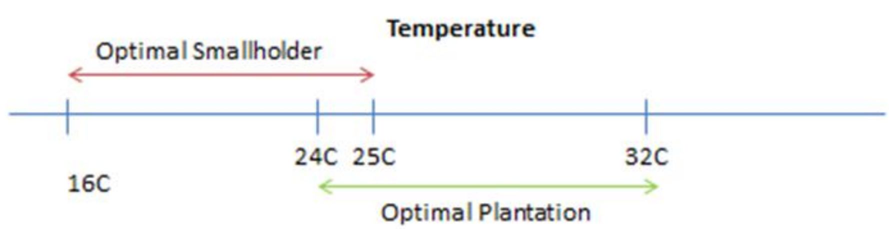
$T^{SP}_{opmin} = 24C$

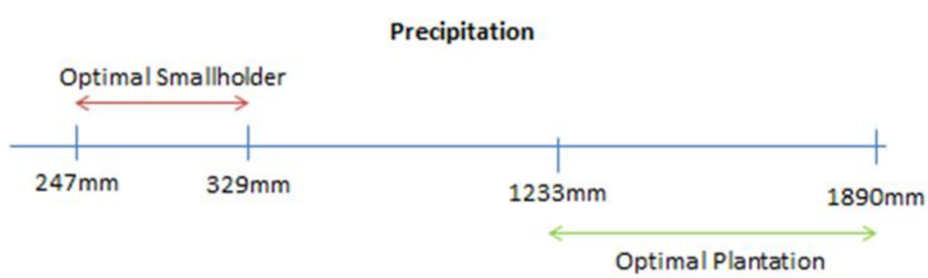
$T^{SP}_{opmax} = 32C$

$R^{SP}_{opmin} = 1233mm$

$R^{SP}_{opmax} = 1890mm$

The graphs below offer a visualization that helps put these values in context:





Much in the same way that optimal temperatures and precipitation were calculated for the synthetic plantation and small-holder crops, an optimal growing season can be derived as well. The FAO lists the minimum and maximum growing seasons needed for a variety of crops. The same procedure described above can be used to assign a growing season to the SS and SP crops: That is, the maximum of the minimum growing seasons for all crops constitutes ${G^{SS}}_{min}$ /${G^{SP}}_{min}$; the minimum of all maximum growing seasons will constitute ${G^{SS}}_{max}$/ ${G^{SP}}_{max}$. Using this methodology, the growing season for the SS and SP crops would be 120 days (4 months) and 300 days (10 months), respectively.

Calculating the optimum conditions for a synthetic smallholder and plantation crop in the manner described above has both advantages and drawbacks. On the one hand, it is quite flexible in that other crops (or their subfamily and genera) can easily be incorporated into the model. Similarly, it does not put too much emphasis on any one crop; therefore, excluding some crops should not greatly alter the index. And as stated previously, it is based on parameters calculated by the FAO, which is consistent with other literature.

Nevertheless, in relying on subfamilies of plants, the procedure is in effect based on some crops that were never grown in Brazil (barley, for example). It also requires making some assumptions about the nature of small-holder and plantation production between crops of different subfamilies and genera (for example, we must assume that some crops, such as wheat and beans, are produced on an equally small scale).

Two points can be raised in response to these criticisms. First, it may be worthwhile to construct an index with a certain element of external validity. Basing the index entirely on crops from Brazil would make the index, in effect, country-specific. And while a country specific index is fine in this study, it may be worthwhile to construct a measure applicable to other countries and contexts, if it is possible to do so without sacrificing the internal validity of the index for this research. Second, given that individual crops are fairly similar within subfamilies and genera, the inclusion of non-Brazilian crops does not significantly alter the index.

9.5 Derivation of the Non-Optimal Conditions for the Synthetic Crops

The discussion above has focused on identifying the optimum conditions for a synthetic smallholder and plantation crop. Yet it is also necessary to identify the range of conditions under which these crops might grow. That is, it is also necessary to derive non-optimal conditions for these synthetic crops, as well as to identify the bounds beyond which these plants would not survive.

The underlying rationale for constructing this suitability index is that climatic conditions in part determine whether efficiency in agriculture lies in smallholder or plantation production. The "ideal" smallholder crops presented above (those belonging to the Pooideae and Phaseolus subfamilies) all thrive with lower rainfall and generally lower temperatures. Thus, at very cold temperatures and/or low levels of precipitation, the only agricultural activity that can take place, if any can take place at all, will be small-holder. Note that even though small-holder farming may be the only viable agricultural activity, this does not necessarily mean smallholder farmers will dominate the economy; instead, other non-agricultural livelihoods may develop (this will be discussed below). A similar story applies to those regions with high rainfall and high temperatures; plantation crops (especially coffee) can withstand exceptional amounts of rainfall, much more than many food crops; here large-scale agriculture may be unambiguously more efficient.

Identifying a lower bound on the temperature and precipitation requirements for the synthetic smallholder crop, and an upper bound for the synthetic plantation crop, can be done in a straightforward manner. In particular, the minimum temperature and precipitation for the synthetic smallholder crop can be calculated as the minimum of the minimum requirements for all crops in the *Phaseolus* and *Pooideae* groups. Using this methodology, the minimum rainfall requirement for the SS crop, $\mathrm{R^{SS}}_{\mathrm{min}}$, would be 66mm and the minimum temperature requirement, $\mathrm{T^{SS}}_{\mathrm{min}}$, would be 2C. More generally, for individual smallholder crops a through z, $R^{SS}{}_{min}$ and $T^{SS}{}_{min}$ can be calculated as:

$R^{SS}{}_{min}$= Min $[R^{a}{}_{min}, \ldots, R^{z}{}_{min}]$

$T^{SS}{}_{min}$= Min $[T^{a}{}_{min}, \ldots, T^{z}{}_{min}]$

Identifying the upper bound on temperature and rainfall for the synthetic plantation crop can be done in an analogous manner. The maximum temperature for the SP crop $T^{SP}{}_{max}$, and the maximum rainfall for the plantation crop $R^{SP}{}_{max}$, can be calculated as the maximum of all individual crops in the *Caffea* and *Saccharum* genera. This would yield $T^{SP}{}_{max}$ = 41C, and $R^{SP}{}_{max}$= 4110mm. The general method for calculating these two parameters would be similar to the above; that is, for plantation crops a through z:

$R^{SP}{}_{max}$= Max $[R^{a}{}_{max}, \ldots, R^{z}{}_{max}]$

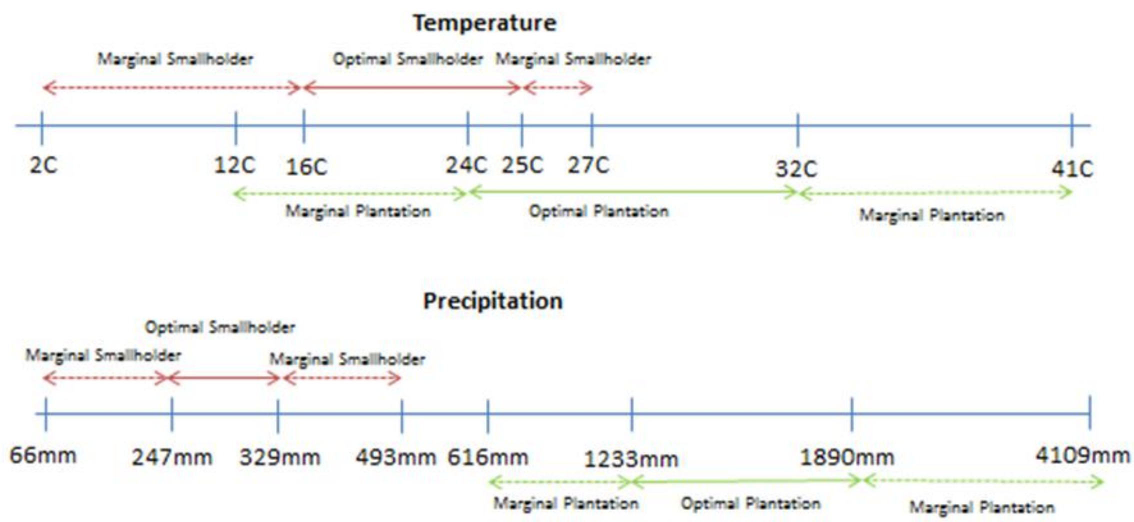
$T^{SP}{}_{max}$= Min $[T^{a}{}_{max}, \ldots, T^{z}{}_{max}]$

Calculating the maximum temperature and rainfall values for the SS crop, and minimum values for the SP crop is more difficult. In essence, this involves a determination as to when the climate shifts from offering efficiency in smallholder agriculture to efficiency in plantation agriculture. That is, as one moves away from the "smallholder" end of the spectrum, conditions become more and more suitable for plantation agriculture. When some maximum values for temperature and precipitation for the SS crop are surpassed, then plantation production will be the most suitable form of agricultural activity. An analogous argument can be made when one moves away from the plantation end of the spectrum.

I therefore calculate $R^{SS}{}_{max}$, $T^{SS}{}_{max}$, $R^{SP}{}_{min}$, and $T^{SP}{}_{min}$ in the following manner. To obtain values for $R^{SS}{}_{max}$ and $T^{SS}{}_{max}$, I first calculate a maximum precipitation and temperature for both the *Pooideae* and *Phaseolus* crop groups. I set the maximum temperature for the *Pooideae* group ($T^{po}{}_{max}$) to the minimum T_{max} for all crops within this group; the same methodology would apply for the *Phaseolus* group ($T^{ph}{}_{max}$). TSSmax is then set equal to the minimum of $[T^{po}{}_{max}, T^{ph}{}_{max}]$. Similarly, I obtain an $R^{po}{}_{max}$ and an $R^{ph}{}_{max}$ by choosing the minimum R_{max} in each group. I then take $R^{SS}{}_{max}$ to be the minimum of $[R^{ph}{}_{max}, R^{po}{}_{max}]$. This yields $R^{SS}{}_{max}$ = 493mm, and $T^{SS}{}_{max}$ = 27C.

A similar method can be used to calculate $R^{SP}{}_{min}$ and $T^{SP}{}_{min}$. I first define a minimum temperature for both the *Caffea* and *Saccharum* genera ($T^{c}{}_{min}$ and $T^{s}{}_{min}$, respectively), by choosing the minimum T_{min} within each genus. I then set $T^{SP}{}_{min}$ equal to the maximum of [$T^{c}{}_{min}$, $T^{s}{}_{min}$]. I identify an $R^{c}{}_{min}$and $R^{s}{}_{min}$ by choosing the minimum R_{min} within each genus of crops; I then set $R^{SP}{}_{min}$ equal to the maximum of {$R^{c}{}_{min}$ and $R^{s}{}_{min}$}. This yields $R^{SP}{}_{min}$= 616mm and $T^{SP}{}_{min}$ = 12C.

Again, a graph offers a visualization that may help illuminate the conditions derived above:



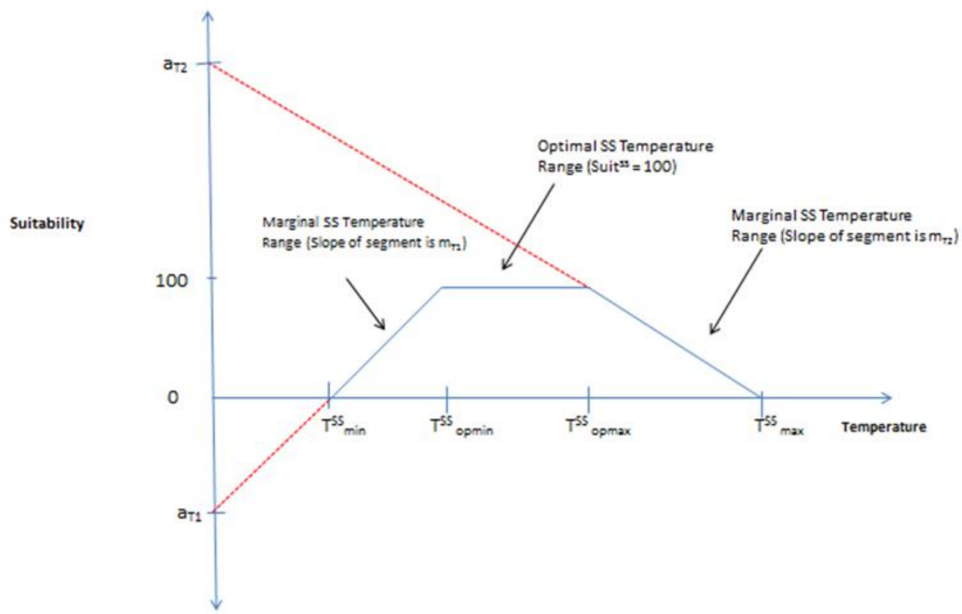
9.6 Mapping Environmental Conditions into a Spectrum

A discussion of the index is incomplete without presenting a method for mapping these temperature and precipitation conditions into an index. The first step is to calculate the average length of the growing season for the SS and SP crop. Recall that temperature and precipitation data is available at a monthly level. For a municipality to be suitable for growing a certain crop it may not be necessary for climatic conditions to be suitable all year around; they only need to be suitable for the length of time it takes to grow and harvest the given agricultural product. In this case the growing season would correspond to the SS and SP crops. For example, the values for $G^{SS}{}_{min}$ and $G^{SS}{}_{max}$ were derived above; the average length of the growing season for the SS crop would simply correspond to the average of these two numbers.

Separate indices are calculated for temperature and precipitation. Each month of the year is treated as the potential start of the growing season. Therefore, for each municipality, twelve temperature suitability indices are calculated, one for each month i. For each month i of municipality p, the temperature suitability index for the synthetic smallholder crop would be calculated as:

$T^{SS}_{suit,\,i} =$	0	$T_{average\ p,i} < T^{SS}_{min}$
	$a_{T1} + m_{T1} * T_{average\ p,i}$	$T^{SS}_{min} \leq T_{average\ p,i} < T^{SS}_{opmin}$
	100	$T^{SS}_{opmin} \leq T_{average\ p,i} < T^{SS}_{opmax}$
	$a_{T2} + m_{T2} * T_{average\ p,i}$	$T^{SS}_{opmax} \leq T_{average\ p,i} < T^{SS}_{max}$
	0	$T^{SS}_{max} \leq T_{average\ p,i}$

Where $T_{average\ p,i}$ is the average temperature in in municipality p during month i. aT1 and mT1 are the intercept and slope of the regression line between $[T^{SS}_{min}, 0]$ and $[T^{SS}_{opmin}, 100]$. Similarly, aT2 and mT2 are the intercept and slope of the regression line between $[T^{SS}_{opmax}, 100]$ and $[T^{SS}_{max}, 0]$. These piecewise linear portions are used to interpolate the suitability of observed temperatures that fall between T^{SS}_{min} and T^{SS}_{opmin}, or between T^{SS}_{opmax} and T^{SS}_{max}. The graph below illustrates possible values of the index given the key temperature parameters described above. Note that the shape of the graph (the blue line) is entirely determined by these parameters; that is, the structure of the index is unrelated to any observed municipal temperatures. It is instead based on the optimal and non-optimal temperature and precipitation requirements of the synthetic smallholder and plantation crops.



The precipitation suitability index is calculated using the total rainfall during a crop's growing season. There are twelve candidate growing seasons, since each month is treated as if it could be the start of the season. Thus I calculate twelve precipitation suitability indices according to the table shown below.

$R^{SS}{}_{suit}$ =	0	$R_{total,\ p} < R^{SS}{}_{min}$
	$a_{r1} + m_{r1}{}^{*}R_{total,\ p}$	$R^{SS}{}_{min} \leq R_{total,\ p} < R_{SS\,opmin}$
	100	$R_{SS\,opmin} \leq R_{total,\ p} < R^{SS}{}_{opmax}$
	$a_{r2} + m_{r2}{}^{*}R_{total,\ p}$	$R^{SS}{}_{opmax} \leq R_{total,\ p} \leq R^{SS}{}_{max}$
	0	$R_{total,p} \geq R^{SS}{}_{max}$

Where $R_{total,\ p}$ is the total amount of precipitation in a growing season for the synthetic smallholder plant. a_{r1} and m_{r1} are the intercept and slope of the regression line between [$R^{SS}{}_{min}$, 0] and [$R_{SS\,opmin}$, 100] and a_{r2} and m_{r2} are the intercept and slope of a regression curve between [R^{SS}opmax, 100] and [$R^{SS}{}_{max}$, 0].

There are thus twelve temperature suitability indices (one for each month) and twelve precipitation suitability indices (one for each potential growing season). To obtain temperature indices on the basis of a growing season, I take each month of the year as the potential start of the season, and choose the lowest monthly suitability index within that season as the season's temperature suitability index. The lowest monthly index is chosen so as to be consistent with "Liebig's Law", or the principle that plant development will be contingent on the limiting factor for its growth (a common assumption in the agronomic literature). For each potential growing season g, a suitability index for the synthetic smallholder crop is calculated as $Suit_g^{SS} = T_g^{SS}suit*R_g^{SS}suit$. The final suitability index for the SS crop is taken to be the maximum of these twelve suitability indices. The suitability index for the SP crop would be calculated in the same manner.

The calculations above all utilize data on average monthly temperatures and precipitations. Unfortunately, such data omits useful information regarding climatic variation. This is a particularly important point in regard to temperature, since regions with temperatures that regularly exceed (or fall below) a crop's T_{max} (or T_{min}) are unlikely to be amenable for growing that plant (even if, on average, temperatures are close to the optimum). Such fluctuations will cause the plant to experience a heat or cold shock, in which growth stops and yields are considerably diminished. If these extremes are reached frequently enough, it would likely alter the relative suitability between plantation or smallholder agriculture.

To account for this, I utilize data on average monthly minimum and maximum temperatures; that is, the mean daily high (or low) temperature for a given month. I integrate this information into the index in the following manner. Recall that T_{max} is defined as the minimum of [$T^{po}{}_{max}$, $T^{ph}{}_{max}$]. In this case, I set a given month's temperature suitability index for the SS crop equal to zero if the mean daily maximum temperature exceeds the maximum of [$T^{po}{}_{max}$, $T^{ph}{}_{max}$], which is 30C. Similarly, TSPmin is defined as the maximum of [$T^{c}{}_{min}$, $T^{s}{}_{min}$], so I set a month's temperature suitability index for the SP crop equal to zero if the mean daily minimum temperature is below the minimum of [$T^{c}{}_{min}$, $T^{s}{}_{min}$], which would be 9C. In essence, I assume that these temperatures are thresholds which, if exceeded often enough, will cause the SS or SP crops to become unviable. Note that I do not calculate a threshold minimum temperature for the SS crop, or maximum temperature for the SP crop, since temperatures do not reach these extremes in Brazil.

Finally, to obtain the PSI I take the ratio of the indices calculated above. That is:

$$PSI = \frac{1 + Suit^{SP}}{1 + Suit^{SS}}$$

This functional form is similar to that of Easterly (2007). I use a ratio of the indices since the hypothesis is that the *relative* suitability of the climate for plantation agriculture is the key determinant of inequality. While most regions are far more suited to one type of agriculture over the other, a few areas are suitable for both. Brazil's dry season, combined with the short growing times for smallholder crops, mean that many tropical areas which are otherwise well-suited for plantation agriculture have short periods of cool and dry weather, and hence become amenable to smallholder production. In such cases, the index will become approximately equal to one. Interestingly, when the index is between .95 and 1.05 (which occurs in 32 AMCs), the average land Gini is .61, fairly close to its country-wide mean value of .62.

9.7 Index Data

To calculate the suitability index, I use data compiled by Hijmans et al (2005). This dataset contains information on mean monthly precipitation and temperature, as well as monthly average maximum and average minimum temperature, on a global grid of thirty arc-second resolution (roughly 1 km by 1 km at the equator). The data is drawn from major climate databases compiled by the Global Historical Climatology Network, the Food and Agriculture Organization, the World Meteorological Organization, the International Center for Tropical Agriculture, R-Hydronet, and more minor country-level databases. The authors only use records for which there are at least ten years of data, and the grids are generally representative of climate from the years 1950-2000. The final dataset incorporates precipitation records from 47, 554 locations, and mean temperature records from 24,542 locations. The authors then use other geographic data, along with a smoothing algorithm, to interpolate these temperature and precipitation values to other grid cells. This data has previously been used in the economics literature by Gennaioli et al. (2013) and Dell, Jones, and Olken (2009).

9.8 Note on late-19th Century & Early 20th Century Immigration to Brazil

Overwhelming evidence suggests that immigration policies were driven by the economic interests of the agrarian elite; furthermore, the effects of this immigration boom are consistent with factor endowments shaping long-run development. In particular, the immigration policies of the late Empire and First Republic sought to subsidize the elites' costs of production through cheap immigrant labor. The Land Law of 1850 was explicit in codifying the social and economic status of the agrarian oligarchy: The law banned the acquisition of land by any means except cash purchase, and increased the valuation of land above its market value so as to limit the ability of immigrants and other small-holders to acquire it for themselves. But the law also dedicated some of the profit from land sales to programs that subsidized European immigration. Importantly, this law was not enacted in isolation, but at a time when the institution of slavery was widely recognized to be in terminal decline (DeWitt 2002). Subsidizing the immigrants' cost of passage to Brazil in return for a number of years of labor was thus a response to the anticipated end of slavery and the availability of a low-cost agricultural workforce.

Slavery and subsidized immigration were thus different means to a similar end: Both served to preserve the economic power of the agrarian elite. Furthermore, the working conditions faced by immigrants, while not as bad as those of slaves, could still be harsh: Plantation owners demanded total servility from their immigrant laborers, and subjected them to imprisonment for non-fulfillment of their contracts. Many landowners resorted to hired gunmen to ensure no revolts or insurrection took place among the immigrant workers (Bethell 1989). The Prinetti decree, issued by the Italian government in 1902, briefly prohibited Italian citizens from accepting subsidized passage to Brazil due to working conditions on the coffee plantations of São Paulo. Yet overall the policy of subsidized immigration was remarkably successful from the point of view of the landed elite; estimates suggest wages for agricultural workers in the coffee producing Southeast did not rise between 1870 and 1914 (Hall 1969), and immigrants were apparently cheaper to employ than the ex-slave share-croppers that dominated in the Northeast (Bethell 1989).

Immigrants also spent only a limited amount of time in forced labor, before earning the freedom to pursue other livelihoods. Often, this involved relocation farther south, to states such as Rio Grande do Sul, and a transition to independent small-holder farming (Bethell 1989). Yet here too geography played a critical role in determining the structure of the local economy. Whereas the plantation economies farther north specialized in cash crops for export, immigrant farmers in the South and in rural São Paulo specialized in food crops for domestic consumption (wheat and rice, as well as beef). Furthermore, these small-holder farmers did not seek economic autarky and self-sufficiency like plantations in the Northeast (Bethell 1989). This led to a specialization and division of labor which encouraged the formation of non-agricultural employment. The economic and social impact of immigration to Southern Brazil thus appears to have been shaped by the South's geography; or more specifically, by the suitability of the land for small-holder agriculture.

But more importantly, there is a recognizable continuity in the history of southern Brazil (Paraná, Santa Catarina and Rio Grande do Sul); That is, the social and economic structure of the south in the early 20th century remained similar to that which manifested hundreds of years before. In the colonial south, "the general poverty of the regions, its small European population, and the need for military cooperation against hostile Indian tribes tended to level social differences among the Europeans" (Bethell 1987, pp. 112). Although European institutions were always present, they were limited by the general isolation of the region.

Thus the colonial south was primarily a simple pastoral society in which cattle-raising, limited farming, hunting, and smuggling constituted the major economic activities (Bethell 1987). The continuity between the colonial and early 20th century can be characterized by the general absence of a monoculture plantation-style economic and social system, in favor of a more independent, small-holder one. Again, for the purposes of this analysis, the important thing is that the social and economic system of the south was shaped by factor endowments, and immigrants conformed to this system rather than upset it.