DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 10-04

The Economic Geography of Race in the New World: Brazil, 1500-2000

Justin Bucciferro
University of Colorado at Boulder

October 2010

Department of Economics

University of Colorado at Boulder
Boulder, Colorado 80309

The Economic Geography of Race in the New World: Brazil, 1500-2000[†]

Justin Bucciferro[‡]
University of Colorado at Boulder

October 20, 2010

Abstract

This paper examines the economic factors behind the geography of race in the Americas. It hopes to resolve the apparent paradox that many areas once occupied by Natives are now inhabited by peoples of predominantly European and African descent. A simple model is presented in which the racial composition of the labor force depends on the feasibility of slavery (factor endowments) and the relative cost of Native and African slaves (a function of the ratio of the distance to the frontier to the slave ports in Africa). The predictions of the model are tested using a newly-created database on the racial composition of twenty Brazilian states from 1500 to 2000. The (inverse) labor cost ratio is found to be positively related to the ratio of Africans to Natives, controlling for factor endowments. The results suggest that for the average state, a 1% increase in the cost of Native labor (a 4.6 km shift in the frontier) corresponded to a 2.6% increase in the ratio of Africans to Natives (an additional 18,114 Africans), all else equal.

Keywords: Colonization, Race, Economic Development, Native Americans, Ethnic Diversity, Inequality, Economic Geography, Historical Demography, Institutions, Brazil.

[†] I am especially grateful to Murat Iyigun (my advisor), Ann Carlos, Carol Shiue, Robert McNown, Lee Alston, Francisco Barbosa, and Frank Witmer for their feedback and support. I also thank Barbara Buttenfield, David Jacks, Angela Redish, John Wallis, and Ran Abramitzky for valuable comments. Any errors are my own.

[‡] Department of Economics, 256 UCB, Boulder, CO 80309; Justin.Bucciferro@Colorado.EDU.

I. Introduction

Estimates of the Native American population of the western hemisphere circa 1500 range from 8.4 to 100 million people (Ubelaker, 1992). Over five hundred years later, some countries are of predominantly Amerindian descent, e.g. Bolivia, Ecuador, while others are of mostly European, e.g. Argentina, United States, or African descent, e.g. Haiti, Jamaica. Although intermixture has made the concept of race ambiguous, people continue to identify themselves within a racial category in the census, and race has important economic and political implications.[1] If the Americas were inhabited by Natives five hundred years ago, what can explain the variation in racial composition observed across countries today? Why would some countries that were once entirely 'Native American' now be primarily 'African' or 'European'?

One hypothesis based on the Economics literature suggests that factor endowments – in particular, the suitability of land for producing staple crops (like sugar or coffee), or rich mineral deposits (like silver or gold) – determined the use of coerced labor during the colonial period (Sokoloff and Engerman, 2000). A competitive international market in African slaves allocated labor to the most productive regions; in a few areas, the initial Native population was large enough to meet the demand for labor. This theory suggests that enslaved Native Americans and Africans would comprise a larger proportion of an economy more reliant on export agriculture or mining, and the share of Natives in the labor force would be related to their initial population densities in an area.

The consensus in the History literature is that, initially, indigenous labor was exclusively used; first through barter, then enslavement. In areas with smaller initial populations, the combined effect of forced labor, warfare, and disease often led to extinction. Such was the case in the Caribbean, where the Taino, the aboriginal group that populated the Antilles, were almost extinct fifty years after the arrival of Columbus. In the rest of

[1] Webster's Dictionary (2004) defines race as "a group of persons related by common descent or heredity" and ethnicity as "ethnic (pertaining to or characteristic of a people, esp. a group sharing a common and distinctive culture, religion, language, etc.) traits, background, allegiance, or association." Although the terms are often used interchangeably, for this study the term 'race' is more appropriate.

Spanish America, the pre-conquest populations were much larger and more resistant to European diseases (Symcox and Sullivan, 2005). In Brazil, one estimate suggests that the Native population decreased by two-thirds in only forty years of permanent settlement (Marcílio, 1984). On the Atlantic coast of South America, which includes Brazil, and in the Caribbean, African slaves took the place of the Natives on the sugarcane plantations and in the mines, giving these regions a strong African influence (Stein and Stein, 1970).

This paper provides a simple model to explain how the racial composition of countries in the Americas was and is determined, to some extent, by their economic organization during the colonial period. The hypothesis is that in areas with a high marginal productivity of labor, slavery was common; where slavery prevailed, the racial composition of the labor force depended on the relative cost of African and Native labor. In areas where labor was less productive, alternative labor arrangements, such as debt servitude, were adopted (Grubb, 2000). The predictions of the model are tested using historical data on the racial composition of Brazilian states from 1500 to 2000. The results suggest that the racial composition of Brazilian states over time was closely related to the relative cost of African and Native labor; as the frontier shifted into the interior, Native labor became scarce in coastal cities and African labor predominated. This relationship is robust to changes in the price of output on the world market, to shifts in the legal status of slavery in Brazil, and a range of other relevant variables.

Racial diversity has important implications for overall development, and since this paper explores the circumstances under which different racial groups became established in the Americas, it will further understanding of the Americas' roots of racial geography.[2, 3] A

---

[2] The racial composition of society from the colonial period to the present has an impact on development, including the level and growth rate of income per-capita, the quality of institutions, and the provision of public services. Using contemporary cross-country data, Easterly and Levine (1997) conclude that ethnic diversity slows economic growth by making it difficult to agree on public-good provision and to adopt pro-growth policies (measures include school attainment, financial depth, the black market premium, and the number of telephones). Alesina, Baqir, and Easterly (1999) find that across U.S. cities, metropolitan areas, and urban counties, racial diversity is associated with less spending on education, roads, and sanitation (but more on police). Miguel and Gugerty (2005) discover a similar relationship between ethnic fragmentation and funding for public education and well maintenance in rural western Kenya. Finally, Alesina and LaFerrara (2005) argue that there are benefits (productivity gains) and costs (low public goods) of diversity; however, the net effect is positive only at relatively high levels of per-capita income.

simple theoretical model is provided based on the hypothesis that 1. Factor endowments determined the demand for coerced labor, and 2. African labor substituted for Native American labor depending on the initial Native population density, its rate of decline, and the cost of African slaves. The model is tested using data on the racial composition of Brazil from 1500 to 2000; the compilation of this dataset from historical sources is a separate contribution of this work. Please note that although the model suggests that owning slaves may have been rational, slavery is certainly an abhorrent institution.

The rest of this paper will proceed as follows: In Section II, the context in which the Americas were settled will be described; in Section III, the two main economic sectors in colonial Brazil – gold mining and sugarcane production – will be explored; in Section IV, a theoretical model will be presented; in Section V, the sources of data and methodology of constructing the demographic database will be detailed; in Section VI, the estimation strategy and empirical results will be discussed; in Section VII, the robustness of the model to alternative assumptions about the Native population will be tested; and, finally, in Section VIII, conclusions and directions for future research will be suggested.

II. Historical Background

The discovery of the Americas occurred during a period of European overseas expansion. Prior to the discovery of the Americas, Portugal had already settled the Atlantic islands of Madeira, the Azores, and Cape Verde; created trading posts along the West African coast; and established a trade route around South Africa to the Indian Ocean. Spain occupied the Canary Islands and, having expelled the Moors from the Iberian Peninsula, became intent on expanding its possessions overseas (Mauro, 2000; Elliott, 1984).

[3] Apart from the evidence that there is a negative contemporaneous effect of racial heterogeneity on public goods provision, and in turn on incomes, there is also some indication that racial fragmentation during the colonial era is causally related to underdevelopment today. Specifically, racial fractionalization in the past could have an independent effect on current development through poor institutions. Sokoloff and Engerman (2000) claimed that in highly unequal societies in the Americas, the (usually white) elites were able to maintain their advantage by restricting the right to vote and not investing in public education, resulting in lower per-capita incomes today. Therefore, it is possible that in multi-racial societies, the dominant group would establish institutions more conducive to patronage than growth. Two recent papers that find quantitative support for this hypothesis are by Easterly (2007) and Nunn (2007).

The economic model that Portugal and Spain applied in their previous expansion would greatly influence their conquest of the New World. On the African coast, Portugal established trading posts to exchange European products for gold, gems, and slaves; on the Atlantic islands, the most gain could be obtained through sugarcane cultivation, which required settlement. The islands were developed by assigning the lands to powerful *donatorios*, a practice adopted by Spain (Elliott, 1984).

Under Columbus' contract with the Spanish Crown for finding a westward route to Asia, Columbus was entitled to 10% of the profits of the enterprise as well as the title of Governor of all the lands he discovered (Moya Pons, 2000). Columbus never reached Asia, but laid claim to several islands in the Caribbean in 1492. In only fifty years, similar public/private ventures in search of gold and new territories resulted in the settlement of nearly all of the major cities in Spanish America today, from Mexico City to Santiago, Chile (Elliott, 1984). The Portuguese explorer Cabral reached the east coast of South America in 1500 while en route to India. During the following 30 years, trading posts were established to trade with the Natives for brazilwood.

In 1530, the Portuguese adopted the *donatorio* system and divided the land into fourteen captaincies where sugarcane plantations were established (De Abreu, 1997). These grants divided South America to the east of the Line of Tordesillas by lines of latitude. For practical purposes, the line connecting the mouth of the Amazon and the Río de la Plata divided the Portuguese and Spanish territories (Johnson, 1987). Areas far to the west of the Line of Tordesillas were eventually occupied and became official Portuguese lands under the Treaty of Madrid in 1750 (Mauro, 1987).

The early colonial history of South America reflects the pattern of European expansion in the Atlantic islands and on the African coast. The islands of the Caribbean and the coast of Brazil provided the ideal climate for growing sugarcane and other cash crops. In central Mexico, the highlands of Peru, and south-central Brazil, there was abundant mineral wealth, and, finally, the Natives themselves were a source of wealth.

III. Production in Colonial Brazil

During the sixteenth and seventeenth centuries, sugar was the largest component of national output and exports in Brazil (Mauro, 1987); it was not until the 1830s that the value of coffee exports surpassed that of sugar (Bethell and Murilo de Carvalho, 1989; Klein and Vidal Luna, 2010). The first major gold strike occurred in 1695, and gold production grew at a rapid pace over the following sixty years (Mauro, 1987). Plantation agriculture (namely sugar, but also coffee, tobacco, cotton, and cacao) and mining (primarily gold, but also diamonds, minerals, and gemstones) were the two primary activities in the colonial economy.

What was the nature of production on a sugarcane plantation or in a goldmine? The three main inputs in either economy were land, labor, and capital. Land was the abundant factor; capital and labor were scarce. Capital was limited to the personal wealth and connections of the settlers. Labor, on the scale required, could only be obtained through coercion of the Natives or by importing African slaves. In both sectors, Native labor was initially used and then substituted by African slaves. The production technology of each sector will now be considered in detail, beginning with sugarcane production.

i.) Sugarcane Production

The general process of sugarcane production was quite uniform throughout Brazil. According to Schwartz (1987), sugarcane needs to grow for fifteen to eighteen months before it can be harvested; it can be cut again after another nine months. At harvest, gangs would cut the cane, perhaps 4,200 canes a day per man; the women would tie the cane in bundles. At the mill, the cane juice would be expressed, then moved through copper kettles in which it was purified, and then put in molds to be purged (a process in which molasses is removed and the remaining sugar crystallizes). After drying, the sugar was separated by quality, crated, labeled, and sent to port.

According to Blume (1985), the size of an early sugar plantation ranged from 10,000 to 13,000 hectares (100 to 130 km^2). The plantation owner leased the land to *lavradores* who owned their own slaves and basic tools; the *lavradores* then processed their cane at the owner's mill. Subsistence crops were cultivated through sharecropping contracts on marginal parts of the plantation; slaves had to produce their own food. Sugarcane cultivation, and processing to a lesser extent, was extremely labor-intensive. A large plantation employed 120 to 160 slaves in the fields and about 20 whites; another 40 slaves were employed in the milling process.

A single *lavrador* usually worked one 6.1 hectare (ha) *partido* with 20 slaves. Based on an average cane yield of 53.51 tons/ha and a 5% recovery rate, each slave would have produced 51 *arrobas* (one arroba≈32 pounds) of sugar (Blume, 1985). Combined with data on the prices of sugar and African slaves from Schwartz (1987), this suggests that the average value of output per slave between 1608 and 1751 was about 57$018 Réis a year, compared to an average price per male slave of 63$889 Réis. The only major technological innovation during the colonial period was the introduction of a mill press with three vertical rollers at the turn of the seventeenth century. These were less reliant on water and led to a large expansion in the number of plantations (Schwartz, 1987).

ii.) Gold Production

Gold production, on the other hand, began on a much smaller scale. Some *bandeirantes* (slave-hunters) encountered placer gold during their expeditions and panned it with Native slaves using few or no tools. There were several such minor discoveries of gold in Brazil during the 1690s, but the discovery that beckoned a "gold-rush of epic proportions" was that of Ouro Prêto ('Black Gold'), "the greatest gold-mine in all South America," in 1701 (Hemming, 1978). The Native population in this region, which became known as Minas Gerais ('General Mines'), declined rapidly because of enslavement (albeit illegal), disease, and flight. High mortality among Native slaves working in the mines led to the importation of African slaves (Hemming, 1978).

According to A.J.R. Russell-Wood (1987), gold production dominated the economies of the states of Minas Gerais, Mato Grosso, and Goiás during the first half of the eighteenth century. Upon discovery of gold, an individual was usually granted two *datas* (1,080 sq. ft.) and was required to pay one-fifth of the extracted gold as tribute to the Crown. Gold production during the colonial period was very rudimentary; it involved rocking the river sediment in a wooden or metal pan so that the gold would sink to the bottom and other materials would wash over the sides. More sophisticated variants included *taboleiros*, where a whole river bed was worked, or *grupiaras*, where sediment was panned on the riverbanks or nearby hillsides. Openings in the hillsides, or *catas*, were worked by bringing the dug-out rock to be panned in the river or by using hydraulic pressure to bring water to the hillside. In such a *lavra*, sediment was passed through sluice boxes, gradually separating gold, and the residue was then panned. This was the most productive method of gold mining; however, it was also the most capital-intensive.

As on the plantations, Native labor was first used, but eventually African labor predominated. The supply of labor was a limiting factor in the scale of production, as was the lack of tools and other equipment. The only substantial technological innovation during the colonial period was the development of hydraulic machines to increase the availability of water for panning or to remove water from the *catas*. The average yield of a slave varied greatly over time; according to Teixeira Coelho, the average annual yield was 20 drams of gold (one dram=1/8 ounce) in 1780, but reports from the 1720s and 1730s range from one-half to one-and-a-half drams a week (27-81drams/year). Of course, the value of this gold depended on its purity, which varied across regions.

IV. Model

Both in the gold mines (*M*) and on the sugar plantations (*P*) in colonial Brazil, land was the abundant resource, slave labor predominated, and capital and skilled labor were crucial but scarce factors of production. Specifically, the owner of a mine or plantation had to choose the number of Native (*N*) and African (*A*) slaves to own, the number of hours of wage labor (*L*) to hire, and the amount of land (*T*) and capital (*K*) to employ.

The level of technology (A) determined overall productivity. Total output (Y) for each sector i ($i \in \{P,M\}$) in period t can be represented using the following Cobb-Douglas production function:

$$(1) \qquad Y_{it} = \Lambda_{it} F_{it}[K, L, A, N, T] = \Lambda_{it} K_{it}^{\alpha_i} L_{it}^{\beta_i} A_{it}^{\gamma_i} N_{it}^{\tau_i} T_{it}^{\mu_i}$$

No restrictions will be placed on the coefficients, except that they are not equal within or across sectors. Native (N) and African (A) labor could be considered as a single factor but the evidence suggests that they were not perfect substitutes. According to Ribeiro (2000), "No colonist ever doubted the usefulness of indigenous labor, even though he might prefer black slaves for mercantile export products. The Indian, on the contrary, was considered an ideal worker for transport of loads or people over land and water, for cultivation of commodities, for preparation of food, and for hunting and fishing," p. 63. Although Natives were more productive in subsistence farming and hunting, monoculture was alien to them and they had greater incentives to flee (Hemming, 1978). Africans were unfamiliar with the environment and had less motivation to run away; furthermore, some brought experience in mining and metallurgy (Klein and Vidal Luna, 2010).

The use of slave labor (N, A) was fundamentally different from the use of free labor (L). The cost of a slave was paid up-front and the rents from the slave accrued to the slave owner. Slaves were not paid their marginal product, although one could argue that wage laborers were. There are several factors that influenced how much a miner or sugar producer would be willing to pay for a slave. These included: slave characteristics (e.g. age, gender, and size), maintenance costs, the value of children, and the risk of legal changes (e.g. abolition) that would limit (or eliminate) ownership rights. In this model, slaves are assumed to be uniform and not differ according to gender, only race.

Each period, the slave owner would purchase new slaves (I_{it}^{A} , I_{it}^{N}) that would increase the stock of slaves. Absent new investment, the slave population would decline due to net mortality (a_i, n_i) and, potentially, flight or manumission.[4] This model will consider the transition of the stock of African and Native slaves in sector i in time t (A_{it}, N_{it}) in terms of new investment (I^A_{it} , I^N_{it}) and (negative) population growth rates (a_i, n_i). The transition equations for the stock of African and Native slaves are as follows:

$$
\begin{aligned}
(2) \qquad A_{i,t+1} &= I^A_{\text{it}} + (1 + a_i)\, A_{\text{it}} \\
N_{i,t+1} &= I^N_{\text{it}} + (1 + n_i)\, N_{\text{it}}
\end{aligned}
$$

A plantation or mine owner would choose the amounts of inputs to maximize the net present value of profits. The choice variables in this problem are the levels of capital (K_{it}), wage labor (L_{it}), land (T_{it}), and the numbers of African (I^A_{it}) and Native (I^N_{it}) slaves to purchase. The state variables are the level of technology (Λ_{it}) and the current stock of African (A_{it}) and Native (N_{it}) slaves. The prices of output (p_{it}) and African (p^A_{it}) and Native (p^N_{it}) slaves are exogenous (this will be discussed later). The prices of capital (r_{it}), free labor (w_{it}), and land (ι_{it}) – net interest, the wage rate, and rent, respectively – are set in the domestic market (endogenous). There is full depreciation and land is rented, not owned. Let the subjective discount rate be ρ. The Lagrangean for the firm's profit-maximization problem is

$$
\begin{aligned}
\mathcal{L} = \max\nolimits_{(K_{\text{it}}, L_{\text{it}}, A_{\text{it}}, N_{\text{it}}, T_{\text{it}} \geq 0)} \\
\textstyle\sum_{t=0}^{\infty} \rho_i{}^t \{ p_{\text{it}}\, \Lambda_{\text{it}}\, K^{\alpha_i}_{\text{it}}\, L^{\beta_i}_{\text{it}}\, A^{\gamma_i}_{\text{it}}\, N^{\tau_i}_{\text{it}}\, T^{\mu_i}_{\text{it}} - r_{\text{it}}\, K_{\text{it}} - w_{\text{it}}\, L_{\text{it}} - \iota_{\text{it}}\, T_{\text{it}} \\
(3) \qquad - p^A_{\text{it}} (A_{i,t+1} - (1 + a_i)\, A_{\text{it}}) - p^N_{\text{it}}\, (N_{i,t+1} - (1 + n_i)\, N_{\text{it}}) \}
\end{aligned}
$$

The first-order conditions are

[4] Manumission, the movement from a slave to free laborer, will not be explicitly modeled here. If one includes manumission rates, the predictions for the racial composition still hold; however, the analysis is complicated by the fact that wages are inversely related to the quantity of slaves. This is an interesting relationship that could be investigated further.

$$(4) \qquad K_{it} : \alpha_i \, p_{it} \Lambda_{it} \, K_{it}^{\alpha_i - 1} \, L_{it}^{\beta_i} \, A_{it}^{\gamma_i} \, N_{it}^{\tau_i} \, T_{it}^{\mu_i} - r_{it} = 0$$

$$L_{it} : \beta_i \, p_{it} \Lambda_{it} \, K_{it}^{\alpha_i} \, L_{it}^{\beta_i - 1} \, A_{it}^{\gamma_i} \, N_{it}^{\tau_i} \, T_{it}^{\mu_i} - w_{it} = 0$$

$$T_{it} : \mu_i \, p_{it} \Lambda_{it} \, K_{it}^{\alpha_i} \, L_{it}^{\beta_i} \, A_{it}^{\gamma_i} \, N_{it}^{\tau_i} \, T_{it}^{\mu_i - 1} - \iota_{it} = 0$$

$$A_{i,t+1} : -\rho_i{}^t \left(p_{it}^A\right) + \rho_i{}^{t+1} \Big(\gamma_i \, E_t \left[p_{i,t+1} \Lambda_{i,t+1} \, K_{i,t+1}^{\alpha_i} \, L_{i,t+1}^{\beta_i} \, A_{i,t+1}^{\gamma_i - 1} \, N_{i,t+1}^{\tau_i} \, T_{i,t+1}^{\mu_i} \right] + E_t \left[p_{i,t+1}^A \right] (1 + a_i) \Big) = 0$$

$$N_{i,t+1} : -\rho_i{}^t \left(p_{it}^N\right) + \rho_i{}^{t+1} \Big(\tau_i \, E_t \left[p_{i,t+1} \Lambda_{i,t+1} \, K_{i,t+1}^{\alpha_i} \, L_{i,t+1}^{\beta_i} \, A_{i,t+1}^{\gamma_i} \, N_{i,t+1}^{\tau_i - 1} \, T_{i,t+1}^{\mu_i} \right] + E_t \left[p_{i,t+1}^N \right] (1 + n_i) \Big) = 0$$

As evident from the first three conditions, capital, labor, and land are paid their marginal products. African and Native slaves, however, are not. The slave-owner would choose the next-period stock of slaves to equate their expected marginal product to their expected marginal cost: a function of the current and expected price of slaves, the subjective discount factor, and the population growth rate. By solving the last two first-order conditions for the ratio of Africans to Natives in period *t+1* ($A_{i,t+1}/N_{i,t+1}$) assuming that slave prices follow a random walk and that output prices and the technology level are independent of the number of slaves, one obtains

$$(5) \qquad \frac{A_{i,t+1}}{N_{i,t+1}} = \frac{p_{it}^N \, \gamma_i \, (1 - \rho_i - \rho_i \, n_i)}{p_{it}^A \, \tau_i \, (1 - \rho_i - \rho_i \, a_i)}$$

The ratio of African to Native slaves in sector *i* in time *t+1* ($A_{i,t+1}/N_{i,t+1}$) is equal to the inverse ratio of their prices in the previous period (p_{it}^N / p_{it}^A) times a constant that captures relative productivity (γ_i/τ_i) and population growth rates (n_i/a_i), plus a subjective discount factor (ρ_i). Note that a change in the price of output (p_{it}), the level of technology (Λ_{it}), or the level of capital (K_{it}), free labor (L_{it}), or land (T_{it}), has no effect on the ratio of

African to Native workers ($A_{i,t+1}/N_{i,t+1}$). The "race ratio" reacts negatively to increases in the price of African workers (p^A_{it}) and positively to increases in the price of Native workers (p^N_{it}). Finally, the ratio reacts positively to an increase in the share of African workers in production (γ_i) and negatively to the share of Native workers (τ_i). It is also clear from equation (5) that a decrease in the Native population (N_{it}) would increase equilibrium demand for African labor ($A_{\mathrm{it}}{}^*$), *ceteris paribus*, and vice-versa. All intermediate steps for solving the model above are shown in Appendix I.

The population growth rates will not affect the estimation of equation (5) in the following section, provided that this ratio is, in fact, constant. The evidence suggests that the Native population declined at a faster rate than the African population on the plantations and in the mines ($n_i/a_i>1$). This is partially due to an asymmetric impact of disease on the Natives. In both the plantation and mining sectors, mortality rates were high: estimates of the working life of a slave in the mines range from seven to twelve years (Russell-Wood, 1987) and the mortality rate on a sugarcane plantation was estimated by planters at five to ten percent a year (Schwartz, 1987).

The Substitution of African for Native Labor

It comes as no surprise that the optimal ratio of Native to African labor is inversely related to the ratio of their prices. The only novelty in this profit maximization problem is that the price of each type of labor is determined exogenously. The price of African slaves was determined in a competitive international market and plantation owners in the Americas were essentially price-takers (Sokoloff and Engerman, 2000). It reflected, to a large extent, the cost of the trans-Atlantic voyage and high rates of attrition. The price of Native labor, on the other hand, was related to the costs of slave-raiding expeditions into the interior; as new lands were settled and Native populations declined, the colonists had to travel increasing distances over rugged terrain in order to find Native villages.

During the early phase of settlement, the coast supported large Native populations and their labor was easily obtained through barter or enslavement. The relatively low price of Native slaves during the sixteenth century is documented in the will of Governor Mem de Sá (1569), who valued his African slaves at thirteen to forty times more than unskilled Native slaves (Hemming, 1987). When new settlements were established a century later in the interior, at a comparable distance to the ever-shifting frontier, Native labor was also many times less costly than African labor. For example, the price of renting Native labor from the Jesuit missions in 17th century Pará was fixed at only two and a half yards of cloth per month – almost valueless for barter (Hemming, 1987).

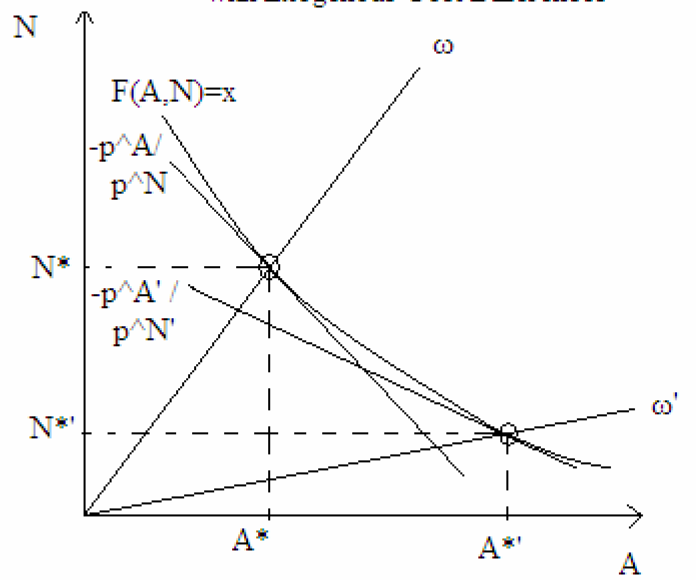
In the settlements, Native labor soon became scarce. According to Pedro Calmon (1939), "on the coast of Bahia and Pernambuco, the substitution of Native slaves by African was most rapid during the decade of 1590-1600," p. 347 [author's translation]. In a letter to the King of Portugal in 1617, the Bishop of Lisbon wrote, "in the whole district [of the city of Maranhão], there is not a single Indian village left. Within a hundred leagues of Pará there is not a single Indian who is not at peace or has not been domesticated by the Portuguese, whom he fears even more than a slave fears his master. In the district of Ceará, there used to be 60 villages within a radius of 60 leagues. Today, not one remains, for they have all disappeared as a result of the activities of the slave hunters," p. 42 (Marcílio, 1984). Hemming (1978) writes, "[By the end of the seventeenth century] *bandeirante* activity in the south was subsiding, with the disappearance of most Indians in the hinterland of Paraná and São Paulo and the realization that Indian captives were scarcely worth the effort now needed to capture them," p. 180.

Thus, in the colonial cities that continue to be the population centers of Brazil today, Native American labor became increasingly difficult to obtain and plantation and mine-owners substituted African slaves. Native American labor was not traded between colonies hence its price varied across regions.[5] As an illustration, compare a hypothetical coastal city in seventeenth-century Brazil to a new settlement in the interior. Suppose

[5] In this respect, this is related to the approach that Goldin and Sokoloff (1984) employ to explain why industrialization first occurred in the Northeast United States.

that, consistent with the historical evidence, the coastal city is far enough away from the frontier such that African slaves are relatively less expensive than Native labor. In the city in the interior, the proximity to the frontier makes Native labor relatively inexpensive as compared to the coast. In this scenario, *ceteris paribus*, the first-order conditions of the profit-maximization problem would yield equilibria like those shown in Figure I for a given level of output x.

Figure I. Optimal Labor Ratios in Two Settlements with Exogenous Cost Differences



Lemma 1:

The coastal city that is distant from the frontier has an optimal labor ratio of Africans to Natives that is higher than the city in the interior ($p^{N}{}'/p^{A}{}' > p^N/p^A \rightarrow A^*{}'/N^*{}' > A^*/N^*$).

The model presented in this section has shown the optimal decision for the representative firm in sector i, where the two sectors are plantation agriculture and mining. Considering that each settlement in Brazil (then captaincy and eventually state) specialized in only one sector, i can also be interpreted as a settlement (captaincy, and then state) without changing the key result in equation (5) and Lemma 1. The geography of each settlement (its location and endowments) may explain its observed racial composition over time.

The discussion above supports the assertion that the ratio of the price of Native labor to the price of African labor in settlement *i* in time *t* (p_{it}^{N}/p_{it}^{A}) is a function of the ratio of the travel costs to the frontier to the travel costs to Africa. Of course, the transport costs are increasing with distance. Also, the cost of traveling by sea was much lower than traveling by land; however, mortality rates during the sea voyage were high (Galeano, 1971).[6] It follows that the price ratio of Africans and Natives is some function of the ratio of the distance to the frontier to the slave ports in Africa – in fact, it could simply be equal to the ratio of the distances times some factor capturing the markup, the relative cost of land and sea travel, and the attrition of slaves during the trip from Africa.[7] This is represented by the following equation:

$$(p_{it}^{N}/p_{it}^{A}) = f\,(d^{F}_{it}/d^{A}_{i}) = \eta_i\,(d^{F}_{it}/d^{A}_{i}) \qquad (6)$$

The term d^{F}_{it} is the distance from the principle city of settlement *i* in time *t* to the frontier, d^{A}_{i} is the distance from that city to the port-of-origin of slaves in Africa (note: there is no time subscript), and η_i is some constant. Is this reasonable? Consider that Mem de Sá valued some African slaves forty times more than Native slaves in Bahia in 1569. Given that the ratio of the distances was almost 1:100 (there were un-contacted Natives within 40 km from the capital but Africa was over 4,000 km away), the implied value of η_i is 2.5. This is the value one would expect assuming that overland travel (with no roads) was five times more expensive than overseas travel and the attrition rate of African slaves was 50%. Although this is oversimplified, transportation costs were certainly a major component of the price of slaves.

---

[6] One shocking example is provided by the records of the Royal African Company: out of 70,000 slaves that embarked between 1680 and 1688, only 46,000 survived the voyage. In the specific data for Brazil, as compiled by Klein (2010) from the Transatlantic Slave Trade CD-ROM (Eltis et al., 1998), between 1750 and 1850 Portuguese slave ships had an average 8.7% mortality rate during the passage to America.

[7] If the transportation costs increased exponentially, then the relationship between distance-to-origin and slave prices would not be linear. This does not appear to be the case. According to Klein (2010), the distance traveled was less important in explaining mortality than the conditions prevalent in the port of departure: "In regard to length of voyage, there was some increase in mortality with numbers of days at sea, particularly for those unexpectedly long voyages on which water and provisions ran short and accelerated the spread of disease. For the great number of voyages, however, there was little variation that could be directly explained by the differences in the number of days at sea… there was little difference that the length of a voyage from a given port made in the mortality rates," p. 144.

The following section will describe the data that will be used to test the predictions of this model; in particular, whether the ratio of the distance to the frontier to the coast of Africa can explain the racial composition of the labor force (and the total population) over time.

V. Demographic Data

The unit of analysis for the present study is the modern Brazilian state-year, spanning 1500 to 2000. A separate contribution of this paper is the compilation of all published data on the population and racial composition of Brazil over its history. Obtaining data on the disaggregated population of Brazil prior to 1800, much less its racial composition, is difficult. According to Marcílio (1984), "Without data, however, there is no demography, and in the case of Brazil there is practically no statistical information for the first 250 years of its existence," p. 37.

Although it is impossible to fully consider population dynamics during the colonial period, it *is* possible to obtain snapshots of the racial composition of society at different points in time. The first national census data are available in 1872 (Bergad, 2007), but there are regional censuses beginning in 1775 (Alden, 1963); official documents, correspondence, and chronicles by contemporaries provide population figures by city and/or captaincy going back to 1545. The accuracy of the data decreases further back in time, but because the key variable of interest here is the ratio of the population of two racial groups, not the total headcount, the noise in the data is lessened. After reviewing all published primary and secondary sources, about twenty were identified that together contain enough information to construct a modest panel of population by race. Some data is available for twenty Brazilian states at or around fourteen time periods: 1545, 1570, 1585, 1625, 1675, 1725, 1775, 1800, 1825, 1850, 1875, 1900, 1950, and 2000.

Unfortunately, this dataset is far from complete. The data coverage by period, race, and number of settlements is shown in Table I. The sources of demographic data and the construction of the database are detailed in Appendix II.

Table I. Data Coverage by Race over Time†			
Period	Obs. # "European"	Obs. # "African"	Obs. # "Native"
1545	2	2	1
1570	6	1	3
1585	6	2	5
1625	8	3	2
1675	3	3	3
1725	2	5	2
1775	6	6	2
1800	14	14	12
1825	10	10	9
1850	10	9	8
1875	20	20	20
1900	20	20	20
1950	20	20	0
2000	20	20	20
Total # Obs.	**147**	**135**	**107**

†Note: The data described is from multiple sources (detailed in the appendix) and only includes published historical statistics. For early periods, the observations correspond to the population of settlements within the boundaries of twenty modern Brazilian states.

i.) Estimating the 'Missing' Native American Population

A problem with the historical sources is that the number of Native Americans is often unreported. Nevertheless, the current literature on the size of Native populations in the Americas before and after colonization can be used to infer the Native population of the settlements for periods in which data is missing. The robustness of these estimates will be demonstrated by their close correspondence to the actual data in cases where they overlap. Furthermore, Section VII will demonstrate that the empirical results in the next section are not sensitive to a wide range of assumptions regarding the natural rate of increase of the Native population.

The estimates of the Native population in some Brazilian settlements during the colonial period are based on the following logic: 1. Natives previously living in areas settled by the Portuguese were incorporated into the local economy, and, 2. The natural growth rate of Native populations in the settlements was negative due to the effects of disease and enslavement, therefore, 3. The population of Natives living in Portuguese settlements could only be maintained or increased by bringing in more Natives from the interior.

A brief historical account supports these conclusions. Following Brazil's discovery, the Portuguese bartered Europeans goods with the Natives in exchange for food and labor, but by 1549, there is evidence of Native enslavement (Marchant, 1942). The Portuguese (and French) presence led to hostilities with the Natives, creating a vicious cycle of warfare and enslavement that vacated many Native villages by the end of the sixteenth century. Natives were forced to relocate to *aldeias* where they provided food and other services for the colonists. The effect of European diseases was devastating: according to Marcílio (1984), the Bahian smallpox epidemic of 1562-65 may have claimed 30,000 Native lives.[8] In response to the shortage of labor, colonists ventured beyond the frontier to raid villages for slaves under a pretext of war against hostile Natives (Langfur, 2006).

In order to estimate the number of Natives living in colonial settlements, this paper will use estimates of the population density of Brazilian tribes in 1500 combined with data on the areas settled by the Portuguese and estimates of the rate of decline of the Native population over time.[9] The population density estimates come from Julian Steward (1946) and are based on detailed ethno-historical data. A digital version of his original hardcopy map is presented in Appendix III. Data on the areas settled over the colonial period are taken from John Hemming (1978). A digital version of his original hardcopy map is presented in Appendix IV. Finally, the average rates of decline for the Native Brazilian population are adopted from Rosenblat (1954).

---

[8] The testimony of the Jesuits suggests that only nine or ten thousand Natives survived in the proximity of Bahia; the original population was estimated at 80,000 (Marchant, 1980).

[9] Although Natives in the settlements may have had higher mortality rates, even Natives that had no contact with Europeans were still affected by European diseases and inter-tribal warfare.

The data was combined to estimate the number of Native Americans in each settlement as follows: assuming that the natural rate of increase of the un-acculturated Native population was negative in the magnitude suggested by Rosenblat, then it follows that the number of Natives in a settlement in any period was roughly equal to the number of un-acculturated Natives in the new regions settled since the previous period. The estimates obtained using this methodology are surprisingly close to the number of Natives recorded by contemporary observers when such data is available.[10] The data are only included in the database when other data is unavailable. In total, these estimated figures comprise 49 out of 312 points, or 15.7 percent, giving a total of 156 observations.

ii.) Estimating the 'Missing' African Population

As with the Native population, the African population is also under-reported over the course of Brazilian history, albeit to a lesser extent. In a separate work entitled "Food, Drugs, and Gold: Coerced Labor in Colonial Brazil," the author examines the labor requirements in the sugar and gold sectors, as well as the importation of African slaves, in order to estimate the African workforce in each settlement during the colonial era. This work identifies a close association between the labor requirements in these two sectors and the observed African population. As with the Natives, the flow of Africans into the settlements can also be used to project a possible population range.

The economy was centered on sugar production until the beginning of the gold rush at the end of the seventeenth century. This is reflected in the close correspondence between the estimated workforce in sugar production and the African population. For instance, given the number of plantations in Bahia/Ilhéus in 1545, 320 slaves were required – compare this with 400 Africans recorded by a contemporary document. In São Vicente, an

[10] The dataset contains a combined 18 overlapping observations for Native populations during the sixteenth, seventeenth, and eighteenth centuries. In São Paulo, the 'observed' Native population was 3,000 in 1545, 1,000 in 1570, and 1,000 in 1585 – compare this with estimates of 1,324, 1,291, and 1,272, respectively. In Bahia in 1585, the average observation of 7,335 Natives compares with an estimate of 2,272. That year in Pernambuco, the 2,000 observed Natives compares with an estimated 1,254. In Maranhão in 1725, there were supposedly 12,500 Natives compared to an estimate of 6,354. In Pará in 1775, 28,391 recorded Natives compares with the estimate of 23,405. Over these 18 sets of competing figures, the average deviation is 8,033 people, a small margin given the noise in the early statistics.

estimated 480 slaves were needed on the plantations – 500 were actually documented in the captaincy at this time. In 1570, 1585, and 1625, the estimates based on the number of plantations are relatively accurate in predicting the African population.[11] Accordingly, 17 of the production-based African population estimates will supplement the dataset.

In addition, four estimates based on the inferred African population from African slave import records will also be added to the database.[12] These are built upon conservative assumptions regarding the natural rate of increase (decrease) of the slave population and are shown to correspond closely with the observed data when it is available. For instance, scholars have placed the African population of Bahia in 1675 at 25,000 – based on slave import records and moderate negative rates of natural increase over time, the estimated African population is 25,381. The numbers correspond well for subsequent periods from 1725 to 1850.[13] With four additional estimates for the dataset, there are a total of 21 new observations on the African population, or 6.7 percent of all data points.[14]

iii.) Major Demographic Trends

To illustrate some trends in the dataset, the evolving racial compositions of three geographical regions – the North, the Northeast, and the Southeast – are shown in Figure II. The "North" includes the states of Pará and Maranhão; the "Northeast" consists of the states of Paraíba, Pernambuco, and Bahia; and, the "Southeast" comprises Espírito Santo, Rio de Janeiro, and São Paulo. Outliers are omitted. In 1775, national statistics become

[11] Although the estimate of 320 slaves in São Vicente in 1570 understates the 2,000 observed, for 1585, the estimates of 4,240 in the Bahia region and 5,280 in Pernambuco are quite close to the average 4,165 and 5,665, respectively, observed by contemporaries; in 1625, the estimate of 15,486 slaves in Pernambuco compares unfavorably with the 5,000 Africans supposedly in the captaincy at this time. Nonetheless, the early observations have a wide margin of error; therefore, the production-based estimates are useful checks for the available population statistics and, when such data is lacking, serve as reliable substitutes.

[12] The slave import records are from the Trans-Atlantic Slave Trade Database (http://slave voyages.org) accessed August, 2010. Based on evidence cited in the aforementioned work, the growth rate of the African population is assumed to be -5%, -4%, and -2%, respectively, in the 16th, 17th, and 18th centuries.

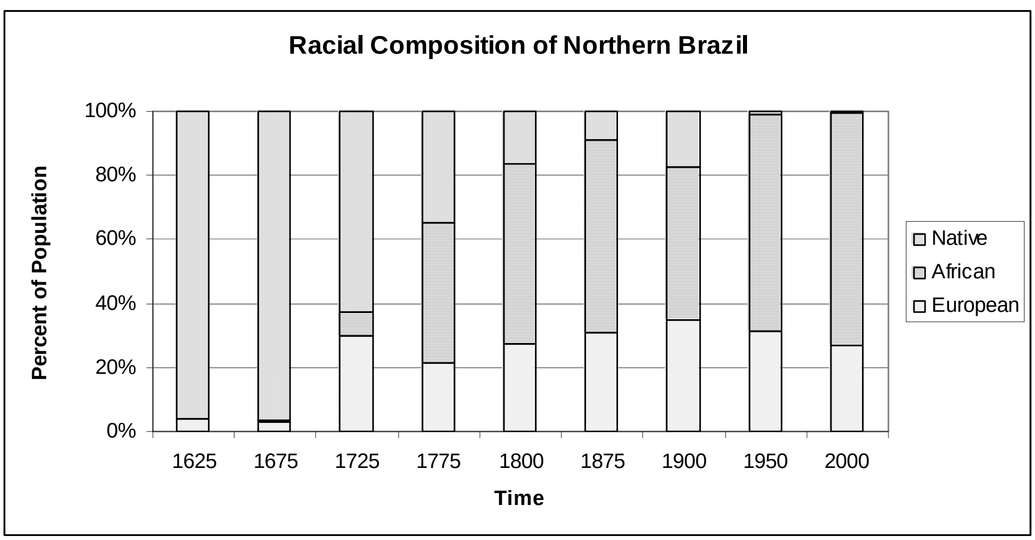
[13] In 1725, the estimated African population in "Amazonia" of 5,685 compares with 3,000 recorded; in Bahia that year, the estimate of 36,594 compares to an observed 45,482; in 1775, 33,997 Africans in Amazonia compares with a recorded 45,164; the same year in Bahia, the estimate of 280,342 is close to 216,636 recorded; finally, from 1800 to 1850, the estimates are within 25-50% of the observed population.

[14] It will be shown that the imputed data is not driving the results. The model will be tested with the base dataset described in Table I (with 100 obs. included) as well as the enhanced dataset (all 156 obs. included).

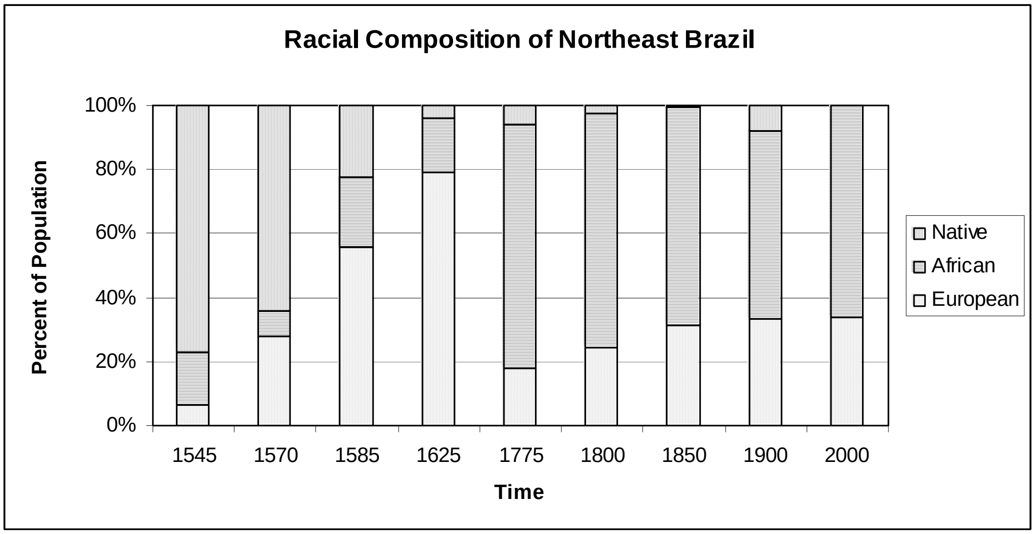
available with “mulatto” (of African and European descent) as an additional racial category – these individuals are classified as being of predominantly “African” descent.[15]

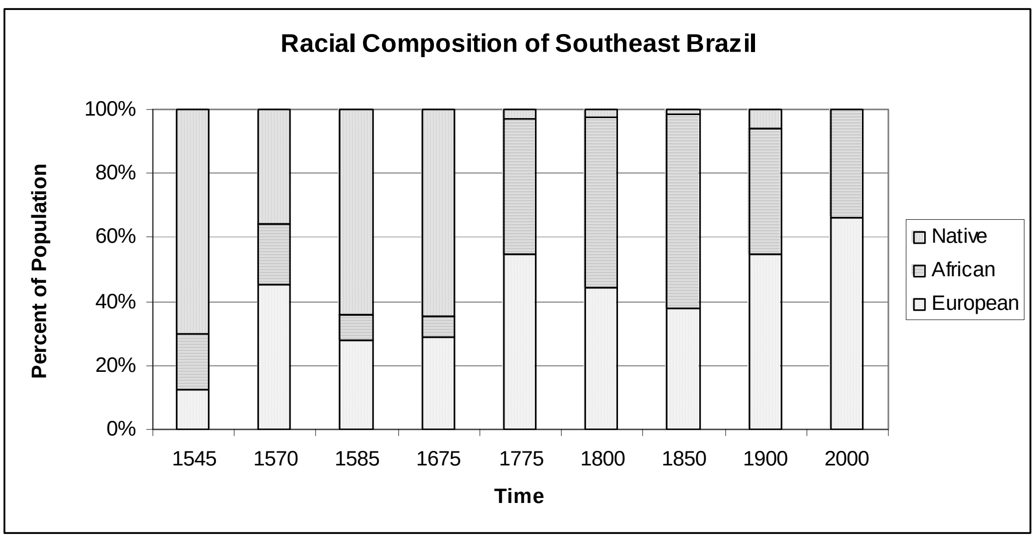
The three regions shown in Figure II all display a shift from societies in which the majority of the population was of Native descent into ones of predominantly African descent – and eventually in the Southeast, of predominantly European descent. This transition was not uniform across regions. In the North, with relatively high initial Native populations and a less-developed export industry, the transition into an African-American society occurred between 1725 and 1775. In the Northeast, the prime sugar-producing region of Brazil, the African population exceeded the Native between 1585 and 1625. Finally, in the Southeast, which despite a colder climate was still export-oriented, this occurred sometime between 1675 and 1775.

Figure II. The Racial Composition of Three Regions of Brazil

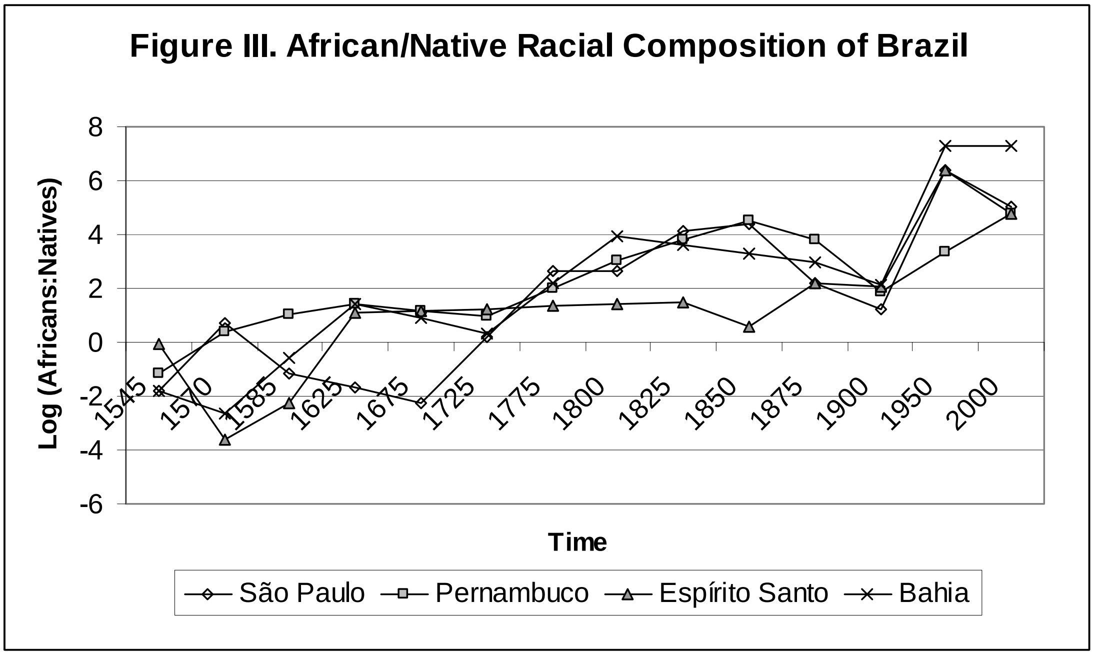


[15] “Mestizo,” or of European and Native descent, is not a separate category in the modern census. Most Brazilians have some Native ancestry (Ribeiro, 2000); however, in the census they self-identify as either “white” (branco), “yellow” (amarelo), “black” (preto), “brown” (pardo), or “indigenous” (indígena).





Across Brazil, the Indigenous majority had become a minority by 1675. This is consistent with the observation of Schwartz (1987): "The transition from Indian to African labour, although underway from the 1570s on, was slow and was not fully achieved in the plantation areas until the third decade of the seventeenth century," p. 82. This conversion was not in effect in all places at all times, but in the long-run it is undeniable. For the four states with the best data coverage – São Paulo, Pernambuco, Espírito Santo, and Bahia – the shift toward an African population is visible in Figure III (four outliers are excluded plus nine other data points are smoothed or 23% of total).



A challenge to this data is that the social construction of race makes it biased, particularly in the later periods. The fraction of the population classified as "white" increased notably beginning with the Trans-Atlantic Migration at the end of the nineteenth century. However, the "whitening" of the population was at once a physical and psychological phenomenon. According to Abdias do Nascimento (1995), in spite of government statistics indicating that the majority of Brazil's population today is "white", Africans still constitute the majority. He writes, "In this distortion of statistics we come upon a cornerstone of 'Latin' racism: the psycho-social whitening of Africans in these societies. The compulsion to identify with European values, aesthetics and criteria of personal beauty create various negative psychological complexes …" (p. 102).

This incentive to become "white" creates a bias in the data, predominantly for the nineteenth and twentieth centuries, in which Africans are underreported. Nonetheless, this bias would work against the proposed relationship between racial composition and labor costs because one would expect greater numbers of Africans, not fewer, as the frontier extends over time. The next section will consider potential proxies for the variation in labor costs across the two races.

iv.) Distance Measures

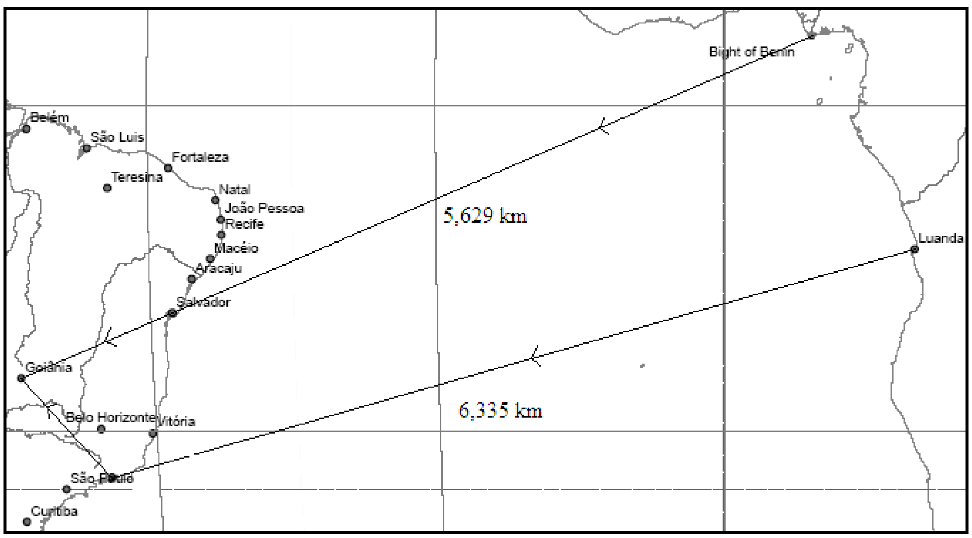
The population counts by race for each settlement in the fourteen periods are used to calculate the dependent variable: the ratio of Africans to Natives ($A_{i,t}/N_{i,t}$). The independent variable is the relative price of Native and African labor ($p_{i,t}^{N}/p_{i,t}^{A}$). As discussed, a natural proxy for the relative labor cost is the ratio of the distance to the frontier to the port-of-origin of African slaves ($d^{F}_{i,t}/d^{A}_{i}$). The distance to the frontier (d^{F}_{it}) is the more subjective of the two measures because the "frontier" has many possible interpretations. The distance to Africa (d^{A}_{i}) would be more straightforward, if not for the numerous Portuguese slave ports on the coast, points of disembarkation in the Americas, and transportation routes to captaincies in the interior.

Two different ways of measuring each distance will be considered. In each case, the distances are measured relative to the principle city in each settlement. The "frontier" is regarded as the division between the Native and Portuguese domains; the areas each group frequented, if not inhabited, and was capable of defending. The first method of determining the frontier involves tracking the establishment of new cities in the interior. The "frontier" will be defined as 150 miles (241.5 km) west of the westernmost city that existed during a given time period.[16] This distance is based on what early Jesuit administrators reported as commonly-traversed areas between plantations, parishes, and cities.[17] The second "frontier" comprises the westernmost points settled by the Portuguese as shown in Hemming (1978).[18] Since Hemming's map only spans a few periods (Appendix IV), the distance to the frontier in intermediate periods is assumed to increase linearly. This is not as satisfactory, since the frontier expanded and contracted.

[16] The cities and their gradual establishment are partially outlined in a series of figures in a previous work by this author, "The Demographic History of Brazil." Since the data is presented by century, the location of the frontier in intermediate periods is averaged.
[17] See Joseph de Anchieta, "Cartas, Informações, Fragmentos Historicos e Sermões," 1585.
[18] Hemming's map shows the limits of Portuguese occupation in 1600, 1700, and 1760. It is assumed that the frontier shifted linearly from an initial value of 0 in 1500 to its location in 1600, and likewise between 1600/1700 and 1700/1760. Thereafter, it is assumed that the frontier continued to expand linearly until reaching Brazil's current political boundaries.

The distance to the origin of slaves in Africa, assuming one may identify the origin, can be proxied by: 1.) The shortest distance from the principal city in each captaincy to the port on the coast of Africa or, alternatively, 2.) The distance from the principal city to the port of disembarkation plus the additional distance from that port to the port of embarkation. In the first scenario, the port of embarkation in Africa will be a common point (5°50'E, 4°12'N) near the Bight of Benin (it is time-invariant within, but not across, states). In the second scenario, the representative port of embarkation in Africa will be Luanda, Angola (13°12’E, 8°40’S). An example using Goiás is shown in Figure IV.

Figure IV. Two Measures of the Distance to Africa



The source of African slaves varied by region and time, but the Bight of Benin is about the average location on the coast where slaves were purchased; slaves were referred to as being from either Guinea, to the north of the Bight, or Angola, to the south of the Bight (Calmon, 1939). The data published by Klein and Vidal Luna (2010) from the Emory Slave Trade Database suggest that West Central Africa (whose main Portuguese port was Luanda) sent the greatest number of slaves to Brazil, 71 percent; substantial numbers were also sourced from the Bight of Benin, 19 percent of all slaves arriving in Brazil.

Descriptive statistics by state for some of the variables defined in this section are presented in Appendix V; aggregate statistics for Brazil are presented below in Table II.

Table II. Descriptive Statistics†									
	# Africans	# Natives	#Africans / #Natives	#Afr. / #Nat. No Outliers[1]	# Europeans		Distance Frontier (km)	Distance Africa (km)	Dist. Frontier / Dist. Africa
Base Data						*Measure One*[3]			
Mean	162,418	24,353	18.3	"	877,958	Mean	457	5,081	0.09
Std. Dev.	(303,684)	(33,307)	(44.5)	"	(2,679,454)	Std. Dev.	(365)	(813)	(0.08)
# Obs.	135	107	100	"	147	# Obs.	280	280	280
All Data						*Measure Two*[4]			
Mean	141,878	18,079	2,269	23.0	"	Mean	723	5,538	0.14
Std. Dev.	(287,269)	(29,385)	(18,084)	(70.0)	"	Std. Dev.	(828)	(773)	(0.18)
# Obs.	156	156	156	140	"	# Obs.	280	280	280
Base Incl. Mulatto[2]									
Mean	783,273	"	96.8	"	"				
Std. Dev.	(1,662,195)	"	(289.8)	"	"				
# Obs.	135	"	100	"	"				
All Incl. Mulatto									
Mean	679,156	"	3,504	101.9	"				
Std. Dev.	(1,568,045)	"	(23,498)	(309.1)	"				
# Obs.	156	"	156	140	"				

†The unit of observation is the state-period. There are 20 states included across 14 periods, or potentially 280 obs. for each variable.

[1] The estimates of the Native Population in some periods are near-zero because the frontier did not move, the main factor determining the number of new Natives introduced into society during each period. In this case, the numerator in the third column is potentially thousands of times greater than the denominator, creating large variation in the data. Observations over ten times greater than adjacent values are removed.

[2] This paper is concerned with the ratio of individuals of primarily African descent to those of primarily Native descent. The separate statistics for Mulattos, available beginning in 1775, are added to the African population for the purpose of capturing the entire population of primarily African descent. It is preferable to include mulattos to avoid grossly understating the size of the African-American population in Brazil.

[3] The distance to the frontier is the great-circle distance from each state's capital to the western edge of a 241.5km-radius circle buffering the westernmost city established up until each period. The distance to Africa is the great-circle distance from each state's capital to the Bight of Benin (6°E, 4°N).

[4] The distance to the frontier is the great-circle distance from each state's capital to the frontier as mapped by Hemming (1978); a linear expansion of the frontier is assumed for intermediate periods when the frontier is unknown. The distance to Africa is the great-circle distance of each state's capital to the nearest port of disembarcation of African slaves plus the great-circle distance from that port to the city of Luanda, Angola (13°E, 9°S).

VI. Empirical Analysis

With data on the racial composition of up to twenty Brazilian states in fourteen periods, it is possible to empirically test the hypothesis that variation in the racial makeup of countries in the New World was determined, to some extent, by the marginal product of labor and its supply. In the model above, the ratio of African to Native workers ($A_{i,t}/N_{i,t}$) in settlement (or state) i ($i \in \{1,2,\ldots,I\}$) is equal to the ratio of their respective production coefficients (γ_i/τ_i) times the inverse ratio of their extra-market prices ($p_{i,t}^N/p_{i,t}^A$); the 'race ratio' is also a function of the subjective discount (ρ_i) and mortality rates (n_i, a_i), as shown in equation (5). This paper argues that the relative price of Native and African labor ($p_{i,t}^N/p_{i,t}^A$) during the colonial era was a function of the ratio of the distance to the frontier (beyond which there were un-contacted Natives) to the African coast (where there was a relatively elastic supply of slaves) – $d^F_{i,t}/d^A_i$ as shown in equation (6). Plugging equation (6) into equation (5) and taking logs of both sides, the hypothesized relationship is

$$\ln\left(\frac{A_{i,t}}{N_{i,t}}\right) = \ln\left(\frac{1-\rho_i-\rho_i n_i}{1-\rho_i-\rho_i a_i}\frac{\gamma_i}{\tau_i}\eta_i\frac{d^F_{i,t-1}}{d^A_i}\right) = \ln\left(\frac{1-\rho_i-\rho_i n_i}{1-\rho_i-\rho_i a_i}\frac{\gamma_i}{\tau_i}\eta_i\right) + \ln\left(\frac{d^F_{i,t-1}}{d^A_i}\right)$$

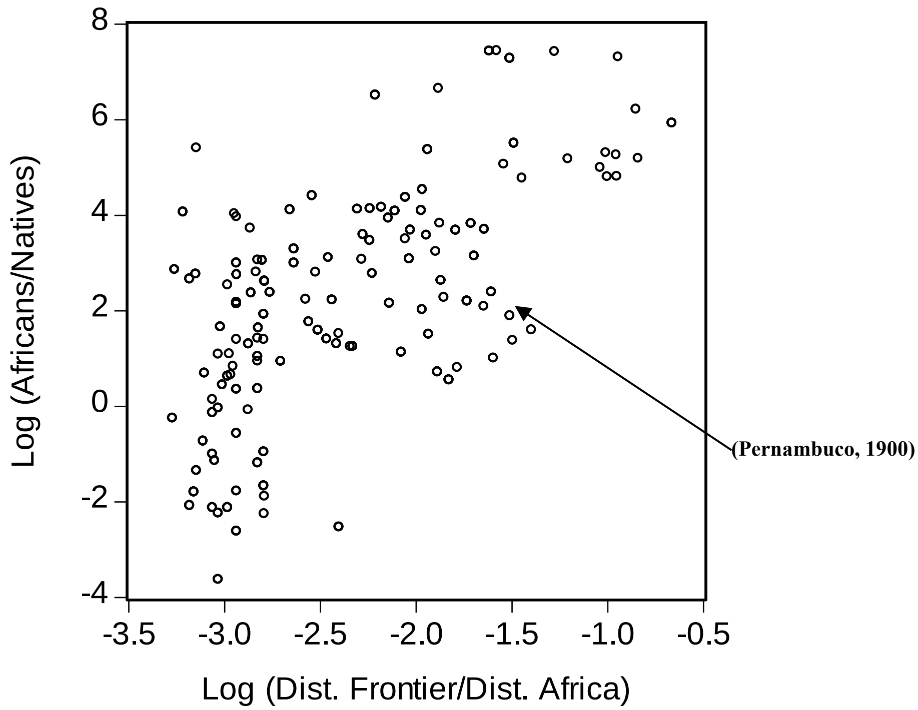
or

$$\ln\left(\frac{A_{i,t}}{N_{i,t}}\right) = c_i + \ln\left(\frac{d^F_{i,t-1}}{d^A_i}\right) \qquad (7)$$

, where c_i is a constant.

Figure V presents a scatterplot of available data of the log ratio of Africans to Natives ($\ln(A_{i,t}/N_{i,t})$) – including Mulattos and removing outliers (n=140) – on the log ratio of the distance to the frontier to the distance to Africa ($\ln(d^F_{i,t}/d^A_i)$) as included in equation (7) for twenty Brazilian states ($i \in \{1,...,20\}$) over fourteen time periods (*t*∈*{1545,1570,1585, 1625,1675,1725,1775,1800,1825,1850,1875,1900,1950,2000}*).

Figure V. The Geography of Race in Brazil, 1545-2000



The figure implies that, all else equal, the further a state is from the frontier relative to Africa (moving rightward on the x-axis), the greater the share of African Americans in the population (moving upward on the y-axis). Conversely, the closer a state is to the frontier relative to Africa, the lower the percentage of Africans. Consider the observation for Pernambuco in 1900 (-1.5, 1.9): the capital is nearly five times closer to the frontier than to Africa, but has almost seven times more African than Native inhabitants. This figure suggests that there is a positive linear relationship between the log ratio of 'Africans' to 'Natives' and the inverse log ratio of their "prices." If this relationship is linear, as it appears, then it is appropriate to use OLS to estimate the following equation:

$$(8) \qquad \ln(A_{i,t}/N_{i,t}) = c + \varphi \ln(d^{F}_{i,t-1}/d^{A}_{i}) + \boldsymbol{\Phi}\,(\mathbf{X}_{i,t}) + \varepsilon_{i,t}$$

This equation says that the log ratio of African to Native laborers in state i at time t ($\ln(A_{i,t}/N_{i,t})$) is equal to a constant (c), plus a coefficient (φ) times the log ratio of the distance to the frontier to the distance to the slave ports in Africa ($\ln(d^{F}_{i,t-1}/d^{A}_{i})$), plus the product of a diagonal matrix of coefficients ($\boldsymbol{\Phi}$) and a vector of control variables ($\boldsymbol{X}$), plus an error term ($\varepsilon_{i,t}$). Note that the distance term is lagged one year but, since the movement of the frontier in any given year is negligible, the value of the variable for the current year will be used in all specifications.

Equation (7) suggests that the coefficient φ is equal to one; however, this optimality condition is drawn from a model in which each state specializes in either mining or plantation agriculture. If the economic activity was principally subsistence farming, then there should be no relationship ($\varphi=0$) other than that stemming from internal migration of Africans from neighboring states relative to the initial number of Natives. Therefore, to identify the 'true' relationship between the race ratio ($A_{i,t}/N_{i,t}$) and the distance ratio ($d^{F}_{i,t-1}/d^{A}_{i}$) using equation (8), the vector $\boldsymbol{X}$ must include controls for whether a given state is suited to mining and/or plantation agriculture and its initial Native population (as well as the values of these variables in neighboring states). Recall that these are the factors that Sokoloff and Engerman (2000) identify as key to explaining the adoption of slavery.

In the case of Brazil, history suggests that the racial diversity of a settlement, i.e. the settlement's reliance on slave labor, can be partially explained by the suitability of land for growing sugarcane or the number of gold deposits in that area. The specific control variables on the right-hand side of equation (8) are the percent of land in each state suitable for growing sugarcane and the number of gold deposits per 1,000 km^2. These are cross-section fixed effects that isolate whether or not slavery was feasible (whether the relationship in equation (8) would hold in a given state) and what the relative productivity difference between sectors is (part of the constant in equation (8)). The former effect would influence the coefficient estimate; the latter, the constant.

The data are contemporary observations by state, but this is not problematic. With remote-sensing imagery, it has only recently become possible to estimate the prevalence of gold deposits, whether or not they have been exploited. This control captures whether or not there is or was gold in a given state, but not the number of gold mines that exist or existed; therefore, the observations will be uncorrelated with the error term – it is an exogenous endowment. The data on the location of gold deposits was obtained from the Brazilian Ministry of Mines and Energy online database (GEOBANK).

A similar criticism of the second control would be that the amount of land suitable for growing sugarcane in each state changes over time. The criteria used to classify land as suitable for growing sugarcane are: 1. the temperature must be at least 59°F to 60.8°F (15-16°C), the average annual rainfall must be between 47.24" and 59" (1,200-1,500mm), and the land must be somewhat level (Blume, 1985). Although Brazil has suffered substantial deforestation and erosion over the centuries, it is unlikely that the temperature, precipitation, or gradient isoquants have shifted significantly (global warming and the natural heating/cooling cycle of the Earth may have shifted the average temperature by only a few degrees over this period). The percentage of land in each state satisfying these criteria was calculated by geo-referencing, overlaying, and querying temperature, precipitation, and vegetation maps from the CIA World Factbook.

The data for the initial Native population of each state in 1500 was calculated using a digitized version of an indigenous population density map by Julian Steward (1946); this was described above and shown in Appendix III. The average population density by state was calculated by overlaying contemporary political boundaries on the population density polygons; the headcount, in turn, is the average population density per state times its area. Descriptive statistics for the percent of land suitable for growing sugar-cane, the number of gold deposits per 10,000 km^2, and the initial Native population for each state are presented along with those for the race ratio and labor cost ratio in Appendix V.

The results of the OLS estimation of equation (8), with and without controls for the three factor endowments, are shown in columns one, three, and four of Table III. Note that by taking the natural logs of the dependent and main explanatory variables, the coefficient estimate (φ) can be interpreted as an elasticity. For the other variables, the semi-log specification means that the coefficient estimates indicate the percentage change in the dependent variable in response to a one-unit increase in the control variable.

Column two separately estimates the coefficients on the log distance to the frontier and the log distance to Africa in order to test the restriction that the coefficient on the log of the ratio of the two variables is the same. The initial population density is added separately from the first two controls in column four; even though the physical geography may be truly exogenous, the population density is not.[19] Thus, there is the potential for slight collinearity or correlation with the error term, and the two columns are separated to evaluate this possibility.

The logic presented here focuses on the movement in the frontier as the driving factor behind changes in the racial composition of Brazilian states. Consider this variable for a moment: The frontier, for the most part, is increasingly distant over time – as the frontier shifted, a plethora of world-changing events were occurring in Europe and elsewhere, beginning with the Enlightenment and then the Industrial Revolution. It is possible that

[19] Humans choose where to live, and that choice depends on the physical geography. In addition, the initial population density estimates, combined with historical mortality rates, were used to calculate some missing values (15.7%) of the dependent variable.

the shifting frontier is simply capturing this historical momentum. Accordingly, the rapid expansion of the slave trade could be related to some overlooked processes evolving in Africa and Europe over time, such as the use of progressively larger slave ships or lower mortality during the passage, not the distance ratio itself. Therefore, the estimation will also explicitly control for time; these results are shown in column five of Table III.

One major political change that would limit the explanatory power of this model is the forced end of the international slave trade by England in 1850 (Bethell, 1970). When new slaves could not be imported, the relative movement of the frontier would be less closely related to the African population. To test this hypothesis, a dummy for the "end of the slave trade" is added in column six.

Finally, the data for this study comes from contiguous states, so it is plausible that the racial composition of any state be influenced by the extent of slavery in its neighbors. For example, if a state has no gold mines but an adjacent state has many, then that state would be likely to have a larger-than-expected African population because of the migration of manumitted slaves or free children of slaves. If this logic holds, then the relationship between the proxy for the labor cost ratio and the ratio of Africans to Natives, controlling for factor endowments, would be underestimated. To control for the presence of spatial autocorrelation, the model was re-estimated with additional controls for the average number of gold mines (per 1,000 km^2) and the average sugarcane suitability (%) in all adjoining states. This specification is in column seven of Table III.

The model was estimated using ordinary-least squares (OLS). Looking at Figure V, however, indicates that heteroskedasticity is present in the data – the average vertical distance between observations is greater further back in time. This was expected given the imprecision of early population statistics. The result is that the standard errors are biased upwards, an issue that may be resolved by using White period standard errors and covariance – the appropriate treatment given an unknown form of heteroskedasticity. By correcting the standard errors, the main difference is that the coefficient on the distance ratio in column five in Table III becomes statistically significant at standard levels.

Table III+. Results for Racial Composition and Labor Cost (Distance Ratio)							
Dependent Variable: ln(# African (A) / # Native (N)) *Mulatto Included; Outliers Omitted*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Dist. Frontier/ Dist. Africa)	2.58***		2.60***	2.50***	0.46*	0.72**	0.79**
	(0.29)		(0.26)	(0.25)	(0.21)	(0.22)	(0.25)
ln(Dist. Frontier)		2.87***					
		(0.27)					
ln(Dist. Africa)		0.97					
		(1.10)					
Gold Deposits (1,000km^2)			1.59**	1.62**	1.43***	1.27***	1.24***
			(0.55)	(0.51)	(0.28)	(0.28)	(0.19)
% Suitable for Sugarcane			0.72	0.42	-0.24	-0.23	-1.96*
			(0.75)	(0.70)	(0.45)	(0.47)	(0.79)
Initial Native Population				-2.6E-6	-8.3E-6***	-7.8E-6***	-8.2E-6***
				(3.2E-6)	(1.3E-6)	(1.4E-6)	(1.5E-6)
Time					0.01***	0.02***	0.02***
					(0.002)	(0.002)	(0.002)
End of Slave Trade Dummy						-1.47***	-1.49***
						(0.36)	(0.37)
Avg. Gold Deposits Neighboring States							0.45
							(0.75)
Avg. Sugarcane Suitability Neighboring States							2.22*
							(0.90)
R-Squared	0.36	0.40	0.41	0.41	0.62	0.64	0.66
No. Observations	140	140	140	140	140	140	140

+Results from ordinary least squares.
++White period standard errors in parentheses. ***P-value < 0.001; **P-value < 0.01; *P-value < 0.05.

Column one quantifies what is very clear visually in Figure V: the relationship between the race ratio ($A_{i,t}/N_{i,t}$) and distance ratio ($d^F{}_{i,t}/d^A{}_i$) is positive and statistically significant. Additional information provided by the regression is that the elasticity is 2.6. This suggests that for the average state, a 1% increase in the cost of Native labor (a 4.6 km shift in the frontier) corresponded to a 2.6% increase in the ratio of Africans to Natives (an additional 18,114 Africans, all else equal). Furthermore, this coefficient is significant at the one-tenth of one percent (0.001) level.

Across all specifications, the coefficient on the price ratio ranges from 0.46 to 2.60 and is significant at the five-percent level in all seven estimations. Is this coefficient significantly different than one? In columns one, three, and four, the coefficient is significantly *greater* than one at conventional significance levels; however, in columns five, six, and seven, one fails to reject the null hypothesis that the coefficient on the price ratio is equal to one, as suggested by the theoretical model.

Column two demonstrates that the coefficients on the log distance to the frontier and the log distance to Africa, when estimated separately, are not significantly different. That is, one fails to reject that the coefficients on the two variables are equal and estimating a single coefficient (φ) on the log distance ratio is not a binding restriction on the model. There is more variation in the distance to the frontier variable than in the distance to Africa variable because the latter is constant for each state; hence, it is not surprising that the coefficient on the distance to the frontier is more precisely estimated.

In column three, the magnitude of the coefficient on the distance ratio increases after controlling for the two factor endowments. The coefficients on “gold deposits” and “sugarcane suitability” are positive, as hypothesized, because slave labor should only be a feature of the mining and plantation sectors in Brazil. Only the coefficient on the gold deposits variable is statistically significant. When deposits of *all* minerals are controlled for, the coefficients on this alternative variable are also significant and don’t alter the other relationships. In this case, and when other measures of gold deposits are used, the coefficient on sugarcane suitability is positive and significant (not presented here).

A control for the initial Native population is added in column four. In accord with the theory, in areas with a higher initial number of Natives, one observes a lower ratio of Africans to Natives. This effect is highly significant in most specifications, but its magnitude is small. This may be because the total initial Native population in Brazil was about 1.7 million (according to Steward, 1946) and was reduced by an estimated eighty percent by 1950 (Rosenblat, 1954), compared with 18.5 million Brazilians who identified themselves as “Black” or “Mulatto” that year (T. Lynn Smith, 1972). Although the average effect is small, in some regions (with poorer endowments) it is still the key determinant of racial composition. Another interesting result is that the coefficient on sugarcane suitability drops when the initial population is controlled for – this confirms that there is slight collinearity between the two variables, but this is not a large concern because the coefficients and errors of the other variables are essentially the same.

Proceeding to column five, controlling for time causes the coefficient on the distance ratio to drop substantially but it remains significant. Time is the only variable that can cause the coefficient on the log distance ratio to become insignificant; however, after adding a dummy for the effective end of the slave trade in 1850, the coefficient on the variable of interest increases substantially and is highly significant, lending credence to the argument that the end of slavery effectively "broke" the structural relationship between these two variables. Note that the explanatory variables capture a large portion of the total variation in the race ratio within this sample (R-squared = 0.64).

Finally, column seven controls for spatial autocorrelation: the overlap effect of slavery in adjacent states that obscures the relationship between geography, prices, and racial composition. In this regression, it is clear that the coefficient estimate on the distance ratio is biased downward in column six. The coefficients on the factor endowments become smaller; however, neighbor states' factor endowments have a moderately strong relationship to the race ratio. Interestingly, while the presence of gold was significant in all specifications, the average gold deposits in neighboring states is not; rather, the coefficient on the neighbor's sugarcane suitability variable is positive and significant.

In sum, the results provide strong evidence that variation in racial composition across states in Brazil is a result of shifting labor costs and initial factor endowments. The relative cost of labor had a large and very significant influence on the relative size of African and Native populations in Brazil. Several potential criticisms of the model were addressed and building these alternative explanations into the model did not significantly alter the results; in fact, they added insight and improved the model's explanatory power.

The results with four alternative combinations of dependent and independent variables are presented in Appendix VI to demonstrate the robustness of those in Table III.

VII. Additional Robustness Checks: Native American Population Estimates

The demographic database outlined in Section V included the estimation of 49 out of 312 observations on the number of Natives and Africans, or 15.7 percent of the total. The inferred data covers the Native American population, when other data is absent, for the sixteenth, seventeenth, and eighteenth centuries, and the turn of the twentieth century. To recap, the 49 observations were calculated based on the following assumptions: If the natural rate of increase of the acculturated population was zero or negative, then the population of Natives in a settlement in any given time period would be limited to the Native peoples in newly-settled areas in that region.

It is generally accepted that the Native population in Brazil has steadily declined to about 1950, contrary to the observed decline and rebound in the population of North American Natives (Ubelaker, 1992). If one accepts Steward's (1946) estimates of the population density of Natives in Brazil in 1500, then there were a total of 1,635,000 Natives within Brazil's contemporary political boundaries. The intersection of state boundaries with the areas settled by the Portuguese over time (Hemming, 1978) yields a set of polygons, each corresponding to the new area settled in each state-period. In turn, the intersection of these polygons with the initial population densities, multiplied by their area, gives their estimated Native headcount in 1500. Thus, the potential Native population in each settlement in subsequent time periods is entirely dependent on the assumptions one makes about the growth rate of the Native population over time.

The estimated data included in the demographic database in Section V and in the estimation and results presented in section VI were based on Rosenblat's (1954) figures for the decline of the Native population of Brazil from 1492 to 1950. He estimated that the Native population, as a percent of the 1492 population, was 80% by 1570, 70% by 1650, 36% by 1825, 20% by 1940, and 20% by 1950. This corresponds to an average annual population growth rate of -0.37%.[20] The robustness of the results to changes in the natural rate of increase of the Native population will now be investigated.

Consider a range of Native population growth rates, ranging from positive ten percent to negative ten percent. If the annual natural rate of increase of the Native population over this period were positive one percent, than an initial Native population of 100,000 in 1500 would have increased to 14.5 million by 2000. Conversely, if the average rate of decline of the Native population were negative one percent, the same initial Native population would have decreased to 657 by 2000. Any growth rate outside of this range would correspond to an explosive increase or decrease of the population that would challenge the conclusions of the model.

If the Native population were to have positive growth, one would not expect the widespread importation of African slaves, all-else-equal. On the other hand, if the Native population were to decline precipitously, then one would not expect the settlement of the interior to contribute as much to the domestic labor force; i.e. the distance ratio would explain less of the racial composition over time because African slaves would have been imported on a large scale very early-on. Therefore, one would expect an upper-bound Native population growth rate beyond which the distance ratio becomes insignificant in explaining the racial composition of each state over time.

Consider the third specification in Table III where the log race ratio is regressed on the log distance ratio, the prevalence of gold, and sugarcane suitability. Using a rate of decline consistent with Rosenblat of -0.1%, the coefficient estimate is 2.6 with a standard error of 0.26. By adding the outliers back in (such that n=156) and repeatedly estimating this specification while varying the growth rate of the Native population, the robustness of the model to changes in the estimated data can be tested. The coefficient estimates, ordinary standard errors, and R-squared for this specification over a range of potential Native population growth rates are shown in Table IV.

[20] The average population growth rate is the coefficient on time in a regression where the dependent variable is the log percent of the initial population and the independent variable is time.

Table IV. Robustness of Results to Native Population Growth Rates[+]			
Growth Rate	Coefficient	Std. Error	R-Squared
+10%	-1.02	1.68	0.004
+5%	0.82	0.83	0.008
+4%	1.20	0.67	0.023
+3%	1.58*	0.51	0.063
+2%	1.96*	0.38	0.16
+1%	2.35*	0.29	0.31
+0.1%	2.71*	0.31	0.344
-0.1%	2.79*	0.32	0.336
-1%	3.14*	0.43	0.27
-2%	3.55*	0.58	0.20
-3%	3.95*	0.76	0.16
-4%	4.36*	0.94	0.13
-5%	4.78*	1.13	0.11
-10%	6.92*	2.12	0.07

[+]Results from specification in column (3) of Table III, Section VI, with ordinary standard errors: $\ln(\#African/\#Native) = c + \varphi \ln(Dist.\ Frontier/Dist.\ Africa) + \zeta_1 Gold + \zeta_2 Sugarcane$. *Significant at 1%.

The results in Table IV suggest that only by assuming an unrealistically high natural rate of increase of the Native population, of four percent or more, does the distance ratio cease to be significant. As one relaxes the assumption of a relatively large negative population growth rate (moving up from the bottom of Table IV), the fit of the model improves and peaks at a growth rate between plus and minus one-tenth of one percent. The distance ratio continues to be significant even at moderate positive natural rates of increase. It is a real possibility that some tribes in the interior could have escaped population shocks until well after 1500, and it's not to say that other contacted tribes could not have had brief periods of growth. Nonetheless, an average growth rate of almost four percent, the level above which the distance ratio is insignificant, is highly unlikely and contradicts available evidence.

VIII. Conclusion

This paper started with a simple but perplexing question: If the Americas were completely inhabited by Native Americans five hundred years ago, why does the racial composition of countries differ so much today? The answer proposed by economists and historians is that in areas favorable for extracting minerals (e.g. gold) or producing staples (e.g. sugar), high demand for labor and low supply led to the widespread adoption of slavery. Slaves were either Native American or African; however, since Native American lands were directly settled by Europeans, the Natives were disproportionately affected by disease, warfare, and enslavement. As Native populations declined, African labor became predominant in the mines and on the plantations.

This paper began by placing the economic organization of the American colonies within the context of European worldwide expansion that started in the thirteenth century. A simple profit-maximization problem was presented in which the prices of two types of labor are given; in a twist on the standard result, the optimal ratio of labor inputs is inversely related to the ratio of exogenous prices, not wages. It was the importation of labor into the colonies that would influence their racial composition ever since. The prices of the two types of labor, Africans and Natives, are assumed to be proportional to the location of each state relative to the frontier and the slave ports of Africa.

To test the predictions of the model, a new dataset on the racial composition of Brazil by state was compiled that spans the history of Brazil from its discovery in 1500 to the present. Given the limited amount of demographic data available for colonial Brazil, the dataset was complemented by using novel methods of estimating the population density of Native Americans and Africans based on ethnohistorical evidence, and the writings of chroniclers and slave import records, respectively. The inferred data comprises only about ten percent of the database, but it permits the construction of indices of race. Also, new data series detailing spatial variation in gold deposits, the suitability of land for growing sugarcane, and the relative distance to the sources of labor were constructed.

As hypothesized, the (inverse) labor cost ratio is significantly related to the ratio of Africans to Natives, controlling for factor endowments. Controlling for additional variables, such as time, output prices, the end of the slave trade, or spatial autocorrelation does not eliminate this relationship. Across all seven specifications, the ratio of the distance to the frontier to the distance to Africa is positively correlated to the racial composition of states over time and highly significant. The results fail to discredit the mechanism suggested in the formal model. They are strengthened by restricting the dataset to published statistics and/or omitting the mulatto population. Furthermore, they are robust to alternative assumptions regarding the rate of decline of the Native population and different measures of the distance to the frontier or to Africa.

The evidence supports the overall theory that the origins of the distribution of race in the New World are in the colonial economy. Across the sample of Brazilian states, it was found that a 1% increase in the distance to the frontier was associated with a 2.6% rise in the ratio of Africans to Natives. Future work will continue to investigate the potential effects of competition between sectors over labor in the colonial economies, expand the analysis to additional countries, and, in general, try to better understand the determinants and consequences of racial fractionalization within the context of historical development.

Works Cited

Alden, Dauril. "The Population of Brazil in the Late Eighteenth Century: A Preliminary Study." *Hispanic American Historical Review* 43.2 (1963): 173-205.

Alden, Dauril. Colonial Roots of Modern Brazil. Berkeley: U California P, 1973.

Alesina, Alberto, and Eliana La Ferrara. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature*, 43.3 (2005): 762-800.

Alesina, Alberto, Baqir, Reza, and William Easterly. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics*, 114.4 (1999): 1243-1284.

Bergad, Laird W. The Comparative Histories of Slavery in Brazil, Cuba, and the United States. New York: Cambridge UP, 2007.

Bethell, Leslie. The Abolition of the Brazilian Slave Trade: Britain, Brazil and the Slave Trade Question, 1807-1869. Cambridge: Cambridge UP, 1970.

Bethell, Leslie. The Cambridge History of Latin America. Vol. II. "Colonial Latin America." Cambridge: Cambridge UP, 1984.

Bethell, Leslie and José Murilo de Carvalho. "1822-1850." Brazil: Empire and Republic, 1822-1930. Cambridge: Cambridge UP, 1989. 45-112.

Blume, Helmut. Geography of Sugar Cane: Environmental, Structural, and Economical Aspects of Sugarcane Production. Berlin: Verlag Dr. Albert Bartens, 1985.

Boxer, C. R. The Golden Age of Brazil. Berkely: California UP, 1962.

Brito Freyre, Francisco de. Historia da Guerra Brasilica. Recife: Governo de Pernambuco, 1977.

Calmon, Pedro. Historia do Brasil. Vol. 1 "As Origens (1500-1600)." São Paulo: Companhia Editora Nacional, 1939.

Calogeras, Pandiá. Formação Historica do Brasil. 3rd Ed. São Paulo: Companhia Editora Nacional, 1938.

CIA World Factbook. Brazil Maps. <https://www.cia.gov/library/publications/the-world-factbook/print/br.html>

De Abreu, Capistrano. Chapters of Brazil's Colonial History, 1500-1800. Oxford: Oxford UP, 1997.

Do Nascimento, Abdias. "The African Experience in Brazil." African Presence in the Americas. Ed. Carlos Moore. Trenton: Africa World P, 1995.

Easterly, William. "Inequality Does Cause Underdevelopment: Insights from a New Instrument." *Journal of Development Economics*, 84.2 (2007): 755-776.

Easterly, William and Ross Levine. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics*, 112.4 (1997): 1203-1250.

Elliott, J. H. "The Spanish Conquest and Settlement of America." The Cambridge History of Latin America. Vol. 1. Cambridge: Cambridge UP, 1984. 149-206.

Galeano, Eduardo. Las Venas Abiertas de América Latina. México, DF: Siglo Veintiuno Editores, 1971.

Goldin, Claudia and Kenneth Sokoloff. "The Relative Productivity Hypothesis of Industrialization: The American Case, 1820 – 1850." *Quarterly Journal of Economics*, 99.3 (1984): 461-487.

Grubb, Farley. "The Transatlantic Market for British Convict Labor." *Journal of Economic History*, 60.1 (2000): 94-121.

Hemming, John. Amazon Frontier: The Defeat of the Brazilian Indians. Cambridge: Harvard UP, 1987.

Hemming, John. Red Gold: The Conquest of the Brazilian Indians, 1500-1760. Cambridge: Harvard UP, 1978.

Instituto Brasileiro de Geografia e Estadística (IBGE). Séries Estatísticas & Séries Históricas. <http://www.ibge.gov.br/series_estatisticas/>

Johnson, H.B. "A Preliminary Inquiry into Money, Prices, and Wages in Rio de Janeiro, 1763-1823." Colonial Roots of Modern Brazil. Ed. Dauril Alden. Berkeley: U California P, 1973. 231-283.

Johnson, H.B. "Portuguese Settlement, 1500-1580." Colonial Brazil. Ed. Leslie Bethell. Cambridge: Cambridge UP, 1987. 1-38.

Klein, Herbert S. and Francisco Vidal Luna. Slavery in Brazil. New York: Cambridge UP, 2010.

Klein, Herbert S. The Atlantic Slave Trade. New York: Cambridge UP, 2010.

Langfur, Hal. The Forbidden Lands: Colonial Identity, Frontier Violence, and the Persistence of Brazil's Eastern Indians, 1750-1830. Stanford: Stanford UP, 2006.

Magalhães de Gândavo, Pêro de. Tratado do Provincia do Brazil. Rio de Janeiro: Instituto Nacional do Livro, 1965.

Marchant, Alexander. Do Escambo á Escravidão: As Relações Economicas de Portugueses e Indios na Colonição do Brasil, 1500-1580. São Paulo: Companhia Editora Nacional, 1980.

Marchant, Alexander. From Barter to Slavery: The Economic Relations of Portuguese and Indians in the Settlement of Brazil, 1500-1580. Baltimore: John Hopkins P, 1942.

Marcílio, Maria Luiza. "The Population of Colonial Brazil." The Cambridge History of Latin America. Vol. 2. Ed. Leslie Bethell. Cambridge: Cambridge UP, 1984. 37-66.

Mauro, Frédéric. "La Expansión Europea desde los Orígenes hasta Finales del Siglo XV." Historia General de América Latina. Vol. 2. París: Ediciones UNESCO, 2000. 49-70.

Mauro, Frédéric. "Political and Economic Structures of Empire, 1580-1750." Colonial Brazil. Ed. Leslie Bethell. Cambridge: Cambridge UP, 1987. 39-66.

Melville, Elinor G. K. "Land Use and the Transformation of the Environment." The Cambridge Economic History of Latin America. Vol. 1. New York: Cambridge UP, 2006. 109-142.

Metcalf, Alida C. Family and the Frontier in Colonial Brazil. Berkeley. U California P, 1992.

Miguel, Edward and Mary Kay Gugerty. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics*, 89.11-12 (2005): 2325-2368.

Ministry of Mines and Energy, Government of Brazil. GEOBANK. <http://geobank.sa.cprm.gov.br/>

Moya Pons, Frank. "El Caribe, Tierra Firme, Darién y Centroamérica." Historia General de América Latina. Vol. 2. París: Ediciones UNESCO, 2000. 109-134.

Nunn, Nathan. "Slavery, Inequality, and Economic Development in the Americas: An Examination of the Engerman-Sokoloff Hypothesis." *Working Paper* (2007).

Poppino, Rollie E. Brazil: The Land and People. New York: Oxford UP, 1973.

Prado Jr., Caio. The Colonial Background of Modern Brazil. Berkeley: U California P, 1967.

Restall, Matthew. Beyond Black and Red: African-Native Relations in Colonial Latin America. Albuquerque: U New Mexico P, 2005.

Ribeiro, Darcy. The Brazilian People: The Formation and Meaning of Brazil. Gainesville: UP of Florida, 2000.

Rönnbäck, Klas. "The Price of Sugar in Sweden: Data, Sources, and Methods." *Göteborg Papers in Economic History*, 10 (2007).

Rosenblat, Angel. La Población Indígena, 1492-1950. 2 vols. Buenos Aires: Editorial Nova, 1954.

Russell-Wood, A.J.R. Society and Government in Colonial Brazil, 1500-1822. Brookfield: Variorum, 1992.

Russell-Wood, A.J.R. "The Gold Cycle, c. 1690-1750." Colonial Brazil. Ed. Leslie Bethell. Cambridge: Cambridge UP, 1987. 190-243.

Schwartz, Stuart B. "Free Labor in a Slave Economy." Colonial Roots of Modern Brazil. Ed. Dauril Alden. Berkeley: U California P, 1973. 147-198.

Schwartz, Stuart B. "Plantations and Peripheries, c. 1580 – c. 1750." Colonial Brazil. Ed. Leslie Bethell. Cambridge: Cambridge UP, 1987. 67-144.

Smith, T. Lynn. Brazil: People and Institutions. Baton Rouge: Louisiana State UP, 1972.

Sokoloff, Kenneth L. and Stanley L. Engerman. "History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives*, 14.3 (2000): 217-232.

Stein, Stanley J. and Barbara H. Stein. The Colonial Heritage of Latin America. New York: Oxford UP, 1970.

Steward, Julian H. Handbook of South American Indians. 7 vols. Washington: Government Printing Office, 1946.

Symcox, Geoffrey and Blair Sullivan. Christopher Columbus and the Enterprise of the Indies: A Brief History with Documents. New York: Palgrave Macmillan, 2005.

Ubelaker, Douglas H. "North American Indian Population Size: Changing Perspectives." Disease and Demography in the Americas. Ed. John W. Verano and Douglas H. Ubelaker. Washington: Smithsonian Institution P, 1992. 169-176.

Varnhagen, Francisco Adolfo de. História Geral do Brasil. 5th Ed. São Paulo: Edições Melhoramentos, 1956.

Webster's Universal College Dictionary. New York: Random House, 2004.

Appendix I. Model

The dynamic optimization problem for each economy i ($i \in \{P,M\}$) is

(1)

$$\text{Max}_{(K_{\text{it}}, L_{\text{it}}, A_{\text{it}}, N_{\text{it}}, T_{\text{it}} \geq 0)}$$

$$\sum_{t=0}^{\infty} \rho_i{}^t (p_{\text{it}}\, \Lambda_{\text{it}}\, F(K_{\text{it}},\, L_{\text{it}},\, A_{\text{it}},\, N_{\text{it}},\, T_{\text{it}}) - r_t\, K_{\text{it}} - w_t\, L_{\text{it}} - p_t^A\, I_{\text{it}}^A - p_t^N\, I_{\text{it}}^N)$$

subject to

(2)

$$Y_{\text{it}} = \Lambda_{\text{it}}\, F_{\text{it}}[K,\, L,\, A,\, N,\, T] = \Lambda_{\text{it}}\, K_{\text{it}}^{\alpha_i}\, L_{\text{it}}^{\beta_i}\, A_{\text{it}}^{\gamma_i}\, N_{\text{it}}^{\tau_i}\, T_{\text{it}}^{\mu_i}$$

$$A_{i,t+1} = I_{\text{it}}^A + (1 + a_i)\, A_{\text{it}}$$

$$N_{i,t+1} = I_{\text{it}}^N + (1 + n_i)\, N_{\text{it}}$$

The Lagrangean is

(3)

$$\begin{aligned} \mathcal{L} = \max_{(K_{\text{it}}, L_{\text{it}}, A_{\text{it}}, N_{\text{it}}, T_{\text{it}} \geq 0)} & \\ \sum_{t=0}^{\infty} \rho_i{}^t \big\{ & p_{\text{it}}\, \Lambda_{\text{it}}\, K_{\text{it}}^{\alpha_i}\, L_{\text{it}}^{\beta_i}\, A_{\text{it}}^{\gamma_i}\, N_{\text{it}}^{\tau_i}\, T_{\text{it}}^{\mu_i} - r_{\text{it}}\, K_{\text{it}} - w_{\text{it}}\, L_{\text{it}} - \iota_{\text{it}}\, T_{\text{it}} \\ & - p_{\text{it}}^A\, (A_{i,t+1} - (1 + a_i)\, A_{\text{it}}) - p_{\text{it}}^N\, (N_{i,t+1} - (1 + n_i)\, N_{\text{it}}) \big\} \end{aligned}$$

The First-Order Conditions are

(4)

(i) $K_{\text{it}} : \alpha_i\; p_{\text{it}}\, \Lambda_{\text{it}}\, K_{\text{it}}^{\alpha_i - 1}\, L_{\text{it}}^{\beta_i}\, A_{\text{it}}^{\gamma_i}\, N_{\text{it}}^{\tau_i}\, T_{\text{it}}^{\mu_i} - r_{\text{it}} = 0$

(ii) $L_{\text{it}} : \beta_i\; p_{\text{it}}\, \Lambda_{\text{it}}\, K_{\text{it}}^{\alpha_i}\, L_{\text{it}}^{\beta_i - 1}\, A_{\text{it}}^{\gamma_i}\, N_{\text{it}}^{\tau_i}\, T_{\text{it}}^{\mu_i} - w_{\text{it}} = 0$

(iii) $T_{\text{it}} : \mu_i\; p_{\text{it}}\, \Lambda_{\text{it}}\, K_{\text{it}}^{\alpha_i}\, L_{\text{it}}^{\beta_i}\, A_{\text{it}}^{\gamma_i}\, N_{\text{it}}^{\tau_i}\, T_{\text{it}}^{\mu_i - 1} - \iota_{\text{it}} = 0$

(iv)
$$\begin{aligned} A_{i,t+1} : -\rho_i{}^t (p_{\text{it}}^A) + \rho_i{}^{t+1} \Big(& \gamma_i\, E_t\Big[p_{i,t+1}\, \Lambda_{i,t+1}\, K_{i,t+1}^{\alpha_i}\, L_{i,t+1}^{\beta_i}\, A_{i,t+1}^{\gamma_i - 1}\, N_{i,t+1}^{\tau_i}\, T_{i,t+1}^{\mu_i} \Big] \\ & + E_t[p_{i,t+1}^A](1 + a_i) \Big) = 0 \end{aligned}$$

(v)
$$\begin{aligned} N_{i,t+1} : -\rho_i{}^t (p_{\text{it}}^N) + \rho_i{}^{t+1} \Big(& \tau_i\, E_t\Big[p_{i,t+1}\, \Lambda_{i,t+1}\, K_{i,t+1}^{\alpha_i}\, L_{i,t+1}^{\beta_i}\, A_{i,t+1}^{\gamma_i}\, N_{i,t+1}^{\tau_i - 1}\, T_{i,t+1}^{\mu_i} \Big] \\ & + E_t[p_{i,t+1}^N](1 + n_i) \Big) = 0 \end{aligned}$$

Assume that prices (p_{it}, p_t^A , p_t^N) and technology (Λ_{it}) follow a random walk and the price of output (p_{it}) and the technology level (Λ_{it}) are independent; also, assume that they (p_{it}, Λ_{it}) are independent of the number of slaves ($A_{i,t+1}$, $N_{i,t+1}$). That is,

$E_t[p_{i,t+1}] = p_{\text{it}}$, $E_t[\Lambda_{i,t+1}] = \Lambda_{\text{it}}$, $E_t[p_{i,t+1}\,\Lambda_{i,t+1}] = E_t[p_{i,t+1}]\,E_t[\Lambda_{i,t+1}]$,

$\mathrm{E_t}[\mathrm{p^A_{t+1}}] = \mathrm{p^A_t}$, $\mathrm{E_t}[\mathrm{p^N_{t+1}}] = \mathrm{p^N_t}$

$$E_t\left[p_{i,t+1}\,\Lambda_{i,t+1}\,K_{i,t+1}^{\alpha_i}\,L_{i,t+1}^{\beta_i}\,A_{i,t+1}^{\gamma_i}\,N_{i,t+1}^{\tau_i}\,T_{i,t+1}^{\mu_i}\right] = E_t\left[p_{i,t+1}\,\Lambda_{i,t+1}\,K_{i,t+1}^{\alpha_i}\,L_{i,t+1}^{\beta_i}\,T_{i,t+1}^{\mu_i}\right]A_{i,t+1}^{\gamma_i}\,N_{i,t+1}^{\tau_i}$$

Substituting and rearranging, (iv) and (v) become

$$\rho_i\left(E_t\left[\gamma_i\,p_{i,t+1}\,\Lambda_{i,t+1}\,K_{i,t+1}^{\alpha_i}\,L_{i,t+1}^{\beta_i}\,A_{i,t+1}^{\gamma_i-1}\,N_{i,t+1}^{\tau_i}\,T_{i,t+1}^{\mu_i}\right] + E_t[p_{t+1}^A(1+a_i)]\right) = p_t^A$$

$$\gamma_i\,p_{\text{it}}\,E_t\left[\Lambda_{i,t+1}\,K_{i,t+1}^{\alpha_i}\,L_{i,t+1}^{\beta_i}\,A_{i,t+1}^{\gamma_i-1}\,N_{i,t+1}^{\tau_i}\,T_{i,t+1}^{\mu_i}\right] = \frac{p_t^A}{\rho_i} - p_t^A(1+a_i)$$

$$= \quad p_t^A\left(\frac{1}{\rho_i} - (1+a_i)\right) \quad = \frac{p_t^A(1-\rho_i-\rho_i\,a_i)}{\rho_i}$$

(5) $$\gamma_i\,E_t\left[\Lambda_{i,t+1}\,K_{i,t+1}^{\alpha_i}\,L_{i,t+1}^{\beta_i}\,T_{i,t+1}^{\mu_i}\right]A_{i,t+1}^{\gamma_i-1}\,N_{i,t+1}^{\tau_i} = \frac{p_t^A}{p_{\text{it}}}\,\frac{(1-\rho_i-\rho_i\,a_i)}{\rho_i}$$

$$\rho_i\left(E_t\left[\tau_i\,p_{i,t+1}\,\Lambda_{i,t+1}\,K_{i,t+1}^{\alpha_i}\,L_{i,t+1}^{\beta_i}\,A_{i,t+1}^{\gamma_i}\,N_{i,t+1}^{\tau_i-1}\,T_{i,t+1}^{\mu_i}\right] + E_t[p_{t+1}^N(1+n_i)]\right) = p_t^N$$

$$\tau_i\,p_{\text{it}}\,E_t\left[\Lambda_{i,t+1}\,K_{i,t+1}^{\alpha_i}\,L_{i,t+1}^{\beta_i}\,A_{i,t+1}^{\gamma_i}\,N_{i,t+1}^{\tau_i-1}\,T_{i,t+1}^{\mu_i}\right] = \frac{p_t^N}{\rho_i} - p_t^N(1+n_i)$$

$$= \quad p_t^N\left(\frac{1}{\rho_i} - (1+n_i)\right) \quad = \frac{p_t^N(1-\rho_i-\rho_i\,n_i)}{\rho_i}$$

(6) $$\tau_i\,E_t\left[\Lambda_{i,t+1}\,K_{i,t+1}^{\alpha_i}\,L_{i,t+1}^{\beta_i}\,T_{i,t+1}^{\mu_i}\right]A_{i,t+1}^{\gamma_i}\,N_{i,t+1}^{\tau_i-1} = \frac{p_t^N}{p_{\text{it}}}\,\frac{(1-\rho_i-\rho_i\,n_i)}{\rho_i}$$

Dividing (6) by (5) and rearranging

(7) $$\frac{A_{i,t+1}}{N_{i,t+1}} = \frac{p_{\text{it}}^N\,\gamma_i\,(1-\rho_i-\rho_i\,n_i)}{p_{\text{it}}^A\,\tau_i\,(1-\rho_i-\rho_i\,a_i)}$$

Appendix II. Data Sources and Construction of the Demographic Database

Knowledge of the population history of Brazil is limited for the sixteenth, seventeenth, and eighteenth centuries; hence, evidence from a variety of sources must be evaluated and combined. For this "pre-statistical period," the data is based on first-hand accounts, correspondence and legal documents, and modern scholarship using archival evidence. For the "statistical period," the nineteenth and twentieth centuries, the data sources are municipal, state, and national census records. The following section will describe the sources of the 'Base Data' – consisting of 147 observations on the "European" population, 135 observations on the "African" and "Mulatto" (post-1775) populations, and 107 observations on the "Native" population – across fourteen periods (1545, 1570, 1585, 1625, 1675, 1725, 1775, 1800, 1825, 1850, 1875, 1900, 1950, 2000).

These figures are not all reported here – a full discussion of the evidence for each period is provided in "The Demographic History of Brazil" (unpublished working paper). The data sources described below include the primary sources (and secondary sources when these provide additional information). For the national census data, secondary sources are adopted when the statistics are identically reported across multiple references, such as scholarly articles and books, or government publications. No inferences will be made for those settlements where figures are not based on confirmed primary documents.

An important clarification is that the modern territory of Brazil is not what 'Brazil' was hundreds of years ago. 'Brazil' in the first few centuries after contact is considered from a Eurocentric perspective as the colonial domain where European, African, and Native elements were integrated. Each colonial settlement's area has a correspondence (one-to-one or one-to-many) with a modern state defined by political borders; to make the data comparable, each settlement or captaincy is associated with the state in which it is chiefly located. The sources of data are discussed below in chronological order.

Beginning with the period ca. 1545, the European and Native populations of São Vicente are described in the letter of Luís de Góis to the King dated May 12, 1548 – also see

Marcilio (1984) and Calmon (1939). The European population of Bahia after re-settlement is indicated in *Documentos Historicos* (XXXVIII) – also see Johnson (1987) and Ribeiro (2000). Primary sources for the African population are *As Gavetas da Torre do Tombo* (Lisbon, 1962) and Paulo Meréa, *Historia da Colonia Portuguesa*.

The main source of data for the European population of Brazil ca. 1570 is the *Tratado da Província do Brasil* (Lisbon, 1576) written by Pêro de Magalhães de Gândavo (he traveled in Brazil from 1568 to 1571). The Native population is described in *Cartas Avulsas de Jesuitas, 1550-1568* and *Cartas, Informaçoes, Fragmentos Historicos e Sermões de Padre Joseph de Anchieta, S.J.* (1554-1594) – also see Marchant (1942). A reference for the number of African slaves in one captaincy is from Marcílio (1984).

The three primary references for the population of Brazil ca. 1585 are Joseph de Anchieta (*Cartas, Informações, Fragmentos Historicos e Sermões*), Fernão Cardim (*Narrativa Epístolar de uma Viagem e Missão Jesuítica*), and Gabriel Soares de Sousa (*Tratado Descritivo do Brasil em 1587*). Anchieta and Cardim were Jesuit administrators; Soares was an explorer and writer who settled in Brazil. This data covers most captaincies and all three racial groups.

The main source of data for the European population in the period ca. 1625 is Francisco de Brito Freyre, Governor of Pernambuco from 1661 to 1664 (*História da Guerra Brasílica*, Portugal, 1675). *Livro que Dá Rezão do Estado do Brasil* (1612) by Diogo de Campos Moreno, *Sargento-Mor* (Principal Sergeant), provides information for the Native and European populations of two captaincies. Some observations for the Native and African populations are available from Jesuits Luíz Figueira (1637) and António Vieira (1655) – also see Schwartz (1987) and Hemming (1987).

Statistics for the European and Native population ca. 1675 are from *Bandeirantes no Paraguai*, by Juan de Mongelos (a Spaniard living in São Paulo), *Anais Historicos* by Governor Bernardo Perreira de Berredo (1718-1722), *Regulamento das Aldeias* (1660) by Jesuit António Viera, and a collection of Jesuit documents in *Suma Histórica* (1938-

1950) by Serafim Soares Leite – also see Boxer (1962), Hemming (1978), and Schwartz (1987). Two officials in Maranhão – Maurício de Heriarte (1662) and Manuel da Vide Souto-Maior (1658) – describe the northern population. Also, another document mentioned by Varnhagen (1956) – from the secretary of the *Conselho Ultramarino* (Lisbon, 1674) – refers to Rio de Janeiro. The African population is partially described by these sources in addition to António Viera (see Goulart, 1975).

For the period ca. 1725, sources for the north-east include *Suma Histórica*, by Serafim Leite, and a letter from the Governor of Maranhão-Pará to the Secretary of State to the King (ca. 1750) – also see Schwartz (1987), Boxer (1962), and Hemming (1978). The data for the center-west comes from Andreoni's account *Cultura e Opulência do Brasil* (Lisbon, 1711) and slave matriculations for 1735 from the Municipal Library of São Paulo's *Codice Costa Matoso* – see Russell-Wood (1987), Poppino (1973), and Boxer (1962). For the south-east, the primary source is a letter from the Archbishop of Bahia to the Crown (1702), cited by Schwartz (1987) and Boxer (1962). The data is relatively scarce, but covers all three racial groups.

The data for the population of Brazil ca. 1775 was compiled by Dauril Alden (1963) in "The Population of Brazil in the Late Eighteenth Century: A Preliminary Study" (Hispanic American Historical Review, 43.2). His sources include state and municipal censuses and parish records. This data includes headcounts for most states, and a racial breakdown for about a third of the sample with notably less coverage of Natives.

The data for the period ca. 1800 is from Alden (1984), whose main source is an enclosure in Lord Strangford to the Marquis of Wellesley dated May 20, 1810; other sources include MacLachlan, "African Slave Trade," the 1800 census for Mato Grosso, and the 1803 census for Rio de Janeiro. Information for Pará comes from Hemming (1978) and for Paraíba, from Marcílio (1984) citing *Paraíba, Mapa de População, 1798*. This data is merged with the population statistics for nearby years from Klein and Vidal Luna (2010). They employ multiple primary sources (listed on p. 186), including archival references, IBGE (Brazilian Institute of Geography and Statistics) publications, and modern texts.

The data for Brazil ca. 1825 is from Joaquim Norberto de Souza e Silva, *População Geral do Império*, and Antonio Rodrigues Velloso de Oliveira, *Memória* –see Marcílio (1984) and Calogeras (1938). Complementary data for this period is from Goulart (1975), who cites the National Library archives, and Metcalf (1992), who cites the 1820 census records for Parnaíba. The statistics published by Klein and Vidal Luna (2010), from these and other sources, have the best coverage so they will be used as the sole data source for this period; please see their book for a detailed list of sources.

For the period ca. 1850, the data also comes from Klein and Vidal Luna (2010). In addition to Norberto de Souza e Silva (São Paulo, 1870), their sources include documents from Archivo Nacional do Rio de Janeiro (ANRJ) and Octávio Ianni: *As Metamorfoses do Escravo, Apogeu e Crise da Escravatura no Brasil Meridional* (São Paulo, 1962).

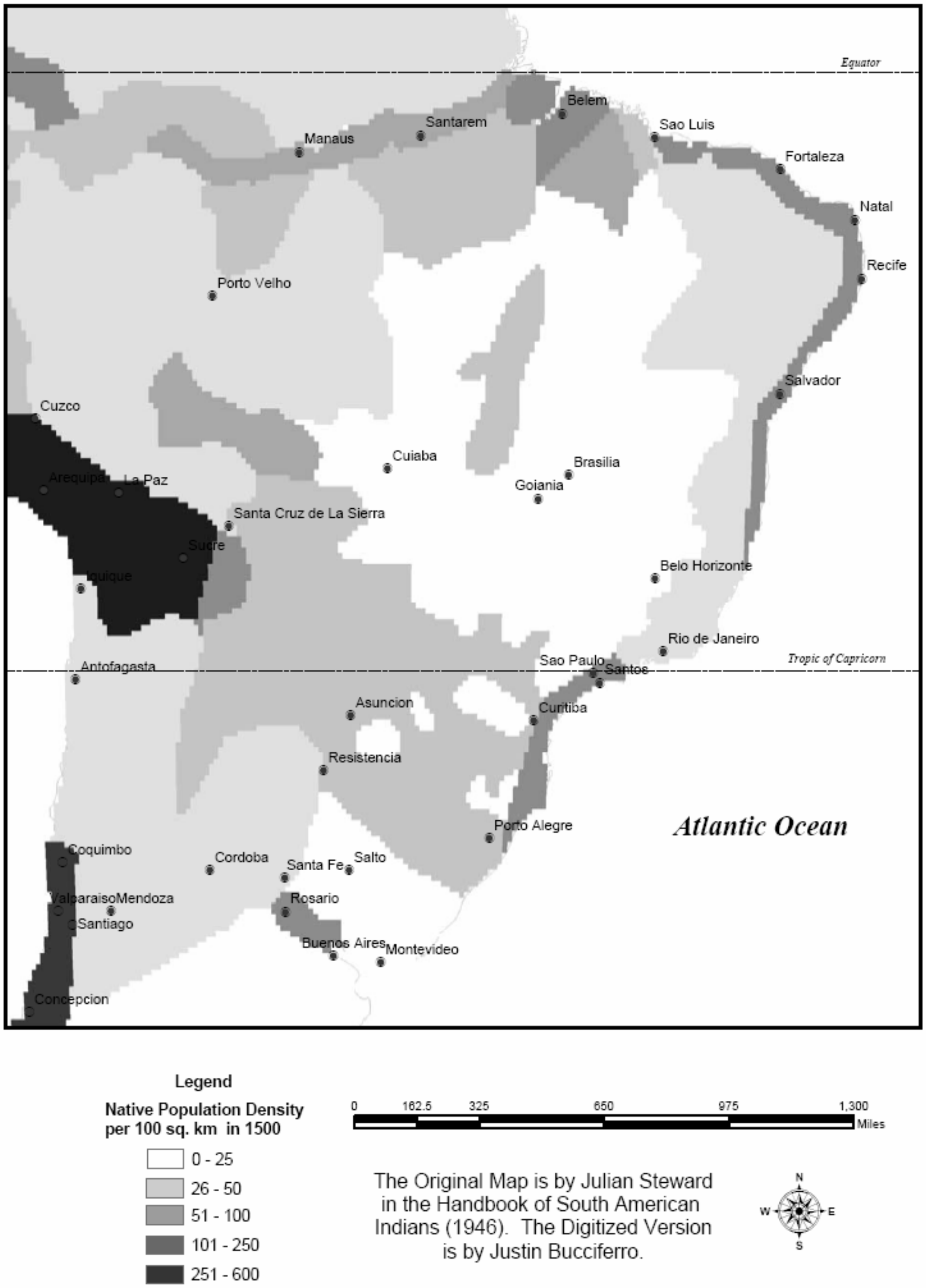
The 1872 census data is presented by Merrick and Graham (1979), *Population and Economic Development in Brazil: 1800 to the Present* and Klein and Vidal Luna (2010). The original data is available as *Census of 1872, Recenseamento da População do Império a que se Procedeu no dia 1 de Agosto de 1872* (Rio de Janeiro, 1873-1876).

The population statistics for the 1900 reference period are from the 1890 census: *Sexo, Raça e Estado Civil, Nacionalidade, Filiação Culto e Analphabetismo da População Recenseada em 31 de Dezembro de 1890* published by the Statistics Office of the federal government of Brazil (Rio de Janeiro, 1898) and available at http://biblioteca.ibge.gov.br.

The data for the total population by state and race in 1950 was consolidated from T. Lynn Smith (1972), *Brazil: People and Institutions*, based on the 1950 Demographic Census, *VI Recenseamento Geral do Brasil, 1950, I* (Rio de Janeiro, 1956) of the Brazilian Institute of Geography and Statistics (IBGE). The 'Native' figures are not reported.

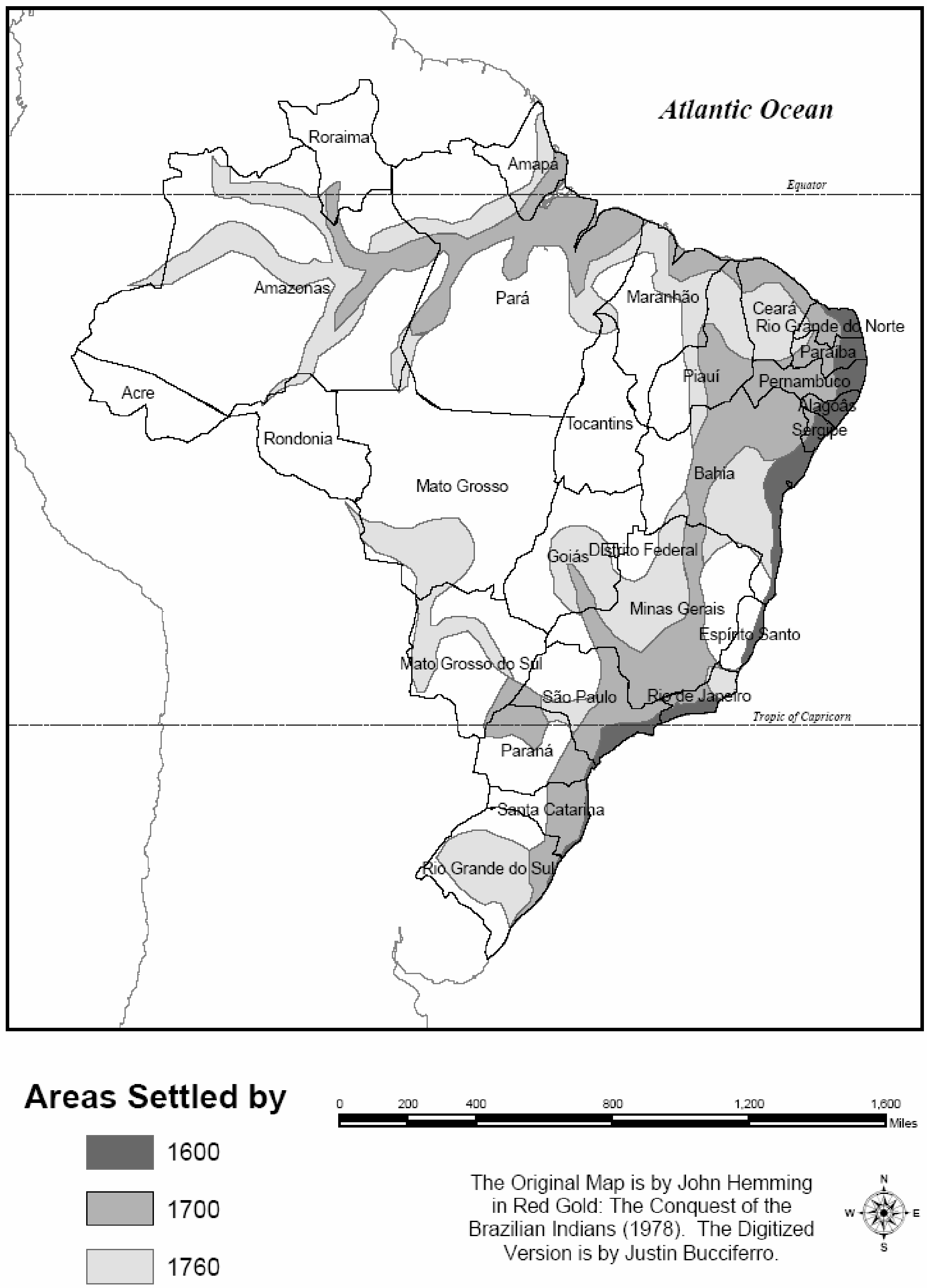
The source of data for the year 2000 is also IBGE (http://www.ibge.gov.br/series_ estatisticas/). The source publication is *Tendências Demográficas: Uma Análise dos Resultados da Amostra do Censo Demográfico 2000* (Rio de Janeiro, 2004).

Appendix III. Native Population Density in 1500, Brazil and Surrounding Areas



The Original Map is by Julian Steward in the Handbook of South American Indians (1946). The Digitized Version is by Justin Bucciferro.

Appendix IV. Areas of Brazil Settled During the First Three Centuries of Colonization



The Original Map is by John Hemming in Red Gold: The Conquest of the Brazilian Indians (1978). The Digitized Version is by Justin Bucciferro.

Appendix V. Descriptive Statistics

State (Capital)		Africans / Natives	Africans (Incl. Mulattos)/ Natives	Dist. Frontier / Dist. Africa	Gold Deposits (10,000 km^2)†	% Suitable for Sugarcane†	Initial Pop. Density (100km^2)†	Size (100km^2)†
1. Alagoâs (Maceió)	*Mean*	8.1	66.3	0.12	0.4	44.6	43.2	278
	Std. Dev.	6.2	91.1	0.10				
	# Obs.	4	4	14				
2. Amazonas (Manaus)	*Mean*	0.4	6.5	0.04	1.1	0.3	21.2	15,707
	Std. Dev.	0.4	8.6	0.00				
	# Obs.	5	5	14				
3. Bahia (Salvador)	*Mean*	64.9	275.9	0.09	3.7	30.3	19.2	5,647
	Std. Dev.	134.9	590.4	0.10				
	# Obs.	11	11	14				
4. Ceará (Fortaleza)	*Mean*	8.9	128.4	0.13	0.9	1.8	31.7	1,488
	Std. Dev.	14.0	213.9	0.14				
	# Obs.	3	3	14				
5. Espírito Santo (Vitória)	*Mean*	3.4	16.6	0.13	1.7	69.5	17.3	461
	Std. Dev.	4.9	39.6	0.10				
	# Obs.	9	9	14				
6. Goiás (Goiânia)	*Mean*	42.0	337.0	0.07	9.4	15.1	10.7	3,401
	Std. Dev.	61.1	673.8	0.05				
	# Obs.	6	6	14				
7. Maranhão (São Luis)	*Mean*	10.7	48.1	0.09	1.0	59.5	23.3	3,320
	Std. Dev.	19.1	87.8	0.09				
	# Obs.	9	9	14				
8. Mato Grosso (Cuiabá)	*Mean*	6.2	23.4	0.07	3.2	10.4	18.7	9,034
	Std. Dev.	5.2	21.5	0.03				
	# Obs.	7	7	14				
9. Minas Gerais (Belo Horizonte)	*Mean*	133.3	493.4	0.09	15.5	70.0	11.2	5,865
	Std. Dev.	227.3	807.4	0.07				
	# Obs.	7	7	14				
10. Pará (Belém)	*Mean*	1.7	18.5	0.09	5.1	0.2	24.8	12,477
	Std. Dev.	2.9	39.6	0.06				
	# Obs.	9	9	14				
11. Paraíba (João Pessoa)	*Mean*	4.3	31.5	0.13	3.7	10.8	27.7	564
	Std. Dev.	4.5	65.5	0.11				
	# Obs.	8	8	14				

† This variable is time-invariant and will serve as a control; therefore, the std. dev. and # obs. are not reported.

Appendix V. Descriptive Statistics (Continued)

State (Capital)		Africans / Natives	Africans (Incl. Mulattos)/ Natives	Dist. Frontier / Dist. Africa	Gold Deposits (10,000 km^2)†	% Suitable for Sugarcane†	Initial Pop. Density (100km^2)†	Size (100km^2)†
12. Paraná (Curitiba)	*Mean*	34.5	70.4	0.05	4.0	99.1	23.4	199
	Std. Dev.	45.2	89.8	0.02				
	# Obs.	5	5	14				
13. Pernambuco (Recife)	*Mean*	8.0	30.0	0.13	2.7	11.6	25.7	983
	Std. Dev.	13.0	43.7	0.11				
	# Obs.	10	10	14				
14. Piauí (Teresina)	*Mean*	41.6	291.7	0.07	0.0	13.0	16.1	2,515
	Std. Dev.	56.3	394.0	0.03				
	# Obs.	5	5	14				
15. Rio de Janeiro (Rio de Janeiro)	*Mean*	29.7	55.1	0.12	3.4	76.9	15.0	437
	Std. Dev.	21.4	55.2	0.07				
	# Obs.	8	8	14				
16. Rio Grande do Norte (Natal)	*Mean*	9.0	90.8	0.14	2.5	3.7	42.7	528
	Std. Dev.	15.3	200.1	0.12				
	# Obs.	6	6	14				
17. Rio Grande do Sul (Porto Alegre)	*Mean*	4.5	9.5	0.06	1.6	65.6	25.0	2,817
	Std. Dev.	5.3	13.4	0.03				
	# Obs.	5	5	14				
18. Santa Catarina (Florianópolis)	*Mean*	13.6	22.2	0.06	1.4	80.0	38.7	953
	Std. Dev.	15.5	22.5	0.03				
	# Obs.	6	6	14				
19. São Paulo (São Paulo)	*Mean*	11.4	28.6	0.08	0.0	90.3	20.2	2,482
	Std. Dev.	16.8	48.7	0.05				
	# Obs.	12	12	14				
20. Sergipe (Aracajú)	*Mean*	9.0	58.3	0.12	1.4	69.1	39.6	219
	Std. Dev.	7.0	71.4	0.11				
	# Obs.	5	5	14				
* Total	*Mean*	23.0	101.9	0.09	3.1	41.1	24.8	3,469
	Std. Dev.	70.0	309.1	0.08	3.6	34.4	9.9	4,340
	# Obs.	140	140	280	20	20	20	20

† This variable is time-invariant and will serve as a control; therefore, the std. dev. and # obs. are not reported.

Appendix VI. Alternative Specifications

Table IIIa+. Results for Racial Composition and Labor Cost (Distance Ratio)

Dependent Variable: ln(# African (A) / # Native (N))

Base Dataset (n=100)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Dist. Frontier/ Dist. Africa)	2.20***		2.24***	1.93***	0.34	0.50	0.65*
	(0.39)		(0.36)	(0.36)	(0.33)	(0.35)	(0.32)
ln(Dist. Frontier)		2.55***					
		(0.48)					
ln(Dist. Africa)		0.63					
		(1.40)					
Gold Deposits (1,000km^2)			1.68*	1.74**	1.51***	1.33***	1.17***
			(0.72)	(0.56)	(0.33)	(0.30)	(0.29)
% Suitable for Sugarcane			0.99	0.20	-0.23	-0.09	-2.12*
			(0.76)	(0.57)	(0.57)	(0.60)	(0.86)
Initial Native Population				-6.5E-6**	-9.6E-6***	-9.0E-6***	-9.2E-6***
				(2.2E-6)	(1.4E-6)	(1.5E-6)	(1.4E-6)
Time					0.01**	0.02***	0.02***
					(0.005)	(0.006)	(0.005)
End of Slave Trade Dummy						-1.52***	-1.50**
						(0.43)	(0.45)
Avg. Gold Deposits Neighboring States							0.65
							(0.94)
Avg. Sugarcane Suitability Neighboring States							2.7**
							(0.98)
R-Squared	0.33	0.35	0.38	0.41	0.58	0.61	0.64
No. Observations	100	100	100	100	100	100	100

+Results from ordinary least squares.

++White period standard errors in parentheses. ***P-value < 0.001; **P-value < 0.01; *P-value < 0.05.

Table IIIb+. Results for Racial Composition and Labor Cost (Distance Ratio)

Dependent Variable: ln(# African (A) / # Native (N))

All Data (n=156)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Dist. Frontier/ Dist. Africa)	**2.76*** **		**2.79*** **	**2.59*** **	**0.68** **	**0.97*** **	**1.09*** **
	(0.29)		**(0.28)**	**(0.29)**	**(0.26)**	**(0.25)**	**(0.28)**
ln(Dist. Frontier)		**2.97*** **					
		(0.35)					
ln(Dist. Africa)		**-0.29**					
		(1.68)					
Gold Deposits (1,000km^2)			**1.33***	**1.40** **	**1.22*** **	**1.03** **	**1.03*** **
			(0.55)	**(0.47)**	**(0.35)**	**(0.34)**	**(0.28)**
% Suitable for Sugarcane			**0.41**	**-0.15**	**-0.92**	**-0.84**	**-3.01***
			(0.89)	**(0.84)**	**(0.74)**	**(0.74)**	**(1.17)**
Initial Native Population				**-5.3E-6**	**-1.1E-5*** **	**-1.0E-5*** **	**-1.1E-5*** **
				(3.6E-6)	**(2.4E-6)**	**(2.4E-6)**	**(2.7E-6)**
Time					**0.01*** **	**0.02*** **	**0.02*** **
					(0.002)	**(0.002)**	**(0.002)**
End of Slave Trade Dummy						**-1.78*** **	**-1.82*** **
						(0.40)	**(0.41)**
Avg. Gold Deposits Neighboring States							**0.01**
							(0.99)
Avg. Sugarcane Suitability Neighboring States							**2.93***
							(1.17)
R-Squared	**0.32**	**0.33**	**0.34**	**0.35**	**0.47**	**0.49**	**0.51**
No. Observations	**156**	**156**	**156**	**156**	**156**	**156**	**156**

+Results from ordinary least squares.

++White period standard errors in parentheses. ***P-value < 0.001; **P-value < 0.01; *P-value < 0.05.

Table IIIc+. Results for Racial Composition and Labor Cost (Distance Ratio)

Dependent Variable: ln(# African (A) / # Native (N))

Excluding Mulattos; Outliers Omitted (n=140)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Dist. Frontier/ Dist. Africa)	**1.70*** **		**1.74*** **	**1.56*** **	**0.12**	**0.39**	**0.48**
	(0.31)		**(0.26)**	**(0.27)**	**(0.24)**	**(0.26)**	**(0.27)**
ln(Dist. Frontier)		**1.92*** **					
		(0.32)					
ln(Dist. Africa)		**1.03**					
		(1.18)					
Gold Deposits (1,000km^2)			**1.67*** **	**1.73*** **	**1.59*** **	**1.42*** **	**1.40*** **
			(0.49)	**(0.40)**	**(0.30)**	**(0.28)**	**(0.20)**
% Suitable for Sugarcane			**1.33***	**0.78**	**0.31**	**0.32**	**-1.67***
			(0.61)	**(0.53)**	**(0.44)**	**(0.45)**	**(0.66)**
Initial Native Population				**-4.7E-6***	**-8.8E-6*** **	**-8.3E-6*** **	**-8.7E-6*** **
				(1.9E-6)	**(1.6E-6)**	**(1.5E-6)**	**(1.5E-6)**
Time					**.01*** **	**0.01*** **	**0.01*** **
					(0.002)	**(0.002)**	**(0.002)**
End of Slave Trade Dummy						**-1.58*** **	**-1.60*** **
						(0.41)	**(0.42)**
Avg. Gold Deposits Neighboring States							**0.30**
							(0.70)
Avg. Sugarcane Suitability Neighboring States							**2.63*** **
							(0.67)
R-Squared	**0.21**	**0.24**	**0.30**	**0.32**	**0.46**	**0.49**	**0.52**
No. Observations	**140**	**140**	**140**	**140**	**140**	**140**	**140**

+Results from ordinary least squares.

++White period standard errors in parentheses. ***P-value < 0.001; **P-value < 0.01; *P-value < 0.05.

Table IIId[+]. Results for Racial Composition and Labor Cost (Distance Ratio)

Dependent Variable: ln(# African (A) / # Native (N))

All Data (n=156); Mulattos Omitted;

Alternate Distance Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Dist. Frontier/ Dist. Africa)	0.82***		0.90***	0.85***	0.40*	0.34	0.41*
	(0.12)		(0.14)	(0.13)	(0.18)	(0.18)	(0.19)
ln(Dist. Frontier)		0.81***					
		(0.13)					
ln(Dist. Africa)		-1.34					
		(2.43)					
Gold Deposits (1,000km^2)			1.29*	1.45**	1.37**	1.19**	1.18***
			(0.60)	(0.44)	(0.41)	(0.40)	(0.31)
% Suitable for Sugarcane			1.71	0.58	0.003	-0.12	-2.57*
			(0.96)	(0.75)	(0.58)	(0.59)	(1.00)
Initial Native Population				-1.1E-5***	-1.2E-5***	-1.2E-5***	-1.3E-5***
				(2.0E-6)	(1.9E-6)	(1.9E-6)	(1.9E-6)
Time					0.01*	0.01***	0.01***
					(0.003)	(0.003)	(0.003)
End of Slave Trade Dummy						-1.51***	-1.47***
						(0.37)	(0.40)
Avg. Gold Deposits Neighboring States							0.33
							(0.97)
Avg. Sugarcane Suitability Neighboring States							3.28**
							(0.99)
R-Squared	0.18	0.18	0.23	0.30	0.33	0.35	0.37
No. Observations	156	156	156	156	156	156	156

[+]Results from ordinary least squares.

[++]White period standard errors in parentheses. ***P-value < 0.001; **P-value < 0.01; *P-value < 0.05.