EMPIRICAL EVIDENCE OF THE CONNECTIOM BETWEEN SLAVERY AND LETHAL VIOLENCE IN CONTEMPORARY BRAZIL

Bruna Sineta Scarpioni[1*]
Geraldo Edmundo Silva Júnior[2]**

ABSTRACT
The aim of this study is to obtain a causal, significant and positive estimate of the effects of slavery in 1872 over the homicide rates of Brazilian cities in 2000. To compare the data of both periods, the territorial division of the Demographic Census of 2010 is applied to all previous Census. The treatment variable is the percentage of enslaved people relative to the total population of the city in 1872 considering its borders in 2010. Due to the specificities of a continuous treatment, generalized propensity score weighting is employed with help from a machine learning algorithm. The causal effect is given by a dose-response function. Controlling the socioeconomic factors most employed in the Economics of Crime literature, it is noted that slavery is related to higher homicide rates in all models, though its effect cannot be considered causal due to specificities of the weighting model employed. A self-authored index of public safety and an interaction term between inequality and income are introduced among the controls. The former is an alternative to the lack of indicators of government action in public safety and the latter contributes to represent more accurately the economic diversity of the cities. This work collaborates to fill the gap of empirical studies about historical causes of contemporary violence in Brazil.
Keywords: Slavery; Violence; Brazil; Generalized propensity score.

JEL Classification: K42, N36, Z13.

1 INTRODUCTION

In 2022, Brazil reached the lowest level of violent deaths in twelve years (FBSP, 2023). The decline has been observed since 2018, although it lost momentum between 2021 and 2022. The homicide rate per hundred thousand inhabitants decreased from 23.9 in 2021 to 23.4 in 2022. Among the victims, 91.4% are male; 76.9% are black, and 50.2% are between 12 and 29 years old. However, 2022 recorded the highest number of rapes in the historical series, with 68.3% occurring at the victim's residence. Indicators of violence against children and

[1*] Master in Applied Economics from the Federal University of São Carlos (UFSCAR), Sorocaba, SP, Brazil. Email: bruna.scarpioni@gmail.com. ORCID: 0009-0005-2933-8931.
[2**] Professor at the Federal University of São Carlos (UFSCAR), Sorocaba, SP, Brazil. Email: gedmundos@ufscar.br. ORCID: 0000-0002-7110-4689.

adolescents, domestic violence, and violence against women have increased significantly. The volume of ammunition sold has increased by 147% since 2017, and the number of individuals registered as Hunters, Shooters, and Collectors (known as “Caçadores, Atiradores, Colecionadores”, CAC, in Brazilian Portuguese) of firearms has increased sevenfold since 2018[3].

In comparison to the world, Brazil ranks 123rd out of 163 countries analyzed in the 2023 Global Peace Index (IEP, 2023). One of the aspects evaluated is the perception of security. In this regard, Brazil obtained rates of concern higher than the rates of (self-declared) occurrence of violent crimes. The same occurred in 2022 when researchers emphasized that this difference could reflect individual or collective efforts to deal with actual high risks. That is, in the absence of these efforts, occurrence rates would be much higher. Evidence supporting this hypothesis is the high spending on private security in the country (IEP, 2022; 2023). Despite the decline in violent deaths, the increase in other crimes helps to understand the sense of insecurity among Brazilians.

In recent years, the relaxation of firearm carrying conditions and ammunition tracking requirements, as well as the extension of legal justifications for deaths caused by public agents, suggest that atomized, short-term, and force-intensive actions have become alternatives to public safety policies. In the last three years, there have been episodes of sedition and unofficial actions among public security forces (Lima, Bueno, and Alcadipani, 2021). Paramilitary groups have sought to assert themselves as a reaction to the state's ineffectiveness in ensuring

[3] According to the methodology of the Brazilian Institute of Geography and Statistics (IBGE, in Brazilian Portuguese), the category "black" represents the union of self-declared black and brown individuals (IBGE, 2013). Since the differentiation between blacks and browns does not compromise the assumptions of this work, the term "black" has been adopted throughout the text to refer to the population directly descended from African populations enslaved in the Americas. “Brown” will be employed only in case of specific information regard this self-declared group.

peace, gaining strength and now competing for the monopoly of various illegal activities with older criminal factions (Cano, 2012; Manso, 2020).

However, this approach to the public security problem is not exactly unprecedented. Since Brazil began to see itself as a violent country, short-term and highly force-dependent actions have been preferred over integrated policies for preventing and combating crime. This occurred in the early 1980s (Amorim, 2013; Cerqueira, 2014) when the press and the population became aware of the first criminal factions formed at the end of the previous decade (Amorim, 1993). Even after the demilitarization process started in 1985, which marked the gradual return to democracy and direct elections in Brazil, state action continued to be based on the punctual and arbitrary repression of urban crime (Bueno, 2014; Spaniol, Moraes Jr., and Rodrigues, 2020).

In general, it is possible to observe that the approach to public security in Brazil disregards historical and structural causes of violence. Although historical factors cannot be widely considered when formulating specific policies, they have bequeathed the institutions and structures that determine incentives and disincentives to violence and crime today. By ignoring them, the probability that the measures adopted will be ineffective increases, contributing to discredit government actions while incentives for violent behavior remain intact.

Due its authoritarian nature, the Military Dictatorship (1964-1985) was marked by the distance between researchers and public policy makers, which especially compromised quantitative and empirical research on violence in Brazil. This distance helps explain the existence of two strands of literature on the subject produced during the New Republic (1985-2023). The first consists of Sociology and its intersection with other areas (such as Law) in investigating the trajectory and causes of violence in Brazil (Adorno, 1996; Bueno, 2014;

Manso and Dias, 2018; Manso, 2020). Created in 1987, the Center for the Study of Violence of the São Paulo University (in Portuguese, Núcleo de Estudos da Violência da Universidade de São Paulo (NEV-USP)) stands out as a pioneering research center specifically investigating crime in Brazil.

The second strand is essentially composed of Crime Economics, which faces obstacles in obtaining data because of problems in security monitoring processes employed (or not) by the state organs. Under this restriction, this strand mostly analyses contemporary and direct causes of crime (FBSP, 2007; Justus, 2012; Cerqueira, 2014; Justus, Khan, and Cerqueira, 2016). Founded in 2000 with the aim of bringing together different social sectors related to security, the Brazilian Forum of Public Security (Fórum Brasileiro de Segurança Pública (FBSP), in Portuguese) has a huge variety of researchers, but has been particularly decisive in producing quantitative data and monitoring public policies.

The division between the strands is marked by the absence of an empirical approach to the historical causes of violence in Brazil. Therefore, this work focuses on filling this gap and constitutes the first empirical investigation into a historical determinant of contemporary violence by examining the causal role of slavery in the lethal violence levels in municipalities in 2000.

Sociological literature asserts that slavery was the basis of the social structure that provided socioeconomic incentives for Brazil to become so unsafe, naturalizing violence in social relations to the point of making it invisible. Hence the idea that the country would be peaceful until the outbreak of drug trafficking, even with various historical records of violence in family, labor, and social spheres (Fernandes, 1965, 1978; Franco, 1983; Ribeiro, 1995; Souza, 2017).

The analysis of other countries subjected to colonial slavery supports the hypothesis that its damages are long-term. Acemoglu, Johnson, and Robinson (2002) obtained empirical evidence that, in exploitation colonies, extractive and predatory institutions characteristic of the slave period survived abolition and decolonization, perpetuating elite privileges and leading to their low development. Acemoglu, García-Jimeno, and Robinson (2012) found a direct relationship between slave labor in gold extraction and higher poverty rates, lower education levels, lower vaccination rates, and less access to public goods in Colombian municipalities. Educational inequality, which supports income inequality, is also related to slavery in former colonies (Bertocchi and Arcangelo, 2012, 2014).

In the case of Brazil, many socioeconomic aspects have been empirically linked to slavery. Using data from the 1872 Imperial Census on the enslaved population in Brazilian municipalities, Maloney and Valencia-Caicedo (2016) found a significant and decreasing relationship between slavery and municipal per capita income in 2000. Fujiwara, Valencia-Caicedo and Laudares (2017) concluded that slavery had a significant and positive impact on current income and education inequalities in the country.

Compared to other aspects, violence is relatively less analyzed in empirical studies of the effects of slavery. The United States is the main object of these investigations, mostly due to its greater availability of data on the slave period. Gouda and Rigterink (2017) found that the proportion of slaves in U.S. counties in 1860 is positively and significantly related to violent crime rates (homicide, rape, robbery, and assault) in all Censuses conducted between 1970 and 2000.

According to the Transatlantic Slave Database (compiled by Harvard University), the United States received about four hundred thousand slaves, while Brazil received 4.9 million. The American abolition took place in 1865, while Brazilian abolition took place in 1888 under strong international pressure. Given these data and the conclusions of Gouda and Rigterink (2017) for the United States, it is reasonable to expect that the relationship between slavery and contemporary violence in Brazil is more pronounced – even though data limitations may significantly hinder its identification.

In addition to greater dependence on the slave system, social, political, and spatial configurations remained similar to the colonial period even after the Abolition in 1888 (Fernandes, 1965, 1978; Ribeiro, 1995; Schwartz, 2012; Souza, 2017). During slavery, the territorial dominance of sugar cane plantation owners led to the atomization of power, which became an obstacle to the authority of the emerging Republic (1889). The primacy of local powers and the low legitimacy of official powers favored the arbitrary use of force, the appropriation of institutions by the elites and extra-judicial solutions at the expense of due legal process, practices that persisted in the republican period and were normalized throughout Brazil's history.

After Abolition (1888), freed workers faced economic marginalization because the main economic activity of the early 20th-century-Brazil (the coffee agriculture/industry) used free labor from European immigrants. Poverty prevented *de facto* freedom and the overcoming of the disadvantages bequeathed by slavery. The hurdles in economic integration and exercise of citizenship among the majority of the population created a post-colonial society that was segregated, traditionalist and less cohesive, whose individualistic and authoritarian character persisted beneath a modern veneer (Fernandes, 1965, 1978; Schwarz, 2012).

Throughout the 19th and 20th centuries, economic and social policies continued to disregard the legacy of slavery (which contributed to its reinforcement) while its intrinsic violence, normalized over almost four hundred years, remained "invisible." Only after the disturbances of the 1970s and 1980s, reported at the beginning of this section, did Brazilians come to consider that they were in a violent country, despite abundant records of violence (in private, public, and interpersonal spheres) found in the press and literary arts (Souza, 2017).

Therefore, the goal of this work is to empirically verify if slavery has a positive and significant causal effect on municipal homicide rates in the year 2000. The municipal level was considered the most suitable due to differences in Brazil's occupation before and after Abolition (1888). The only source of municipal data related to the slave period is the 1872 Imperial Census, retrieved and published by the Brazilian Institute of Geography and Statistics (in Portuguese, Instituto Brasileiro de Geografia e Estatística (IBGE)). The IBGE also adjusted all previous censuses to the municipal division of the 2010 Demographic Census, the most comprehensive available.

From the 1872 Imperial Census, the treatment variable is obtained: the percentage of enslaved people in the total population of a municipality in 1872, called "slavery density" for greater objectivity. In this year, all occupied municipalities had a share of enslaved population, so they form the treatment group of this study. The counties left compose the control group, since they did not have an enslaved population because they were not occupied at the time.

It is necessary to note that, in 1850, the Eusébio de Queiróz Law prohibited the transatlantic slave trade in Brazil, leading to a considerable decrease in the supply of enslaved workers even though the slave system remained. As the worker was considered a commodity with which masters sought to spend the least possible resources, working conditions led to immense mortality among the trafficked people. As the Imperial Census of 1872 was conducted after the transatlantic trade prohibition, it reflects the reduction of the enslaved population in a country that remained intensely dependent on enslaved labor. Therefore, it cannot account for the role of slavery in Brazilian economy and history. On the other hand, if the hypothesis of this study is valid for the 1872 data, it is possible to speculate that the real impact of slavery on violence is greater than estimated, as the treatment variable does not represent the total causal effect of the observed event due to the data collection difficulties mentioned above.

The proxy variable for violence in Brazil is the municipal homicide rate in 2000, chosen because it's the furthest year from 1872 for which municipal public security indicators are available. Binary variables representing Brazilian states were included to capture the effects of spatial autocorrelation, generating fixed-effects models.

The model specifications are based on Buonnano and Vargas (2017). Their instrumental variable model represented the effect of slavery on crime in Colombia as mediated by income inequality, one of the most common phenomena in empirical studies of the impact of slavery (Fujiwara et al., 2017). In turn, the Economics of Crime literature observes that income inequality has a strong correlation with crime rates (Soares, 2000; Soares and Naritomi, 2010).

As it is associated with the treatment variable and the resulting variable, income inequality is a confounding variable. A causal effects model must meet the unconfoundedness hypothesis, that is: there must be no omitted confounding variable. In this case, the estimated effect will be truly responsible for differences in the values of the outcome variable between control group and treatment group. Therefore, it is essential to include income inequality among the control variables.

This is an observational study in which the division of treatment and control groups cannot be considered random. The occupation of the national territory reflected the goals of the Portuguese metropolis, which focused on the Southeast and Northeast regions of Brazil because the distance from the coast to Africa and Europe is shorter. Thus, municipalities in these regions have a higher probability of belonging to the control group. In addition, first-settlement municipalities tend to have higher levels of economic activity, population density, etc. Therefore, for a given municipality in the sample, belonging or not to the control group is correlated with the observed values of several explanatory variables of the model. Hence, it is necessary to use a causal econometric method that is suitable for continuous treatment variables and considers the selection bias present in the sample.

The generalized propensity score (GPS) method, an extension of the original model designed for binary treatment variables (Rosembaum and Rubin, 1983; Hirano and Imbens, 2004), is employed. In a GPS model, the weighting of the sample units is such that the correlation between the treatment and the control variables becomes null. That is: belonging to treatment or control groups no longer correlates with the observed values of control variables for each municipality (so there is no selection bias in the weighted sample). At the same time, the weighting faithfully reflects the distribution of the treatment variable throughout the sample, which is crucial to obtaining robust estimates of the effect of a continuous treatment variable.

Therefore, to achieve the goal of obtaining a positive and significant causal estimate of the effect of slavery in 1872 over the municipal homicide rates in 2000, this work is structured as follows: the current introduction as a first section; a second section provides a literature review on the topics addressed; the methodology and empirical strategy are described in a third section; subsequently, the results obtained are presented and discussed in a fourth section. Finally, concluding remarks are presented in the fifth and last section.

2 LITERATURE REVIEW

Why is it reasonable to assume that colonial slavery affected (and probably still affects) contemporary Brazil? This section seeks to answer this question. One of the reasons is the existence of sociological and empirical studies linking slavery to contemporary phenomena in nations related to this institution. Obstacles for economic development, racial disadvantages in human capital transmission and political-institutional racial discrimination are the most observed long-term effects. Most empirical literature refers to the United States, former Spanish colonies and African countries. However, empirical analyses of the effects of slavery in Brazil have been growing in number and complexity.

Path dependance studies (Engermann and Sokoloff, 1997, 2002; North, 1990; Acemoglu, Johnson, and Robinson, 2002) observed that, during 18th and 19th century colonialism, areas geographically and climatically suitable for cultivating products highly valued in international markets received enslaved labor, such as Brazil and the Southern United States. There, metropolises established extractive institutions to gain short-term benefits, which led to high levels of income inequality in the colonies. These institutions did not change after the abolition of slavery or the end of colonialism, so they contributed to maintain elite privileges, exacerbate inequality and perpetuate disadvantages in economic development.

Examining this hypothesis empirically, Nunn (2008) found evidence that slavery harmed the economic development of the United States, of ex-Iberian colonies and of African countries affected by transatlantic trafficking. However, it was not possible to detect whether these harms were caused by income or land inequality.

Studying African countries, Nunn and Wantchekon (2011) found that ethnic groups whose ancestors were more affected by trafficking in the past had lower levels of trust in institutional powers and in other social groups. This suggests that slavery compromised the integrity of social fabric in countries victimized by trafficking, with significant consequences for the social and political instability that followed the end of colonialism.

Empirical studies on the impact of forced labor in former Spanish colonies suggest that slavery compromised the development of these regions. Acemoglu, García-Jimeno, and Robinson (2012) found a significant and positive relationship between indigenous forced labor, employed in gold extraction during Spanish colonization, and higher poverty rates, lower education rates, lower vaccination rates, and less access to public goods in Colombian municipalities.

Observing many Latin American countries, Maloney and Valencia-Caicedo (2016) noticed that contemporary population density is correlated with economic activity during the colonial period. Additionally, using municipal data from the 1872 Imperial Census, the authors found a significant negative relationship between slavery and per capita income among Brazilian municipalities in the year 2000.

The slavery in United States is subject of most empirical studies, mainly due to the greater availability of data from the period in that country. Bertocchi and Dimico (2012, 2014) concluded that school inequality between whites and blacks in the United States can be attributed to slavery. Educational inequality sustains income inequality over time, which leads to a disadvantage in transmission and absorption of human capital among enslaved population descendants. According to the authors, political exclusion after freedom prevented the correction of these distortions.

Regarding this human capital lag, Sacerdote (2005) compared the freed black population with the free black population and their respective descendants before and after the American Abolition (1865). Observing literacy rates, school attendance and employment levels, the author concluded that it took two generations of descendants of the freed black population to achieve the same living conditions as the descendants of the free black population. Thus, it was possible to conclude that the damage caused by slavery persisted at least until 1920.

Using municipal data, Jung (2019) found empirical evidence that the incentives to absorb human capital were reduced among the freed population due to discriminatory labor legislation, which prevented access to the free labor market. By estimating the return on education for US municipalities in 1940, the author found that, in regions highly dependent on slavery in the past, black individuals had more difficulty converting their education into income and social mobility.

The political attitudes of the American population were also influenced by slavery, which contributed to the persistence of its effects. Through surveys answered by around forty thousand people, Acharya, Blackwell and Sen (2016) concluded that white inhabitants of cities with a higher proportion of enslaved population in 1860 tended to express negative views about the black population and not support affirmative actions.

According to the authors, the legacy of slavery is transmitted through formal and informal social institutions. The Jim Crow segregation laws and the criminalization of unemployment after the American Abolition (1865) are examples of formal institutions. The latter allowed for low wages for freed individuals, which led to conditions akin to enslaved labor. Examples of informal institutions are (overt or covert) social punishment for interracial relationships and distrust of authorities regarding the black population (commonly associated with poverty and considered prone to criminality). The fact that penalties for the same crime are higher for black perpetrators than for white perpetrators suggests that this distrust has concrete implications. Formal and informal mechanisms run parallel in history, simultaneously interacting and reinforcing each other (Acharya, Blackwell, and Sen, 2016).

In the case of Brazil, the impacts of slavery have been analyzed by leading figures in sociology since the early 20th century. Freyre (1973), Holanda (1984), Fernandes (1965, 1978), and Ribeiro (1995) asserted that slavery is essential to understanding contemporary Brazil. The first two were exponents of the so-called culturalist paradigm, for which culture is the main determinant of national identity. Though he acknowledged the weight of slavery, Holanda emphasized Iberian heritage, while Freyre emphasized the hybrid character of the Brazilian population, the "miscegenation," and the "harmony of opposites" as the main national traits. According to Freyre, the absence of open racial conflicts and the primacy of personal relationships over public space rules led to a peaceful society after Abolition (1888).

In the mid-20th century, scholars with materialist approaches moved away from culturalism, seeking to understand the consequences of Brazilian slavery more precisely. In *The Integration of the Negro into the Class Society*, Fernandes (1965) observed that colonial slavery was decisive for the capital accumulation process that financed the industrialization of Europe and its former colonies, albeit at different rates. According to him, Brazil experienced its "bourgeois revolutions" by abolishing slavery in 1888 and establishing the Republic in 1889: these transformations could not disrupt (and did not disturb) the social and economic dynamic of Brazilian slavery system.

For Fernandes, the maintenance of labor conditions previously established by slavery and transatlantic trafficking enabled the investment in large modern agriculture, as well as the beginning of industrialization and urbanization in Brazil. The country did not truly break with slave structures but perpetuated them under nascent capitalism, subverting possible transformations and preventing *de facto* freedoms (Queiróz, 2021).

Marked by cognitive dissonance[4], this historical process is associated with the myth of racial democracy, according to which the legal equality acquired with Abolition (1888) was considered factual equality. The myth naturalized and invisibilized the social exclusion of freed individuals and their descendants, decisively contributing to its perpetuation into the 20th century. In the *The Black in the World of Whites* (written during the Military Dictatorship), Fernandes noted that:

> In summary, urban expansion, the industrial revolution, and modernization have not yet yielded sufficiently profound effects to alter the extreme racial inequality inherited from the past. [...] This statement contradicts what is commonly said about the racial democracy presumed to prevail in Brazil. The issue lies in the conflation of strictly imperative standards in the realm of social decorum with true equality (Fernandes, 1978, p.67).

Combining a rigorous data survey with historical materialism, Florestan Fernandes was categorical in stating that the social ills of 20th-century Brazil trace back to colonial slavery. The relationships between the national economy and the modernization of world economies, between dependence on slavery and nascent capitalism, have remarkable parallels with those described by path dependance studies at the beginning of this section.

[4] Cognitive Dissonance Theory is one of the most important theories in Social Psychology. According to its creator, Leon Festinger (1957), humans align actions and behaviors with their convictions, and vice versa. In case of inconsistency between conviction and action (i.e., cognitive dissonance), discomfort arises. To remedy it, conscious or unconscious attempts are made to reduce this distance by adjusting the involved behavior or convictions. In the case of the emerging Republic, the reality of Abolition and dependence on slavery structures did not correspond to the emerging modernization. To balance this gap without altering the foundations of the economy, a formal (rather than factual) reconciliation effort occurs. Cognitive dissonance can be seen through intellectual works downplaying the pernicious effect of slavery, but mainly in institutional efforts to create a modern national identity that overlooks the differences and disadvantages of the liberated population (Fernandes, 1965, 1978).

In his work *The Brazilian People*, anthropologist Darcy Ribeiro corroborates Florestan Fernandes' conclusions. After highlighting the abyss that separated the living conditions of whites and blacks during slavery, the author states that ‘in these conditions, one should seek the explanation for the glaring discrepancy between the expansion of the white and the black contingents in the development of the Brazilian population’ (Ribeiro, 1995, p. 233). Finally, among the more recent works, Souza (2017) reaffirmed that slavery is the foundation upon which Brazilian society is built:

> In Brazil, from year zero, the institution that encompassed all others was slavery, which did not exist in Portugal except in a very specific and fleeting manner. Our forms of family, of economy, of politics and of justice were entirely based on slavery (Souza, 2017, p.40).

The number of empirical studies on the contemporary impact of Brazilian slavery has been increasing in the last two decades. Summerhill (2010) used available data from fifty thousand farms in the state of São Paulo to test the validity of Engerman and Sokoloff's hypothesis. Ordinary least squares and instrumental variable models could not find a significant relation between slavery and long-term economic development in São Paulo.

In recent years, works on the topic have become increasingly sophisticated. Fujiwara, Laudares and Valencia-Caicedo (2017) arrived at the opposite conclusion when analyzing all Brazilian municipalities in 2000 and 2010. The Treaty of Tordesillas was used as reference to develop a regression discontinuity design model, since the Portuguese portion of Brazil received much more enslaved people than the Spanish portion. The authors found that areas with a higher share of enslaved population in 1872 are wealthier today but have higher income and education inequality, as well as lower-quality public institutions. Studying the emergence of the textile and cotton industries in Brazil, Palma, Papadia, Pereira, and Weller (2021) suggest slavery is related to a delay in the country’s industrialization process.

Social capital is defined as the network of relationships between people living and working together to form a functional society. Uttermark (2019) used LOGIT models to test two hypotheses about the relationship between slavery and interpersonal trust in Brazil and in the United States. The first hypothesis is that the primacy of the slavery-dependent monoculture would have generated economic inequalities that, in turn, decreased long-term social capital. The second claims that slavery would have influenced people's political positions, which, transmitted over time, harmed social capital in these countries. The author concluded that slavery is negatively correlated with social capital: interpersonal trust tends to be lower in regions where slaverym was concentrated. Tests revealed that the second hypothesis was the most suitable to explain this result.

Using instrumental variables, Seyler (2021) observed that slavery and support for coercive institutions in the 19th century are associated with lower levels of economic development in Brazilian municipalities. Support for authoritarian institutions in the late 19th century was measured by the votes of legislators on Abolition (1888) issues, and current municipal economic development, by Gross Domestic Product (GDP), poverty, and income inequality. The author found that, in municipalities with a greater presence of slavery and strong support for authoritarian institutions in the past, residents exhibited lower levels of interpersonal trust, a higher likelihood of not supporting democracy and a belief that corruption is acceptable in some circumstances. This set of behaviors is part of the municipality's social capital, and for the author, this was the channel through which the harms of slavery contaminated economic development over time.

Papadia (2019) emphasizes the difficulties of determining the effects of slavery on contemporary Brazil, both due to the scarcity of data sources and the irregular distribution of the enslaved population across the national territory. In addition to municipal data from the 1872 Imperial Census, the author uses quilombos, refuges for workers escaping slavery. It is assumed that places with a higher number of quilombos had a greater penetration of slavery. Using matching methods to analyze the states of Minas Gerais, Rio de Janeiro and São Paulo, the author found that slavery had a negative impact on economic development during its period of existence and in the next thirty years after its abolition (1888). Municipalities where slavery was more intense presented worse literacy and public management indicators.

Lambais (2020) identified that Brazilian quilombos have a positive and robust relationship with local economic development in recent periods. According to him, quilombos allowed the intergenerational transmission of human and cultural capital – for example, religious beliefs, metallurgy techniques and other skills brought by enslaved African immigrants. These assets would favor the long-term economic development of these regions. To obtain more precise fixe effects, the author divided Brazilian territory into cells of approximately eleven square kilometers. Using nighttime lighting as a proxy variable for economic activity, it was observed that areas with a higher number of quilombos had a more dynamic economy. Obtained through random inference, spatial configurations of 'counterfactual quilombos' showed that proximity to quilombos is related to a higher number of occupations in metallurgy or other human capital-intensive areas, and to higher levels of cultural activity, interpersonal trust, and organized collective action.

Thus, it’s possible to notice that empirical studies of other post-slavery nations show that the effects of slavery persist, especially on economic and social development. When it comes to Brazil, the quantitative and causal analysis of these impacts has been growing in volume of publications and methodological complexity. However, violence has not yet been analyzed from this perspective, and this work aims to fill this gap. The hypothesis that slavery had a significant and positive impact on violence levels in contemporary Brazil is supported by Brazilian sociology, as this section has also shown.

Another reason to consider that the hypothesis of this work is relevant lies in the pernicious effects of racism, a social ill tightly associated with colonial slavery. Empirical studies show that racial discrimination affects violence levels among black population in direct and indirect ways. The latter means that racism contributes to reinforcing inequalities in access to resources such as education and employment, which could significantly reduce victimization levels among discriminated groups. Data about the inequalities between blacks and whites in recent Brazil are provided next. Even though racial inequalities do not imply racial discrimination by themselves, recent empirical studies suggest that many of them are indeed associated with racism.

In 2021, the black population[5] represented 56.1% of total population and 55.2% of the workforce (compared to 43.8% of whites). However, they were also the majority among unemployed (64%) and informal workers. Only 29.5% of managerial positions were occupied by blacks (69% were occupied by whites). The per capita household income of the white population was almost twice that of the black population. Regarding extreme poverty, 72.8% of people living on less than $5.50 per day were black, and 18.6% were white (IBGE, 2022).

In 2021, at all education levels, the real average labor income was higher for whites. Among people with college degrees, whites earned on average 50% more than blacks and 40% more than browns. During the coronavirus pandemic, black students were the majority among

[5] Brown people compose 47,0%, and black people, 9,1% of the total Brazilian population (IBGE, 2022).

those who did not receive school activities. They also had less than 5 days per week dedicated to school and less than two hours per day dedicated to studies. The attendance rates of black students on the National High School Examination (in Portuguese, Exame Nacional do Ensino Médio (ENEM)) were about 10% lower than those of whites.

Inequalities in access to education and employment are associated with poverty, which in turn is associated with higher levels of violence. In 2020, the homicide rate among white individuals was 11.5 victims per hundred thousand inhabitants; 21.9 among brown individuals and 34.1 among black individuals. The male homicide rate was 12.5 times higher than the female this year. For black, brown and white men, the homicide rates were, respectively, 64.3, 41.4, and 21.2 deaths per hundred thousand inhabitants. Among men aged 15 to 29, who form most victims, blacks have 94.4 deaths per hundred thousand inhabitants; browns, 136.5; and whites, 41.6 (IBGE, 2022).

According to the 17th Brazilian Public Security Yearbook (FBSP, 2023), 72% of homicide victims in 2022 were black or brown. Compared to the previous year, homicides decreased by 26.5% among whites and increased by 7.5% for blacks. These inequalities also occur regarding violence against women. In 2021, black women accounted for 62% of femicide victims, 70.7% of intentional violent deaths victims, and 52.2% of rape and vulnerable rape victims (FBSP, 2023).

The strong racial inequalities in poverty and victimization indicators suggest that, in addition to harming economic development, slavery is related to the obstacles the black population faces to this day. As we will see next, present racial discrimination contributes to reinforcing them. This corroborates the premise that slavery affects long-term social relations, which significantly affects the development of a post-slavery society like Brazil.

With a sample of about 277,000 students from public schools in the state of São Paulo, Botelho, Madeira and Rangel (2015) observed that teachers gave lower grades to black students than to white students (who made the same mistakes). The authors emphasized that this bias could lead to higher dropout rates among black students. Recently, Leung-Gagné (2023) pointed out that racial discrimination in Brazilian schools is similar to the one in North Carolina, United States. Analyzing data from *Prova Brasil* (a national standardized testing system) about 5th and 9th-grade students, the author concludes that the country suffers from 'random segregation,' i.e., the distribution of students occurs without direct interventions to segregate them. However, by not acting to correct the racial disadvantages ingrained in Brazilian society, schools end up reproducing them, and the result tends to be very similar to what it would be if segregation measures had been implemented.

In Economics, racial discrimination in the labor market is measured by wage differences that cannot be explained by differences in human capital or contribution to economic outcomes (Fernandes, 2015). Arcand and D'Hombres (2004) were among the first to detect it more accurately. Analyzing the National Household Sample Survey (in Portuguese, Pesquisa Nacional de Amostra Domiciliar (PNAD)), the authors found evidence that wage differences between whites and blacks were related to discrimination. About the female labor market, Fernandes (2015) observed that the racial wage difference among the most efficient workers was high and emphasizes that it could be corrected by policies that disregard color or race. Both works use the Oaxaca-Blinder decomposition to observe the components of the labor income of white and black individuals.

Between 2003 and 2013, Fiuza-Moura, Maia, and Souza (2018) found evidence of racial discrimination in the manufacturing sector, although the wage difference between blacks and whites decreased over the period (especially in more technology-intensive industries).

Analyzing the 1990s process of economic liberalization after the return of democracy in Brazil, Barros and Silva (2021) noted that the salaries of non-white women increased compared to those of white men and women, and racial discrimination decreased among higher salary levels. However, during the economic opening process, racial discrimination at lower salary levels also increased.

Regarding the behavior of Brazilian firms, Gerard, Lagos, Severnini and Card (2018) observed that non-white individuals have lower probabilities of working in establishments that pay more to all racial groups and of receiving financial bonuses. The first conclusion accounts for about 20% of wage differences between groups, and the second, for 5%. About two-thirds of the absence of non-white individuals in companies that pay high salaries is indeed due to differences in human capital and does not involve racial discrimination. However, among workers with college degrees, there is evidence of racial discrimination in hiring and salaries, and therefore the allocative costs of racism can be relatively high in Brazil.

Miller and Schmutte (2021) developed a model to determine if there is racial discrimination in hiring by referral, a common practice in Brazilian labor market. Using matching theory to analyze labor demand, the authors found that employers tend to prioritize hiring people with the same color or race as theirs, but as firms grow in size, they abandon this behavior. At the same time, entrepreneurship is about twice as high among whites compared to non-whites. These two conclusions suggest that hiring by referral is an aggregate disadvantage for non-white workers. According to them, the two facts help explain why non-whites are more likely to be fired, have lower chances of reaching higher positions and tend to seek employment in larger companies (p.28).

Finally, recent empirical studies have observed that racial discrimination does affect violence levels in Brazilian society. Truzzi, Lírio, Cerqueira, Coelho and Card (2021) calculated the racial victimization differential between whites and blacks in Brazil regarding homicide and non-lethal aggression. Based on PNAD and the Mortality Information System of the Health Ministry (in Portuguese, Sistema de Informações de Mortalidade do Sistema Único de Saúde, (SIM-DATASUS)), the authors use the Oaxaca-Blinder decomposition to conclude that part of the victimization differential can be attributed to structural factors. For both crimes, 40% of this differential is related to racial discrimination. The highest levels of discrimination were found in North and Northeast regions, which also have higher social inequality and violence levels. This suggests that, in addition to directly affecting violence, discrimination is correlated with woes that can be traced back to the slavery heritage.

This section sought to demonstrate that empirical studies from other post-slavery nations have already detected the effects of slavery on various contemporary phenomena. Furthermore, the effects of racism in Brazil, one of the main legacies of slavery, have been analyzed empirically lately. All of this contributes to consider that it is possible to detect significant and positive causal effects of slavery on lethal violence in contemporary Brazil.

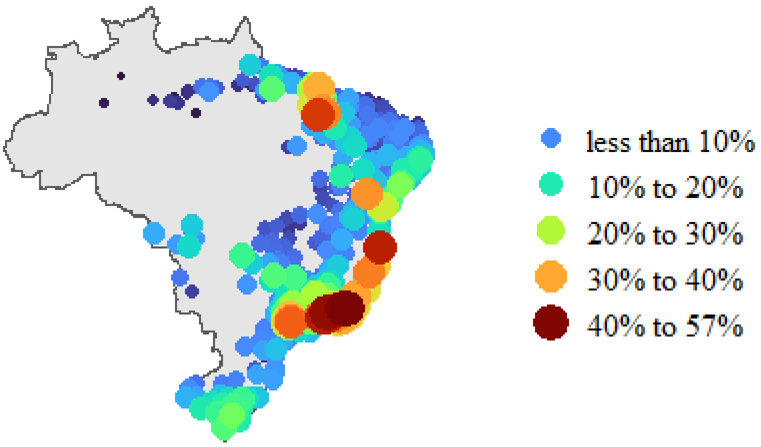
3 EMPIRICAL STRATEGY

3.1 Slavery as a Natural Experiment: Methodological Implications

Brazilian colonial slavery can be considered a natural experiment, that is, an exogenous event that alters the environment in which the studied agents operate (Cook, Campbell, and Shadish, 2002). Settlement was concentrated on the coast (Palma, Papadia, Pereira and Weller, 2021) due to the profitability of transatlantic trade, which made the occupation of the interior territory comparatively more costly for the Portuguese metropolis (Fujiwara, Laudares and Valencia-Caicedo, 2017). Therefore, slavery was located predominantly in the current Southeast and Northeast regions of Brazil.

The Imperial Census of 1872 is the only one that gathers data on the enslaved population. Using Comparable Minimum Areas (or Áreas Mínimas Comparáveis (AMCs), in Portuguese), the Brazilian Institute of Geography and Statistics (Instituto de Geografia e Estatística (IBGE), in Portuguese) applied the municipal division of the 2010 Demographic Census to all previous Censuses (IBGE, 2011). Thus, it was possible to create the map shown in Figure 1, which displays the distribution of the treatment variable across the national territory. Illustrated below, the density of slavery is given by the percentage of enslaved individuals in the total population of a municipality in 1872 under its territorial boundaries in 2010.

Figure 1 – Distribution of slavery density (%) in 1872



Note: created by the author using R Studio.

From the 5,499 municipalities listed in 2010, only 640 were occupied in 1872. All of them had a share of enslaved inhabitants among the total population (IBGE, 2011). Thus, the colored areas in Figure 1 represent the treatment group municipalities, where the density of slavery is greater than zero. Gray areas show the control group municipalities, whose slavery density is null because they were unoccupied in 1872.

In Figure 1, bubbles ranging from dark blue to green represent municipalities with up to 30% of the population enslaved in 1872. More inland, these bubbles are concentrated in the current Northeast, Southeast, and far South, with isolated points in the North and Midwest regions. Orange and red bubbles, indicating municipalities where the enslaved population represents more than 30% of the total inhabitants, are concentrated on the coast of the Southeast region and on the border between the Northeast and North.

Figure 1 suggests that the selection of treatment and control groups cannot be considered random: municipalities in the Southeast and Northeast regions are more likely to belong to the treatment group, while municipalities in the Midwest and North regions are more likely to belong to control group. Moreover, it is reasonable to assume that belonging to one or another group has a significant impact on the observed values of the control variables. For example: the treatment group tends to have higher levels of income, employment and population density because its early occupation allowed for capital accumulation before the control group even came into existence. This difference can significantly affect the observed values in 2000.

There is a strong correlation between the 1872 treatment variable and the 2000 control variables, and so the observed values of the controls are significantly different between groups. Under these conditions, differences in their homicide rates cannot be directly attributed to treatment (in this case, to slavery). Using only sufficiently similar observations would considerably reduce the sample and weaken the representativeness of the results. To overcome these limitations, the Propensity Score Weighting (PSW) method is employed.

The probability that an observation belongs to the treatment group represents an estimate of the correlation between the treatment variable and the control variables. By weighting the observations according to this probability, the correlation between treatment and controls is considered by the model and no longer contributes to estimation bias. After weighting, the weighted sample is said to be balanced: in it, the treatment variable becomes independent of the control variables, and differences in the observed values of the outcome variable can be considered a causal effect of slavery. Thus, the PSW method corrects the limitations of the sample without compromising its size.

PSW methods were developed to address binary treatment variables (Rosenbaum and Rubin, 1983). However, slavery density is continuous: it assumes zero for the control group and varies in the range of 6.87% to 57.41% for the treatment group. Therefore, it is necessary to use the generalization of the PSW method for continuous treatments: the dose-response function (Hirano and Imbens, 2004). To understand it, it is useful to draw an analogy between this study and a common scenario in Health Sciences.

Imagine an experiment where one group of people did not ingest an experimental drug, while the other group ingested different doses of this drug. Considering that its effect may be stronger for some and weaker for others, the goal of the experiment is to obtain the average treatment effect (ATE) of the drug on the observed population, given by the dose-response function.

In this example, the first group represents municipalities that did not exist in 1872 (whose slavery density is null), and the second group represents municipalities that existed in 1872 (whose slavery density varies in a range). The "response" that this study seeks is the ATE[6] of slavery density in 1872 on municipal homicide rates in 2000.

To obtain adequate ATE estimators, it is necessary to correct the limitations of the sample: non-randomness, confounding bias, correlation between treatment variable and controls. By weighting the observed units by the inverse of their generalized propensity scores (GPS), a balanced pseudo-sample is obtained. These scores estimate the probability of belonging to the treatment group while considering that slavery density assumes various values in a range (i.e., is a continuous treatment variable).

Weighting by the inverse of GPSs is known as inverse probability weighting (IPW), so, for simplification, the weighting method used will be represented by the acronym GPS-IPW. Next, the machine learning algorithm employed to is used to accurately estimate GPSs (Zhu, Coffman, and Ghosh, 2015) is described.

[6] A maybe obvious remark: ATE is a causal estimative.

3.2 Obtaining the Dose-Response Function: GPS-IPW with Boosting

The municipal homicide rate in 2000 and the municipal slavery density in 1872 are represented, respectively, by Y and T. $\mathbf{X}$ is a vector of dimension p containing the control variables defined in the next section and the binary variables which control for fixed effects. These represent twenty-five Brazilian states[7]. From (Y, T, $\mathbf{X}$), a random sample $(Y_i, T_i, \mathbf{X_i})$, i = 1, ..., n is extracted, representing the observed data for n municipalities.

The variable T is random and is defined in the interval τ, to which the value t belongs. For each municipality i, there is a set of potential values of Y obtained as a function of the treatment value t applied to the municipality, that is: $Y_i(t)$ para $t \in \tau$. The goal is to obtain the dose-response function $\mu(t)= E[Y_i(t)]$, which states that the average treatment effect (ATE) is given by the expected value of potential outcomes for the treatment levels applied to the n observations. It is assumed that $Y_i(t)$ is defined for every $t \in \tau$, as potential outcomes need to be defined for all levels of treatment observed in the sample[8].

In general, the unconfoundedness condition determines that unbiased causal estimators can only be obtained if the estimation does not depend on variables absent from the model. Thus, it is necessary to control all confounding variables: those that affect both the treatment variable and the outcome variable simultaneously. However, when the treatment variable is continuous, it is not necessary for the potential outcomes of Y to be unconditionally independent of the values of T; it suffices that they are conditionally independent (Hirano and Imbens, 2004). Therefore, Equation 1 is termed the weak unconfoundedness condition:

---

[7] Brazil has twenty-six federative units or states. To avoid multicollinearity, one of them is not represented by a binary variable among the controls.
[8] For simplicity, as from now on variables will be expressed for the full sample, subscript 'i' is no longer employed.

$$f(t|Y(t), \mathbf{X}) = f(t|\mathbf{X}), \qquad \forall\, t \in \tau. \tag{1}$$

In Equation 1, $f(t|.)$ represents the conditional density of a treatment level $t \in \tau$. The weak unconfoundedness condition states that, given the control variables included in the model, the value of slavery density in a municipality is conditionally independent of the potential outcomes of its homicide rate. If this holds, there are no confounding variables. Therefore, belonging or not to the treatment group does not affect potential outcomes (Zhu et al., 2015). In this case, it is considered that the control variables represented in the model are sufficient to ensure the conditional independence of t relative to $Y(t)$.

When analyzing a natural experiment, it is not possible to ensure the randomness of the sample and the conditional independence between T and $\mathbf{X}$. However, it is mathematically possible to find a function $r(\mathbf{X})$ that is conditionally independent of T and use it to weigh the observations (Huling, Greifer, and Chen, 2023). By doing so, estimations based on the weighted sample will generate unbiased ATE estimators, as there will be conditional independence between treatment and control variables.

The generalized propensity scores (GPSs) $r(t,\mathbf{X})$ are the conditional densities of each treatment level t for the given values of the control variables (Zhu, Coffman, and Ghosh, 2015). Based on the characteristics expressed by the controls, these functions determine the probability of each treatment level's occurrence in the sample.

$$r(t, \mathbf{X}) \equiv f_{t|\mathbf{X}}(t|\mathbf{X}) \tag{2}$$

It is possible to express the weak unconfoundedness condition in terms of propensity scores:

$$f(t|Y(t), r(t, \mathbf{X})) = f(t|r(t, \mathbf{X})), \qquad \forall\, t \in \tau. \tag{3}$$

Equation 3 states that, to eliminate confounding bias, it is sufficient for the treatment levels to be conditionally independent of the propensity scores (Zhu et al., 2015). This means that, within each set of municipalities determined by the same score r(t,**X**), the probability of T being equal to t does not depend on **X**. That means that, for example, if a group of municipalities has the same probability of belonging to the treatment group, the value of T does not depend on the observed values of controls for any of the municipalities within the group.

It is not possible to guarantee that T and **X** are independent across the entire sample, but the weak unconfoundedness condition of Equation 3 asserts that it is possible to ensure independence among municipalities with the same probability of belonging to the treatment group, i.e., with the same score of propensity. Therefore, after weighting the sample according to the conditional density functions r(t,**X**), T and **X** become conditionally independent, allowing for the estimation of the causal treatment effect. Based on the observed data $(Y_i, T_i, \mathbf{X}_i)$, $i = 1, ..., n$, the weight assigned to municipality i is given by the inverse of the GPSs:

$$w_i = \frac{f_{T|\mathbf{X}}(T_i)}{f_{T|\mathbf{X}}(T_i|\mathbf{X}_i)} = \frac{r(T_i)}{r(T_i, \mathbf{X}_i)}, for\ i = 1, \dots, n. \tag{4}$$

Equation 4 shows that obtaining the weights depends only on the relationship between T and **X**, so it is essential that the conditional density of T in relation to **X** is realistic. Considering this and the historical context of the data studied, it does not seem reasonable to assume a linear relationship between the density of slavery in 1872 and the observed controls in 2000. Hence, it is necessary to estimate the weights through a non-parametric method, which allows for non-linearity between T and **X**.

In the case of non-parametric models, as the weak unconfoundedness condition cannot be tested, there is a risk of the "curse of dimensionality" (Zhu et al., 2015). That is, there is an incentive to include controls as to avoid hidden confounding variables. The vector **X** can reach a high dimension *p*, which in turn can significantly complicate the estimation without making it more accurate. The generalized boosted model (GBM) selects key controls without compromising the accuracy of the GPSs (McCaffrey, Ridgeway, and Moral, 2004).

Boosting algorithms are machine learning techniques capable of adequately representing a high-dimensional vector **X** (i.e., with a large number of control variables) when its relationship with T is non-linear. Boosting algorithms assume that the relationship represented by Equation 5 is valid for the sample (Zhu et al., 2015):

$$T = m(\mathbf{X}) + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2). \tag{5}$$

In Equation 5, m represents the mean of T as a function of **X** and is unknown. Unlike linear models, **X** is not multiplied by a vector of coefficients; this relationship is open to different distributions of T given **X**. GBM algorithms use functions applied to small groups of observations to estimate a smooth function of controls on **X**. The first partition divides the sample into two groups according to some criterion (for example, the value of variable X_{1i}). In each resulting group, the simple average of T is estimated (i.e., the value of m(**X**) for the group) using regression trees, another non-linear algorithm. Partitions continue, and from all possible ways to divide the sample, those leading to the smallest prediction error of T are chosen (McCaffrey et al., 2004).

Finally, the GBM algorithm makes linear combinations of the regression trees that have the with the smallest prediction errors of T. The goal is to generate smoother functions that can be used to estimate the propensity scores. To determine the final function, the algorithm identifies which of the linear combinations minimizes the differences between the treatment and control groups. It is common to define the optimal model as the one minimizing the average absolute Pearson correlation between the continuous treatment variable and each of the controls, as is done in this study (Zhu et al., 2015). Thus, balance between the two groups can achieved, eliminating the selection bias.

In addition to identifying and ranking the most important variables for predicting the group to which a municipality belongs, GBM can automatically include interactions between variables and higher-order terms (Zhou, 2015). In this study, some controls may exhibit interactions whose relationship with the outcome variable can be significant. An example is the interaction between per capita household income and income inequality. Although this interaction is specified in the models, each control may still interact with the unemployment rate or with the percentage of people who have eight or more years of education. It's impossible to specify all the multiple possibilities of interaction. However, if not addressed, they could act as confounding variables, compromising the quality of the treatment effect estimators.

The GBM algorithm also helps mitigate the trade-off between variance and bias characteristic of GPS-IPW models (Zhu et al., 2015). To eliminate selection bias, these models can generate a very large range of weights. When the variance of the weight range is high, estimation is said to be unstable. The risk of instability is higher when there is a large number of controls, and the groups are very different from each other. In this case, there are few observations that can easily be compared, and so they receive very high weights. While most observations receive lower weights, some observations would significantly influence the estimated treatment effect, and so they could distort it.

When there are no missing values among the observed data, it has been shown that the weights obtained by GBM lead to more accurate treatment estimates and lower mean squared error compared to linear methods (Zhu et al., 2015). For the 5,499 municipalities determined by the 2010 Territorial Division (IBGE, 2010), all independent variables are complete. However, only 2,699 of these municipalities reported their homicide rates in 2000.

Equations 4 and 5 show that weight estimation depends only on T and $\mathbf{X}$. Essentially, correct weights are those that reflect the selection of treatment and control groups established by the natural experiment. Since T and $\mathbf{X}$ data are available for the 5,499 municipalities, the weighted sample has 5,499 units. Reducing it based on missing values of the outcome variable would mean compromising the portrayal of the empirical relationship between slavery in 1872 and Brazilian society in 2000.

To deal with missing data in the outcome, it is possible to limit the sample to complete cases (units that have observed values for all variables) or impute missing values. According to van Buuren (2018), if only the resulting variable has missing values, the two approaches are equivalent and generate unbiased estimators.

The GPS-IPW models were chosen, among other reasons, because they avoid the loss of sample units, so it would not make sense to use only complete cases. On the other hand, imputation would lead to the same estimates as the analysis of complete cases but would increase their variance, generating inconsistent and non-robust estimators (Van Hippel, 2007). Imputation would only be a viable option if other independent variables beyond those contained in $\mathbf{X}$ were available. However, the controls traditionally included in the literature are already specified (Gouda and Rigterink, 2017; Buonanno and Vargas, 2017), and the GBM algorithm is responsible for detecting possible relevant interactions between variables.

The choice was to weigh the observations based on the treatment and control variables for 5,499 municipalities. To get the dose-response function, the weights generated will be applied to the 2,699 municipalities for which there is homicide rate data in 2000. This is enough to assume that GBM algorithm will generate better weights than the linear methods, and that the treatment effect estimators will be unbiased, though referring to a more restricted sample of Brazilian municipalities.

Thus, the methodology of this study consists of the following steps: initially, the relationship between treatment and control variables in the sample is examined by the GBM algorithm, capable of representing non-linear relationships and including interactions between variables, which ensures the stability of the weights. Next, the algorithm generates a linear combination of regression trees: a smooth function that minimizes Pearson correlation between T and $\mathbf{X}$, from which the GPSs are extracted (and by extension, the weights by IPW). Applying the weights generates a balanced pseudo-sample in which differences in the outcome values can be attributed to the treatment. On the weighted sample, a generalized linear model is run to obtain the dose-response function expressed by Equation 6:

$$E[Y(t)] = \alpha_0 + \alpha_1 t. \tag{6}$$

The dose-response function provides the average treatment effect on the outcome variable. This study assumes that the effect of the treatment on the outcome is linear: after all, ATE is an average. This is a reasonable assumption considering the distance between the cause and effect studied, as well as the limitations of the data sources regarding the representation of the natural experiment observed. As time passes, the distance between historical events and contemporary scenario increases, and so lethality levels become relatively more dependent on more recent phenomena. The direct impact of historical events "slows down," so that their verifiable effects can be represented by simpler functional forms. Therefore, the dose-response function represented by Equation 6 provides α_1, the average causal effect of slavery density in 1872 on municipal homicide rates in 2000.

Note that it is not necessary to include the control vector in the generalized linear model expressed by Equation 6. Like Greifer (2019), this study uses robust standard errors when estimating the dose-response function, so that α_1 (the estimated effect of the t values on the expected value of potential outcomes) is causal and robust. As the information contained in vector **X** is adequately represented by the applied weights, it is not necessary to include control variables in the dose-response function to obtain robust estimators.

However, including control variables in the potential outcomes model leads to doubly robust estimators of the average treatment effect (Greifer, 2021), i.e., estimators that are consistent even if the specification of Equation 5 or the estimates of the GPS are incorrect (Graham, McCoy, and Stephens, 2015). Beyond the precautions described earlier, it is impossible to determine if Equation 5 is correctly specified. Therefore, the doubly robust estimator ensures the consistency of the results despite the difficulties inherent to modeling the effects of distant historical events.

Therefore, after executing the GPS-IPW method and weighting the observations, the generalized linear model that includes all explanatory variables provides the doubly robust treatment estimator (Zhang, Zhou, Cao, and Zhang, 2012):

$$\hat{\mu}(t) = \frac{1}{n}\sum_{i=1}^{n} g(t, X_i; \widehat{\beta^*}) \tag{7}$$

In Equation 7, $\widehat{\beta^*}$ represents the estimated coefficients of the weighted regression model $g(t, X_i; \widehat{\beta^*})$. As seen earlier, t represents the treatment levels, and X_i represents the observed value of control X to the municipality *i*. X_i is part of the control vector **X**. The coefficient of the treatment variable is a causal estimate, while the coefficients of the control variables represent the correlation between the control variables and the outcome variable in the pseudo-sample (in which it's possible to obtain the causal effect of the treatment). The models specified in the next section are the generalized linear models $g(t, X_i; \widehat{\beta^*})$ represented by Equation 7.

3.3 Data and model specification

The slavery density *sld* provides the share of enslaved individuals in the total population of a given municipality in 1872. It is based on the 2010 Territorial Division (IBGE, 2010), which readjusted the Imperial Census of 1872 to the most recent municipal division using Comparable Minimum Areas (or Áreas Mínimas Comparáveis (AMCs), in Portuguese). The correspondence between territorial divisions allows the treatment, outcome and control variables to refer to the same areas despite corresponding to different time periods.

The outcome variable *homr* indicates the municipal homicide rate in 2000, i.e., the number of homicide victims per hundred thousand inhabitants of a given municipality that year. It is based on data from the Mortality Information System of the Unified Health System (SIM-DATASUS) and the 2000 Demographic Census (IBGE, 2000). Traditionally, SIM-DATASUS defines homicides as causes of death related to the broad groups X85 to X90, Y35, and Y36 of the tenth International Classification of Diseases (CID-10, 1993). The first interval refers to deaths by Assault, and the last two groups refer to deaths by Legal Interventions (i.e., the perpetrator is a public agent).

The control variables were chosen considering the model developed by Buonanno and Vargas (2017) for Colombia, as well as the specifications of Gouda and Rigterink (2017) for the United States, and Justus, Khan, and Cerqueira (2016) for the state of São Paulo. The year 2000 was chosen as it is most recent one to present the closest data corresponding to the controls used by these authors.

The average per capita household income of a municipality, *ypc,* also comes from the 2000 Demographic Census (IBGE, 2003). The values were adjusted to the 2010 minimum wage (R$ 510.00), the last year of the available historical series, using the National Consumer Price Index (INPC), also calculated by IBGE. The variable *pd* corresponds to the municipal population density in 2000, measured in inhabitants per square kilometer (hab./km²). The total resident population was extracted from the 2000 Demographic Census, and the municipality's area (km²), from the 2010 Territorial Division.

The following were also extracted from the 2000 Demographic Census: the municipal unemployment rate *unmp*, which represents the percentage of the economically active resident population, aged 16 or over, without work in the reference week; the basic education indicator *be*, given by the percentage of individuals aged 16 or over who completed eight years of study or more; and the percentage of young men (15 to 29 years old) in the total resident population, represented by *ym*. This group is considered more prone to adopting risk behaviors that may result in victimization (Buonanno and Vargas, 2017).

Buonanno and Vargas (2017) used an indicator of public safety effectiveness based on the number of solved homicide cases in Colombian municipalities. In the absence of a similar indicator for Brazil, this study introduces the municipal public safety index *cps,* described below, to control for the government's role in municipal security.

Conducted by IBGE, the Municipal Basic Information Survey (called Profile of Brazilian Municipalities until 2015, but currently known as MUNIC) provides detailed municipal data on various topics. The 2001 edition was especially dedicated to public administration (IBGE, 2003). Although justice and security are state and federal competencies in Brazil, municipalities have agencies that can collaborate to reduce crime.

Based on MUNIC-2001, *cps* checks for the presence of the Municipal Guard, Women's Police Station, Civil Defense, Small Claims Court and Guardianship Council in each municipality. For each established agency, the municipality is assigned a value of one, and for each absent agency, a value of zero. The observed value of *cps* is the sum of the values assigned to the municipality, ranging from zero to five. The index consists of counting the present agencies, without specifying them. However, *cps* will be treated as a continuous variable, as explained in the following paragraphs.

Consider a municipality that, for example, has presented three out of the five public agencies observed by *cps*. Is it possible to expect that implementing a fourth agency will bring more public safety and lower homicide rates? That means: is it possible to expect a discrete and inverse relationship between *cps* and *homr*? That's difficult to affirm without further and more specific information. This implementation does not imply a discrete decrease in the homicide rate because the existence of the new agency does not imply the effectiveness of its action in combating violence. However, the increase suggests an expansion of access to conflict mediation and complaint channels, as well as an increase in the probability of institutional punishment of perpetrators. Thus, the effect of a new public safety agency is vaguer and more abstract than that the categories of a discrete variable could represent.

Buonanno and Vargas (2017) worked with the number of solved homicide cases, so they could measure the *effectiveness* of public action when it came to security. Due to the lack of comparable information for Brazilian municipalities, *cps* index can only measure the presence of (municipal) security public agencies, not the impact of their presence on municipal safety. Therefore, it can be unreasonable to expect a perfectly inverse and discrete relationship between *cps* and *homr*.

Compared to an effectiveness index, *cps* probably has a smoother correlation with the municipal homicide rates. This smoothness, which does not characterize the behavior of discrete variables, allows *cps* to be treated as a continuous variable. This choice not only simplifies the model, but also helps to make it more realistic, adhering to the meaning of the proxy variable and not to its specificities. According to recent literature, *cps* is expected to be negatively correlated with *homr*, as there is empirical evidence that greater public safety infrastructure is correlated with lower victimization rates (Justus et al., 2016).

Finally, municipal latitude and longitude, available in the 2010 Territorial Division (IBGE, 2010), were included as geographic controls. A fixed-effects model is obtained by including binary variables related to the states where each municipality is located. One of the twenty-six Brazilian was excluded from the specifications to avoid perfect collinearity. Fixed effects capture hard-to-measure factors that act on municipalities in the same state, i.e., with relatively common history and culture.

Estimating the average effect of slavery in 1872 on Brazilian municipalities (treated or controlled) in 2000 can only reasonable under the assumption that, *ceteris paribus*, these effects spilled over from treated municipalities (occupied in 1872, with enslaved inhabitants among the total population) to controlled municipalities (not occupied in 1872). As fixed effects address the non-specifiable peculiarities of each state, they make it possible to directly compare all municipalities, ensuring that the ATE estimator is valid for the entire population of municipalities in the observed area.

So far, the treatment variable, the outcome variable and the controls common to all specified models have been described. Income inequality, on the other hand, is considered a confounding variable because it is related to both slavery and the municipal homicide rate (Soares, 2000; Soares and Naritomi, 2010; Bertocchi and Dimico, 2010; 2014; Gouda and Rigterink, 2017; Fujiwara, Laudares, and Valencia-Caicedo, 2017; Buonanno and Vargas, 2017). To check the robustness of the estimated treatment effect, two different income inequality indexes were included among the controls of the last two models.

The Gini index was extracted from the 2000 Demographic Census (IBGE, 2003), and *gini* assumes values between zero and one: the former corresponding to perfect equality and the latter to perfect inequality (i.e., only one inhabitant of the municipality would possess all income, characterizing a perfectly unequal municipality). As a robustness check, the Theil index (Buonanno and Vargas, 2017), available in the Human Development Atlas of the United Nations Development Programme (UNDP), replaces the Gini index in the last model. The *theil* variable assumes zero in case of perfect equality and tends to infinity in case of perfect inequality. Table 1 summarizes the variables used.

Table 1 – Description of Variables

Variable	Description
homr	Municipal homicide rate (victims/100,000 inhabitants) Mortality Information System (SIM-DATASUS) 2000 Demographic Census (IBGE, 2003)
sld	Slavery density (%) Percentage of enslaved individuals in the total population of the municipality in 1872. 1872 Imperial Census - 2010 Territorial Division (IBGE, 2010)
ypc	Average per capita household income (R$) Values corrected by the National Consumer Price Index (INPC) based on the minimum wage of 2010 (R$ 510.00). 2000 Demographic Census (IBGE, 2003)
pd	Population density (inhabitants/km²) 2000 Demographic Census (IBGE, 2003) Territorial Division of 2010 (IBGE, 2010)
unmp	Unemployment rate (%) Portion of the resident population in the municipality that, economically active and aged 16 or older, is without work in the reference week. 2000 Demographic Census (IBGE, 2003)
be	Basic education (%) Portion of the resident population aged 16 or older with eight years of completed study. 2000 Demographic Census (IBGE, 2003)
ym	Young men (%) Portion of men aged 15 to 29 in the resident population. 2000 Demographic Census (IBGE, 2003)
cps	Municipal index of justice and public safety (scale from 0 to 5) Presence of Municipal Guard, Women's Police Station, Civil Defense, Small Claims Court, and Guardianship Council in the municipality. Basic Municipal Information Survey/Profile of Brazilian Municipalities (MUNIC-2001) (IBGE, 2003)
long	Longitude (Degrees, minutes, and seconds North-South) Territorial Division of 2010 (IBGE, 2010)
lat	Latitude (Degrees, minutes, and seconds East-West) Territorial Division of 2010 (IBGE, 2010)
gini	Gini income inequality index (range from 0 to 1) Demographic Census of 2000 (IBGE, 2003)
theil	Theil inequality index (greater than or equal to zero) Human Development Atlas in Brazil (UNDP, 2010)

Source: own elaboration.

With the information above, it is possible to specify the generalized linear models of Equation 7: the dose-response functions between Y = *homr* and T = *sld*, whose ATE estimators become doubly robust after the inclusion of the controls described in Table 1. It is useful to recall that the coefficients of *sld* are causal estimators, while the coefficients of the other independent variables are correlations with the outcome variable that were obtained after sample weighting. This process makes *sld* independent of the covariates in the pseudo-sample and balances differences between control and treatment groups.

To simplify the interpretation of the results tables of the next section, the model represented by Equation 8 will be referred to as Model 1. For a given municipality i, we have:

$$\begin{gathered} homr_{i,2000} = \beta_0 + \beta_1\, sld_{i,1872} + \beta_2\, ypc_{i,2000} + \beta_3\, pd_{i,2000} + \beta_4\, unmp_{i,2000} \\ + \beta_5\, be_{i,2000} + \beta_6\, ym_{i,2000} + \beta_7\, cps_{i,2000} + \beta_8\, long_{i,2010} + \beta_9\, lat_{i,2010} \\ + \sum_{j=1}^{25} STATE_{ij} + \varepsilon_{i,2000}. \end{gathered} \tag{8}$$

As shown in Equations 9 and 10, Models 2 and 3 include, respectively, the Gini index and the Theil index to represent income inequality:

$$\begin{gathered} homr_{i,2000} = \beta_0 + \beta_1\, sld_{i,1872} + \beta_2\, ypc_{i,2000} + \beta_3\, pd_{i,2000} + \beta_4\, unmp_{i,2000} \\ + \beta_5\, be_{i,2000} + \beta_6\, ym_{i,2000} + \beta_7\, cps_{i,2000} + \beta_8\, long_{i,2010} + \beta_9\, lat_{i,2010} \\ + \beta_{10}\, gini_{i,2000} + \sum_{j=1}^{25} STATE_{ij} + \varepsilon_{i,2000} \end{gathered} \tag{9}$$

$$homr_{i,2000} = \beta_0 + \beta_1\, sld_{i,1872} + \beta_2\, ypc_{i,2000} + \beta_3\, pd_{i,2000} + \beta_4\, unmp_{i,2000} + \beta_5\, be_{i,2000} + \beta_6\, ym_{i,2000} + \beta_7\, cps_{i,2000} + \beta_8\, long_{i,2010} + \beta_9\, lat_{i,2010} + \beta_{10}\, theil_{i,2000} + \sum_{j=1}^{25} STATE_{ij} + \varepsilon_{i,2000} \quad (10)$$

Besides providing another robustness test and eliminating confounding bias, the inclusion of interactions between average household per capita income *ypc* and income inequality indices *gini* and *theil* aims to acknowledge that relationships between income and inequality may take different forms in different observed areas, and these forms may be significantly related to violence levels. For example, some regions may exhibit high levels of household income, inequality and violence, while others may have low homicide rates despite having high inequality. Such a scenario can be motivated by economic changes that are leading to a long run increase in average income, which is commonly associated with reduced economic incentives for violence.

The adjusted pseudo-R-squared is commonly used to assess the quality of specification in weighted generalized linear models (Nakagawa and Schielzeth, 2012). If it increases after including the interactions (and if other variables remain exhibiting the expected behavior), there is an indication that modeling the relationship between income and inequality contributes to the representativeness of the models. Equations 11 and 12 represent, respectively, Models 4 and 5:

$$homr_{i,2000} = \beta_0 + \beta_1\, sld_{i,1872} + \beta_2\, ypc_{i,2000} + \beta_3\, pd_{i,2000} + \beta_4\, unmp_{i,2000} + \beta_5\, be_{i,2000} + \beta_6\, ym_{i,2000} + \beta_7\, cps_{i,2000} + \beta_8\, long_{i,2010} + \beta_9\, lat_{i,2010} + \beta_{10}\, gini_{i,2000} + \beta_{11}(gini_{i,2000}\; x\; ypc_{i,2000}) + \sum_{j=1}^{25} STATE_{ij} + \varepsilon_{i,2000} \quad (11)$$

$$
\begin{aligned}
homr_{i,2000} = {} & \beta_0 + \beta_1\, sld_{i,1872} + \beta_2\, ypc_{i,2000} + \beta_3\, pd_{i,2000} + \beta_4\, unmp_{i,2000} \\
& + \beta_5\, be_{i,2000} + \beta_6\, ym_{i,2000} + \beta_7\, cps_{i,2000} + \beta_8\, long_{i,2010} + \beta_9\, lat_{i,2010} \\
& + \beta_{10}\, theil_{i,2000} + \beta_{11}(theil_{i,2000}\ x\ ypc_{i,2000}) + \sum_{j=1}^{25} STATE_{ij} + \varepsilon_{i,2000}
\end{aligned} \tag{12}
$$

Models 1 to 5 were estimated for the 2,699 Brazilian municipalities for which *homr* data were available in 2000. The results of the estimations are displayed and analyzed in the following section.

4 RESULTS

On the next page, Table 2 shows that, for all models, the ATE estimator was significant at 5%, which means that the slavery density in 1872 is related to higher homicide rates in Brazilian municipalities in 2000. As more controls were added, *sld* not only remained positive and significant at 5%, but also the pseudo-R-squared and pseudo adjusted R-squared increased. This suggests that the results obtained are robust. However, the balance tables displayed in Appendix B show that the GBM algorithm was unable to provide weights that fully eliminated the correlation between the treatment variable and the controls in the pseudo-sample. Therefore, the results below cannot be considered causal. They represent, however, the first quantitative evidence that slavery relates to violence in contemporary Brazil.

Table 2 – Effect of slavery density on the municipal homicide rate

	Homicides per hundred thousand inhabitants (2000)				
	(1)	(2)	(3)	(4)	(5)
sld	0.117**	0.122**	0.120**	0.119**	0.116**
	(0.052)	(0.052)	(0.052)	(0.052)	(0.053)
ypc	0.001	0.002	0.002	-0.050***	-0.016**
	(0.004)	(0.004)	(0.004)	(0.013)	(0.007)
pd	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
unmp	0.194***	0.208***	0.201***	0.138*	0.146*
	(0.075)	(0.076)	(0.075)	(0.076)	(0.076)
be	-0.149**	-0.163**	-0.159**	-0.053	-0.062
	(0.074)	(0.074)	(0.075)	(0.075)	(0.079)
ym	1.115***	1.054***	1.074***	1.067***	1.103***
	(0.400)	(0.408)	(0.403)	(0.403)	(0.399)
cps	-1.429***	-1.413***	-1.418***	-1.302***	-1.310***
	(0.321)	(0.318)	(0.319)	(0.315)	(0.317)
long	-0.140	-0.160	-0.157	-0.096	-0.113
	(0.208)	(0.210)	(0.209)	(0.210)	(0.209)
lat	-0.397*	-0.386*	-0.388*	-0.447*	-0.451*
	(0.229)	(0.230)	(0.230)	(0.231)	(0.233)
gini		-5.762		-30.697***	
		(5.858)		(8.503)	
theil			-2.214		-8.919***
			(2.676)		(3.243)
*ypc*gini*				0.075***	
				(0.018)	
*ypc *theil*					0.018***
					(0.005)
Intercept	6.314	9.294	7.043	27.942*	13.500
	(15.158)	(15.253)	(15.198)	(15.858)	(15.283)
Observations	2,699	2,699	2,699	2,699	2,699
AIC	173,225	174,980	174,931	177,148	176,494
Pseudo-R²	0.243	0.243	0.243	0.249	0.247
Pseudo-R² Adj.	0.233	0.233	0.233	0.239	0.236

Note: *p<0,1; **p<0,05; ***p<0,01

The sample of the models in Table 2 consists of municipalities affected by slavery in 1872 and municipalities that had not yet been occupied in that year. As results apply to all municipalities in the sample, they suggest that there was a spillover of the effects of slavery not only over time but also across the entire national territory.

Considering that the 1872 data illustrates the state of Brazilian slavery after the end of the transatlantic slave trade (1850) just sixteen years before Abolition (1888), the results gain remarkable strength. It was possible to find a positive and significant correlation between slavery and contemporary violence when this system was in decline. Also, these results are considered causal under partial balance (see balance tables in Appendix B). More than a century later, it is possible to speculate that, for data representative of the peak of slavery, the effects would go in the same directions, but would also be more palpable and concrete. Limitations in the representativeness of the treatment variable must be considered, and even though causality cannot be claimed, this hypothesis cannot be ruled out.

An illustrative exercise helps understanding the results more concretely. Considering that the *sld* coefficient represents a causal effect, according to Table 2, if the share of enslaved people in the population of all municipalities (both treated and untreated) in 1872 increased by 1%, the municipal homicide rates in 2000 would increase, on average, by about 0.12 (since the *sld* coefficients range between 0.116 and 0.122).

In 2000, the 2,699 municipalities that reported data counted 45,433 homicide deaths. Applying the hypothetical 1% increase in slavery density to the results of Table 2, simple calculations show that the absolute number of homicides would increase in 0.37% according to models 1 and 4; 0.38% according to models 2 and 3, and 0.36% according to model 5. Models 1 to 5 indicate that total homicides would increase by 166, 174, 171, 169 and 165 victims, respectively.

Considering the behavior of the control variables (and the values of pseudo-R-squared and adjusted pseudo-R-squared), Model 4 seems to be the most appropriate. In Models 2 and 3, the inclusion of *gini* and *theil* did not lead to significant changes in pseudo-R-squares. Inequality variables did not show a significant or positive correlation with *homr*. This counterintuitive behavior also occurred concerning *ypc*, average per capita household income. The variables that behaved as expected by the literature and showed the highest levels of significance were population density *pd*, unemployment rate *unmp* and municipal public safety index *cps*.

The *cps* coefficients suggest a correlation between higher levels of public safety infrastructure and lower homicide rates in Brazilian municipalities. Given the scarcity of official data on public safety efficiency, especially at municipal level, it is noteworthy that *cps* exhibited reasonable and robust behavior, consistent with expectations.

The percentage of people with at least eight years of education *eb* is significantly related to lower municipal homicide rates in Models 1 to 3. However, this behavior does not withstand the inclusion of controls in Models 4 and 5. On the contrary, the percentage of young men *ym* exhibited significant and consistent behavior with the literature in all models.

In Models 4 and 5, the inclusion of interaction terms between *ypc* and income inequality indexes led to significant changes. In addition to higher pseudo-R-squared and adjusted pseudo-R-squared, *gini* and *theil* became significant, as well as *ypc* and its interaction with inequality. In these models, higher levels of *ypc* are correlated with lower municipal homicide rates, which aligns with the literature. However, the same holds for absolute *gini* and *theil*, contrary to the literature. According to Model 4, the relationship between *ypc* and *gini* with *homr* is given by:

$$\frac{\partial homr}{\partial ypc} = -0{,}050 + 0{,}075gini \tag{13}$$

$$\frac{\partial homr}{\partial gini} = -30{,}697 + 0{,}075ypc \tag{14}$$

Equation 13 shows that if absolute inequality is zero, *ypc* and *homr* are inversely correlated. However, if inequality is high, it can lead to a direct relationship between them. That is: if inequality is low, increments in income are related to lower homicide rates; but if inequality is high, more income relates to higher homicide rates.

On the other hand, Equation 14 suggests that, if income is sufficiently low, inequality is correlated with lower levels of violence. However, it also shows that if income is sufficiently high, it can lead to a positive relationship between inequality and violence, as predicted by the literature.

From Equations 13 and 14, it’s possible to infer that inequality can aggravate violence levels in Brazilian municipalities: if it is high, it can cloud out the usual reductive effect of higher levels of income over homicide rates. In the weighted sample, cities with higher levels of both income and inequality tend to be more violent, such as metropolises like São Paulo and Rio de Janeiro, for example.

The results of model 5, which uses the Theil inequality index, represent the same dynamics, as shown in Equations 15 and 16:

$$\frac{\partial homr}{\partial ypc} = -0{,}016 + 0{,}018theil \tag{15}$$

$$\frac{\partial homr}{\partial theil} = -8{,}919 + 0{,}018ypc \tag{16}$$

Table 3 summarizes the conclusions extracted from interpreting the partial derivatives of Equations 13 to 16. Keep in mind they come from a 2,699 sample of Brazilian municipalities in 2000.

Table 3 –Income, Inequality and Homicide rates: Simple Interpretations

Equation	If…	Then…	Conclusion
13 and 15	$\uparrow inequality$	$\frac{\partial homr}{\partial ypc} > 0$	If inequality is high, more income correlates to higher homicide rates.
	$\downarrow inequality$	$\frac{\partial homr}{\partial ypc} < 0$	If inequality is low, more income correlates to lower homicide rates.
14 and 16	$\uparrow ypc$	$\frac{\partial homr}{\partial ineq.} > 0$	If income is high, more inequality correlates to higher homicide rates.
	$\downarrow ypc$	$\frac{\partial homr}{\partial ineq.} < 0$	If income is low, more inequality is related to less homicides

Table 3 suggests that the dynamics between income, inequality and violence in Brazilian municipalities presents some peculiarities. Literature states that *ypc* and *homr* usually have an inverse relationship: in the analyzed sample, that only occurs under low levels of inequality. Otherwise, the relationship becomes direct, which goes against common sense. At the same time, inequality is usually said to be directly related to violence levels. In the observed sample, that only occurs when average household per capita income is high. If it is low, inequality can be inversely correlated to homicide rates. As significant correlations at 1% were obtained for all these control variables, it’s worthy to take a deeper look at these results through Johnson-Neyman decomposition.

Figure 2 – Johnson-Neyman chart of Model 4

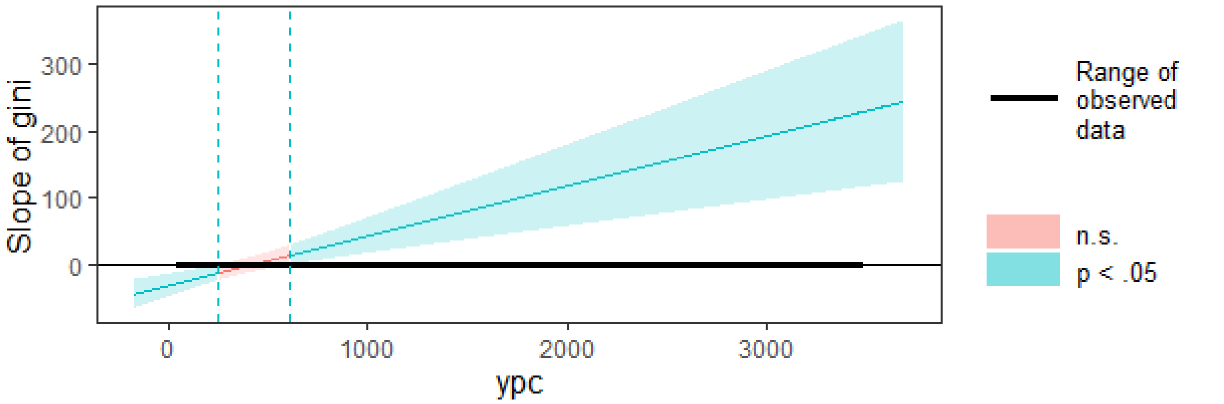


Figure 3 – Johnson-Neyman chart of Model 5

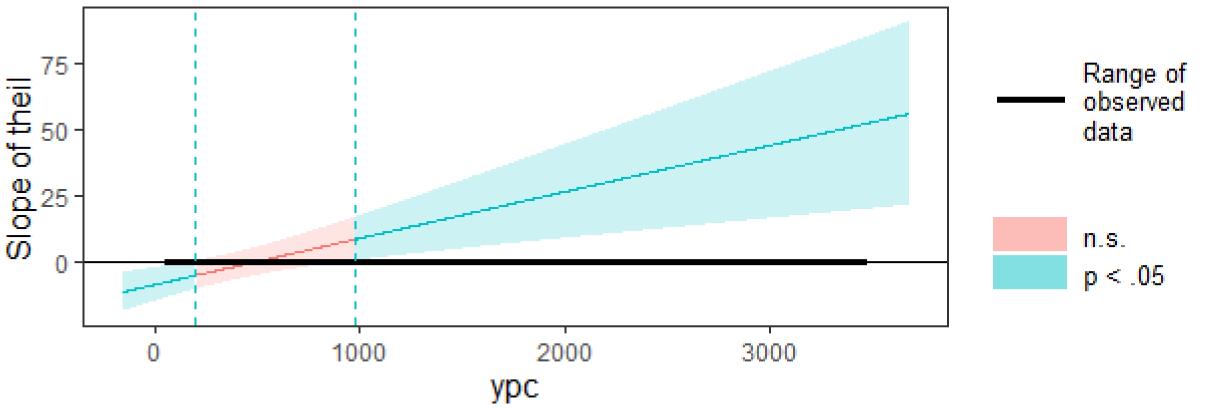


Figure 2 and 3 illustrate Equations 14 and 16, respectively. It can be seen that, for most of the observed *ypc* values in the sample, inequality has a positive and significant relationship with the homicide rates. The variable *ypc* takes values between R$ 54.18 and R$ 3,468.20. According to Equation 14, when *ypc* = R$ 409.29, *gini* starts to have a positive relationship with *homr*. In the case of Equation 16, *theil* starts to have a positive relationship with *homr* when *ypc* = R$ 495.50. Both charts show that the relationship between inequality and homicide rates is not significant throughout all the *ypc* values observed. Figures 4 and 5 dive deeper into this conclusion.

Figure 4– Income, inequality and homicide rates in Model 4

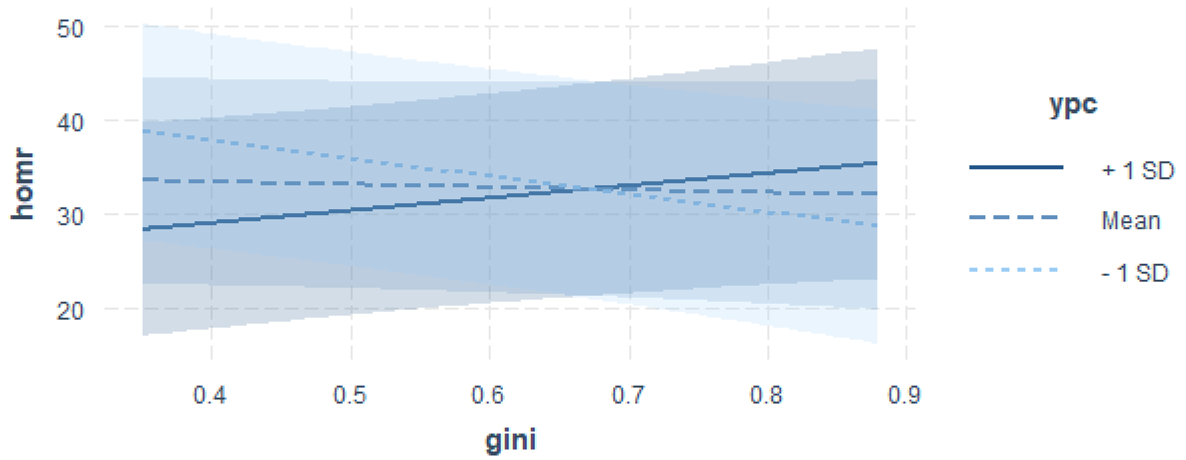
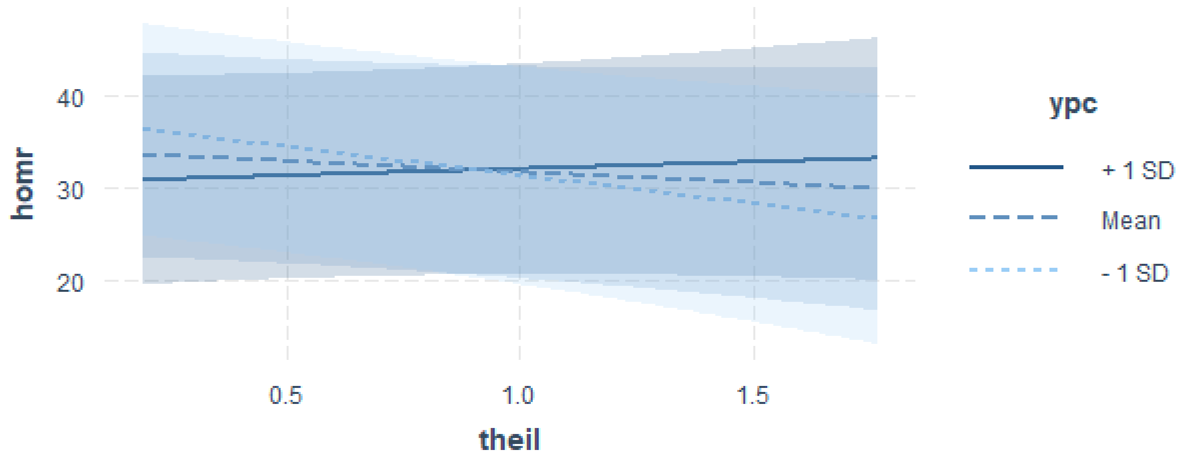


Figure 5– Income, inequality and homicide rates in Model 5



In Figures 4 and 5, the relationship between income inequality and homicide rates is represented for different levels of income observed in the sample. The small-dash light blue line represents municipalities whose income is lower than the sample mean by one standard deviation. Municipalities whose income is equal (or sufficiently close) to the sample mean are represented by the large-dash intermediate blue line. And finally, municipalities with income higher than the sample mean by one standard deviation are represented by the solid dark blue line. The means and their confidence intervals (shaded areas in Figures 4 and 5) are detailed in Table 4 below.

Table 4 - Johnson-Neyman intervals for Brazil

Gini Index (Model 4)				Theil Index (Model 5)			
ypc (R$)	*â*	2.5%	97.5%	*ypc* (R$)	*â*	2.5%	97.5%
157.68	-18.93	-32.05	-5.81	157.59	-6.14	-11.72	-0.56
373.02	-2.86	-14.30	8.58	373.10	-2.33	-7.53	2.86
588.37	13.21	-1.23	27.65	588.60	1.47	-4.27	7.21

In Table 4, *â* is the slope of inequality in relation to *homr*. Combining Table 4 and Figures 4 and 5, we see that municipalities represented by the small-dash light blue line have an income of approximately R$ 157, one standard deviation below the sample average. They represent low-income municipalities, and for them, the relationship between absolute inequality and *homr* assumes negative values across its entire confidence interval. That is: for municipalities with lower income levels relative to the sample, higher inequality is associated with lower homicide rates.

The sample mean of *ypc* is approximately R$ 373. For municipalities with intermediate income (large-dash intermediate blue line), confidence intervals suggest that the relationship between inequality and homicide rates can be either negative or positive. Figures 4 and 5 show that, if *ypc* takes a value close to or equal to the sample average, the relationship between inequality and the homicide rate isn’t very clear: it cannot be considered significant, which Figures 2 and 3 also show.

As Figures 4 and 5 show, inequality and the homicide rates are positively/directly related when *ypc* is one standard deviation above the sample mean, as we can see from Table 4). Figures 2 and 3 suggest that, for these levels of *ypc*, this relationship is significant.

In Figures 4 and 5, high-income municipalities are represented by solid dark blue lines, for which *ypc* is equal to approximately R$ 588, one standard deviation above the sample mean. If we make Equations 14 and 16 equal to zero, they show that the relationship between income inequality and municipal homicide rates only becomes positive from *ypc* = R$ 409.29 on in Model 4 and from *ypc* = R$ on 495.50 in Model 5. So, it's possible to say that the relationship between inequality and *homr* tends to be positive for this group: $\hat{a}$ assumes positive values across most of the confidence intervals corresponding to this income level (Table 4). This suggests that, in high-income municipalities, higher levels of inequality are associated with higher levels of lethal violence, as advocated by the Crime Economics literature.

The Johnson-Neyman results show that, for both inequality indexes observed, the relationship between inequality and homicide rates is only significant for municipalities with higher or lower income levels relative to the sample mean. Higher inequality is correlated to higher homicide rates in richer municipalities, but it correlates to lower homicide rates in poorer municipalities. The causes of these results are beyond the scope of this study, but it is possible to assume that the relationship between income and economic incentives for crime contributes to explaining them, as the following illustrative exercise shows.

Suppose that, in a given municipality with an unknown initial value of *ypc*, an exogenous shock occurs: therefore, poverty decreases among the total population. After the shock, residents who significantly raised their standard of living have fewer incentives to commit violent acts: they now have more to lose and less to gain by acting violently. In other words, an increase in income (or a decrease in poverty) increases the opportunity cost of violence, and thus, contributes to discourage it.

According to Models 4 and 5 in Table 1, the relationship between income and homicide rates is inverse (*ypc* coefficients are negative and significant at 1%), which corroborates this hypothesis. This holds for the observed municipalities regardless of whether they are rich or poor. But the relationship between *ypc* and absolute inequality is different for each of these groups, and that's why the relationship between inequality and *homr* differs by income level.

If the population of a municipality has a generally low level of income and a positive exogenous shock occurs, the improvement in the standard of living of part of the inhabitants tends to generate significant differences between beneficiaries and non-beneficiaries. Therefore, inequality increases. At the same time, the improvement in the standards of living reduces incentives of committing violent acts. This can explain why higher levels of inequality are related to lower homicide rates among poorer municipalities. In this group, a raise in income generates inequality among the population but reduces incentives for violent crime because the marginal utility of income is high for a large part of the inhabitants (who generally have precarious living standards).

In high-income municipalities, the dynamics between income, inequality and lethal violence is different. A positive income shock that is sufficient to discourage violent acts among the most economically vulnerable is likely to reduce inequality, so this becomes correlated with lower homicide rates. That's why, among the municipalities with higher income levels in the sample, it was observed that higher homicide rates are associated with higher levels of inequality, in accordance with Crime Economics literature.

For the richer and poorer municipalities in the sample, we can infer that the marginal utility of income is high. That is: for both groups, an increase in the average income contributes to reduce incentives to commit violent acts associated with economic vulnerability. This increase raises inequality in poorer municipalities, and reduces it in richer ones, but in both cases, it is related to lower levels of lethal violence.

5 CONCLUSIONS

This study partially achieved its goal, as it obtained an empirical, positive, and significant relationship between slavery in 1872 and homicide rates in Brazilian municipalities in 2000, a relatively recent period. The GPS-IPW method was unable to eliminate the confounding bias resulting from the correlation between the treatment variable and the controls. In Appendix B, the balance tables show that, after weighting, some controls are still related to the treatment variable – although the GBM method of calculating weights had remarkable success among most controls. Therefore, it is not possible to assert that the observed effects are indeed causal. Considering the limitations of the data sources on the treatment variable, the distance between the observed events and the strong sociological support for the analyzed hypothesis, it is still not possible to rule it out.

The study provides quantitative evidence that the effects of slavery are related to violence in contemporary Brazil, and the confounding bias suggests that it also affects other socioeconomic factors given by the control variables. Thus, it fills the gap in empirical studies of the historical causes of violence in Brazil. At the same time, the positive and significant results suggest that specifications with other controls and more advanced machine learning methods may yield more robust results in the future (especially considering the increasing concern about data collection on violence in Brazil).

The results suggest that the harms of slavery would have surpassed not only the Abolition (1888) but also institutional and economic changes that lead many observers to believe that the slave legacy had been overcome. Despite the underrepresentation of the treatment variable and its temporal distance from the outcome, the effects were significant one hundred twelve years after the Abolition of Slavery, surpassing the fall of Military Dictatorship (1975-1985), the proclamation of the 1988 Constitution and the success of the Real Plan (1994) in defeating chronic inflation. A more precise and empirical description of the transmission mechanisms between slavery and contemporary obstacles to economic and social development in Brazil is to be explored by Economics of Crime.

Empirical and sociological literature considers income inequality as a legacy of slavery and a contemporary determinant of homicide rates. In addition to contributing to the robustness of the estimated effect, the inclusion of an interaction term between average per capita household income and different inequality indexes among the controls allowed the observation of different dynamics between income, inequality and lethal violence in Brazilian municipalities.

In municipalities where average per capita household income is lower, higher levels of inequality could mean that economic vulnerability is decreasing among a considerable share of the population, reducing their incentives to commit violent acts. In poorer municipalities, increasing inequality can be a sign of increasing income among the most vulnerable, which contributes to turn violent crime more costly. That’s a possible explanation for why inequality is associated with lower homicide rates in low-income municipalities of the sample.

Among the wealthiest municipalities, an increase in income of the poorest has the same effect and is also associated with lower homicide rates. However, in this case, the increase in income is likely to reduce the inequality among the total population. Therefore, lower levels of inequality are associated with lower homicide rates in municipalities with higher relative income in the sample. In municipalities with intermediate relative income, inequality wasn't found to have a significant relationship with homicide rates.

Finally, this study found a significant and positive correlation between the percentage of enslaved people in Brazilian municipalities in 1872 and municipal homicide rates in 2000, using a territorial division that equalizes municipal areas observed in both periods. This relationship is valid for the entire national territory, even though most municipalities in the 2000 territorial division were not yet occupied in 1872. These conclusions contribute to easier refutation of potential questions about the harmful nature of slavery, which may or may not be motivated by racism.

REFERENCES

Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of Modern World Income Distribution. *Quarterly Journal of Economics, 117*(4), 1231-1294. https://economics.mit.edu/files/4127

Acemoglu, D., García-Jimeno, C., & Robinson, J. A. (2012). Finding Eldorado: Slavery and Long-Run Development in Colombia. *Journal Of Comparative Economics*, *40*(4), 534-564. https://economics.mit.edu/files/10401

Acharya, A., Blackwell, M., & Sen, M. (2016). The Political Legacy of American Slavery. *The Journal of Politics, 73*(3), 621-641. https://scholar.harvard.edu/files/msen/files/slavery.pdf

Adorno, S. (1996). A gestão urbana do medo e da insegurança. São Paulo, 281 p. Tese (Livre-Docência). Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo.

Amorim, C. (1993). *Comando Vermelho: A História Secreta do Crime Organizado*. Rio de Janeiro: Record.

Amorim, C. (2013). *CV PCC: A Irmandade do Crime*. Rio de Janeiro: Record.

Arcand, J.-L., & D'Hombres, B. (2004). Racial discrimination in the Brazilian labour market: wage, employment and segregation effects. *Journal of International Development, 16*(8), 1053–1066.

Barros, L., & Silva, M. S. (2020). Between sticky floors and glass ceilings: The effect of trade liberalization on double discrimination in Brazil. Göttingen: Georg-August-Universität Göttingen, Ibero-America Institute for Economic Research (IAI). http://hdl.handle.net/10419/226343

Bertocchi, G., & Dimico, A. (2010). The Racial Gap in Education and the Legacy of Slavery. *Journal of Comparative Economics, 40*(4), 581-595.

Bertocchi, G., & Dimico, A. (2014). Slavery, Education and Inequality. *European Economic Review, 70*, 197-209. https://www.sciencedirect.com/science/article/pii/S0147596712000261

Botelho, F., Madeira, R. A., & Rangel, M. A. (2015). Racial Discrimination in Grading: Evidence from Brazil. *American Economic Journal: Applied Economics, 7*(4), 37–52.

Bueno, S. (2014). Bandido bom é bandido morto: a opção ideológico-institucional da política de segurança pública na manutenção de padrões de atuação violentos da polícia militar paulista. (Master’s thesis). Escola de Administração de Empresas de São Paulo, Fundação Getulio Vargas, São Paulo, Brazil.

Brasil, Ministério da Saúde. (s.d.). *SIM: Sistema de Informações sobre Mortalidade*. http://tabnet.datasus.gov.br/

Buonanno, P., & Vargas, J. F. (2017). Inequality, Crime and The Long-Run Legacy of Slavery. *Journal of Economic Behavior & Organization, 159*, 539-552.

Cano, I., & Duarte, T. (2012). *No Sapatinho: A Evolução das Milícias no Rio de Janeiro (2008-2011)*. Rio de Janeiro: Laboratório de Análise da Violência (LAV-UERJ) e Fundação Henrich Böll.

Case, A., & Deaton, A. (2020). *Deaths of Despair and the Future of Capitalism*. Princeton: Princeton University Press.

Cerqueira, D. (2014). Causa e Consequências do Crime no Brasil. Thesis (Doctoral dissertation). Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, Rio de Janeiro, Brazil. http://twixar.me/s8pn

Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Houghton Mifflin Boston, MA.

Engerman, S., & Sokoloff, K. (1997). Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: a View from Economic Historians of the United States. *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico*. Stanford: Stanford University Press. https://www.nber.org/papers/h0066

Engerman, S., & Sokoloff, K. (2002). Factor Endowments, Inequality and Paths of Development among New World Economies. *Journal of the Latin American and Caribbean Economic Association, 0*, 41-110. https://www.nber.org/papers/w9259

FBSP – Fórum Brasileiro de Segurança Pública. (2023). *17º Anuário Brasileiro de Segurança Pública*. São Paulo: Fórum Brasileiro de Segurança Pública. https://forumseguranca.org.br/wp-content/uploads/2023/07/anuario-2023.pdf

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford: Stanford University Press.

Fernandes, F. (1965). *A integração do negro na sociedade de classes: o legado da "raça branca"*. Vol. 1. São Paulo: Dominus Ed.

Fernandes, F. (1978). *A integração do negro na sociedade de classes: no limiar de uma nova era*. Vol 2. São Paulo: Atica.

Fernandes, G. A. D. A. L. (2015). Brazilian Female Labor Market: Racial-Skin Color Discrimination and Inefficiency. *Economia Aplicada, 19*(2), 241–259.

Fiuza-Moura, F. K., et al. (2019). The luck of being of the right gender and color: a detailed discussion about the wage gaps in the Brazilian manufacturing industry. *Quality & Quantity, 53*(3), 1275–1300.

Franco, M. S. de C. (1983). *Homens livres na ordem escravocrata*. 3ª ed. São Paulo: Kairós.

Freyre, G. (1973). *Casa Grande & Senzala*. 16ª ed. Rio de Janeiro: José Olympio.

Fujiwara, T., Laudares, H., & Valencia-Caicedo, F. (2017). Tordesilhas, Slavery and the Origins of Brazilian Inequality. https://www.insper.edu.br/wp-content/uploads/2018/08/fujiwara-et-al-tordesillas.pdf

Gerard, F., et al. (2021). Assortative Matching or Exclusionary Hiring? The Impact of Employment and Pay Policies on Racial Wage Differences in Brazil. *American Economic Review, 111*(10), 3418–3457.

Graham, D. J., McCoy, E. J., & Stephens, D. A. (2015). Doubly robust dose-response estimation for continuous treatments via generalized propensity score augmented outcome regression. https://arxiv.org/abs/1506.04991

Gouda, M., & Rigterink, A. (2017). The Long-Term Effects of Slavery on Violent Crime: Evidence from US Counties. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2358389

Greifer, N. (s.d.). A Guide to Using WeighIt for Balancing Covariates. https://cran.r-hub.io/web/packages/WeightIt/vignettes/WeightIt_A0_basic_use.html

Holanda, S. B. de. (1984). *Raízes do Brasil*. 18ª ed. Rio de Janeiro: José Olympio.

Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives, 226164*, 73–84.

Huling, J. D., Greifer, N., & Chen, G. (2023). Independence Weights for Causal Inference with Continuous Treatments. *Journal of the American Statistical Association*, 1–14. https://arxiv.org/abs/2107.07086

IBGE – Instituto Brasileiro de Geografia e Estatística. (2020a). *Censo Imperial, 1872.* https://biblioteca.ibge.gov.br/biblioteca-catalogo?view=detalhes&id=225477

IBGE – Instituto Brasileiro de Geografia e Estatística. (2020b). *Perfil dos Municípios Brasileiros: Gestão Pública 2001, 2003.* https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=283993

IBGE – Instituto Brasileiro de Geografia e Estatística. (2020c). *Censo Demográfico de 2000, 2003.* https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial.html

IBGE – Instituto Brasileiro de Geografia e Estatística. (2020d). *Divisão Territorial dos Municípios, 2010.* https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial.html

IBGE – Instituto Brasileiro de Geografia e Estatística. (2020e). *Evolução da divisão territorial do Brasil 1872-2010, 2011.* https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=284481

IBGE – Instituto Brasileiro de Geografia e Estatística. (2020f). *Características étnico-raciais da população: classificações e identidades, 2013.* https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=284235

IBGE – Instituto Brasileiro de Geografia e Estatística. (2020g). *Estimativas Populacionais, 2020.* https://sidra.ibge.gov.br/pesquisa/estimapop/tabelas.

IBGE – Instituto Brasileiro de Geografia e Estatística. (2022). *Desigualdades Sociais por Cor ou Raça no Brasil, 2022*. https://biblioteca.ibge.gov.br/visualizacao/livros/liv101972_informativo.pdf

IEP – Institute for Economics and Peace. (2023a). *Global Peace Index 2022*. Sydney. https://www.visionofhumanity.org/resources/global-peace-index-2022/

IEP – Institute for Economics and Peace. (2023b). *Global Peace Index 2023.* Sydney. https://www.visionofhumanity.org/resources/global-peace-index-2023/.

Jung, Y. (2019). *How the Legacy of Slavery Survives: Labor Market Institutions and Human Capital Investment.* Boston University. https://eh.net/eha/wp-content/uploads/2019/06/Jung.pdf

Justus, M. (2012). *Uma Abordagem Econômica das Causas da Criminalidade: Evidências para a Cidade de São Paulo.* Piracicaba: USP.

Justus, M.; Kahn, T.; Cerqueira, D. (2016). *O "Mistério de São Paulo" e o Papel do PCC na Redução de Homicídios nos Anos 2000.* Campinas: UNICAMP.

Leung-Gagné, J. (2023). *Classroom Segregation without Tracking: Chance, Legitimacy, and Myth in "Racial Paradise".* AERA Open, v. 9, p. 23328584231176456.

Lima, R. S.; Bueno, S.; Alcadipani, R. (2021). *Evolução das mortes violentas intencionais no Brasil.* In: FBSP – FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA. Anuário Brasileiro de Segurança Pública 2021. São Paulo: FBSP, p. 21-35. https://forumseguranca.org.br/wp-content/uploads/2021/07/1-evolucao-das-mortes-violentas-intencionais-no-brasil.pdf.

McCaffrey, D. F.; Ridgeway, G.; Morral, A. R. (2004). *Propensity score estimation with boosted regression for evaluating causal effects in observational studies.* Psychological methods, v. 9

Maloney, W. F.; Valencia-Caicedo, F. (2016). *The Persistence of (Subnational) Fortune.* Economic Journal, v. 126, n. 598, p. 2363-2401. https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12276.

Manso, B. P. (2020). *A República das Milícias: Dos Esquadrões da Morte à Era Bolsonaro.* São Paulo: Todavia.

Manso, B. P.; Dias, C. N. (2018). *A Guerra: A Ascensão do PCC e o Mundo do Crime no Brasil.* São Paulo: Todavia.

Miller, C.; Schmutte, I. M. (2021). *The Dynamics of Referral Hiring and Racial Inequality: Evidence from Brazil.* Econstor (Econstor), 1 set. 2021.

Nakagawa, S.; Schielzeth, H. (2012). *A general and simple method for obtaining R2 from generalized linear mixed-effects models.* Methods in Ecology and Evolution, v. 4, n. 2, p. 133–142.

North, D. C. (1990). *Institutions, institutional change and economic performance.* Cambridge: Cambridge University Press.

Nunn, N. (2008). *Slavery, Inequality and Economic Development: An Examination of the Engerman-Sokoloff Hypothesis.* In: Institutions and Economic Performance, p. 21-180. Cambridge: Harvard University Press.

Nunn, N.; Wantchekon, L. (2011). *The Slave Trade and the Origins of Mistrust in Africa.* American Economic Review, v. 101, n. 7, p. 3221–3252, dez. 2011.

Palma, N.; Papadia, A.; Pereira, T.; Weller, L. (2021). *Slavery and development in nineteenth century Brazil.* Capitalism: A Journal of History and Economics, v. 2, n. 2, p. 372–426.

Papadia, A. (2019). *Slaves, Migrants and Development in Brazil, 1872-1923.* Fiesole. https://cadmus.eui.eu/handle/1814/63508

Queiroz, M. (2021). *Clóvis Moura e Florestan Fernandes: interpretações marxistas da escravidão, da abolição e da emergência do trabalho livre no Brasil.* Revista Fim do Mundo, n. 4, p. 254–280.

Resende, J. P.; Andrade, M. V. (2011). *Crime Social, Castigo Social: Desigualdade de Renda e Taxas de Criminalidade nos Grandes Municípios Brasileiros.* Estudos Econômicos, v. 41, n. 1, p. 173-195.

Ribeiro, D. (1995). *O Povo Brasileiro.* São Paulo: Companhia das Letras.

Rosenbaum, P. R.; Rubin, D. B. (1983). *The central role of the propensity score in observational studies for causal effects.* Biometrika, v. 70, n. 1, p. 41–55.

Rosenthal, C. (2018). *Accounting for Slavery: Masters and Management.* Cambridge: Harvard University Press.

Sacerdote, B. (2005). *Slavery and the Intergenerational Transmission of Human Capital.* Review of Economics and Statistics, v. 87, n. 2, p. 217-234.

Schwarcz, L. M. (2012). *A Abertura para o Mundo: 1889-1930.* Rio de Janeiro: Objetiva.

Seyler, F. (2021). *Slavery, Political Attitudes and Social Capital: Evidence from Brazil.* Journal of Historical Political Economy, v. 1, n. 3, p. 377–409.

Soares, R. (2000). *Development, Crime and Punishment: Accounting for the International Differences in Crime Rates.* Chicago: University of Chicago.

Soares, R.; Naritomi, J. (2010). *Understanding High Crime Rates in Latin America: The Role of Social and Policy Factors.* In: The Economics of Crime: Lessons for and from Latin America. Chicago: University of Chicago Press. http://www.sebh.ecn.br/seminario_5/arquivo1.pdf

Souza, J. (2017). *A Elite do Atraso: da Escravidão à Lava-Jato.* Rio de Janeiro: LeYa.

Souza, S. dos S. (2021). *A influência da familiaridade e localização do fornecedor na decisão de compra em uma organização do sistema S.* (Master's thesis) Centro de Ciências Sócio-organizacionais, Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul, Brazil. http://guaiaca.ufpel.edu.br/handle/prefix/8191

Spaniol, M. I.; Moraes Jr., M. C.; Rodrigues, C. R. G. (2020). *Como tem sido planejada a Segurança Pública no Brasil? Análise dos Planos e Programas Nacionais de Segurança implantados pós-redemocratização.* Revista Brasileira de Segurança Pública, v. 14, n. 2, p. 100–127.

Summerill, W. (2010). *Colonial Institutions, Slavery, Inequality and Development: Evidence from São Paulo, Brazil.*

Transatlantic Slave Database. *Voyages.* http://www.slavevoyages.org/voyage.

Truzzi, B. et al. (2022). *Racial Democracy and Black Victimization in Brazil.* Journal of Contemporary Criminal Justice, v. 38, n. 1, p. 13–33.

Uttermark, M. J. (2019). *What Determines Social Capital? Evidence from Slavery's Legacy in the United States and Brazil.* Social Forces, v. 98, n. 4.

Van Buuren, S. (2018). *Flexible imputation of missing data.* CRC press.

Von Hippel, P. T. (2007). *Regression with missing Ys: An improved strategy for analyzing multiply imputed data.* Sociological Methodology, v. 37, n. 1, p. 83–117.

Zhang, Z. et al. (2016). *Causal inference with a quantitative exposure.* Statistical methods in medical research, v. 25, n. 1, p. 315–335.

Zhou, J. (2015). *Comparison of approaches for handling missingness in covariates for propensity score models.*

Zhu, Y.; Coffman, D. L.; Ghosh, D. (2015). *A boosting algorithm for estimating generalized propensity scores with continuous treatments.* Journal of causal inference, v. 3, n. 1, p. 25–40.

APPENDIX A

Results including binary variables of Brazilian states (fixed effects)

Table 2 - Effect of slavery density on the municipal homicide rate

	Homicides per hundred thousand inhabitants (2000)				
	(1)	(2)	(3)	(4)	(5)
sld	0.117**	0.122**	0.120**	0.119**	0.116**
	(0.052)	(0.052)	(0.052)	(0.052)	(0.053)
ypc	0.001	0.002	0.002	-0.050***	-0.016**
	(0.004)	(0.004)	(0.004)	(0.013)	(0.007)
pd	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
unmp	0.194***	0.208***	0.201***	0.138*	0.146*
	(0.075)	(0.076)	(0.075)	(0.076)	(0.076)
be	-0.149**	-0.163**	-0.159**	-0.053	-0.062
	(0.074)	(0.074)	(0.075)	(0.075)	(0.079)
ym	1.115***	1.054***	1.074***	1.067***	1.103***
	(0.400)	(0.408)	(0.403)	(0.403)	(0.399)
cps	-1.429***	-1.413***	-1.418***	-1.302***	-1.310***
	(0.321)	(0.318)	(0.319)	(0.315)	(0.317)
long	-0.140	-0.160	-0.157	-0.096	-0.113
	(0.208)	(0.210)	(0.209)	(0.210)	(0.209)
lat	-0.397*	-0.386*	-0.388*	-0.447*	-0.451*
	(0.229)	(0.230)	(0.230)	(0.231)	(0.233)
gini		-5.762		-30.697***	
		(5.858)		(8.503)	
theil			-2.214		-8.919***
			(2.676)		(3.243)
state12	-17.316***	-17.340***	-17.348***	-17.592***	-17.846***
	(6.546)	(6.549)	(6.552)	(6.577)	(6.606)
state13	-19.685***	-19.466***	-19.591***	-19.235***	-19.943***
	(5.045)	(5.063)	(5.041)	(5.080)	(5.059)
state14	15.185*	15.301*	15.075*	16.075**	15.099*
	(7.823)	(7.808)	(7.814)	(7.866)	(7.809)
state15	-14.266**	-13.964**	-14.003**	-14.736**	-14.572**
	(5.701)	(5.734)	(5.694)	(5.752)	(5.715)
state16	-1.754	-1.386	-1.411	-2.316	-2.118
	(9.213)	(9.244)	(9.219)	(8.984)	(9.013)

Continued next page

Table 2 (Continued)

state17	-4.348	-3.952	-4.017	-5.387	-5.307
	(5.803)	(5.856)	(5.812)	(5.874)	(5.833)
state21	-19.480***	-19.147***	-19.121***	-21.013***	-20.564***
	(5.976)	(6.014)	(5.973)	(6.019)	(5.980)
state22	-14.504**	-14.168**	-14.129**	-16.241**	-15.707**
	(6.270)	(6.312)	(6.278)	(6.321)	(6.282)
state23	-12.865*	-12.433*	-12.452*	-14.398**	-14.090**
	(6.655)	(6.712)	(6.660)	(6.714)	(6.666)
state24	-10.399	-10.022	-10.080	-12.411*	-12.043*
	(7.098)	(7.143)	(7.099)	(7.146)	(7.106)
state25	-9.780	-9.544	-9.490	-12.191*	-11.437
	(7.234)	(7.261)	(7.233)	(7.268)	(7.243)
state26	4.134	4.582	4.521	2.486	2.751
	(7.043)	(7.109)	(7.054)	(7.112)	(7.064)
state27	-11.241	-10.687	-10.810	-12.412*	-12.491*
	(6.945)	(7.030)	(6.961)	(7.038)	(6.974)
state28	-7.202	-6.881	-6.867	-9.464	-8.942
	(7.136)	(7.187)	(7.154)	(7.180)	(7.161)
state29	-22.083***	-21.684***	-21.704***	-23.814***	-23.472***
	(6.325)	(6.381)	(6.340)	(6.400)	(6.361)
state31	-18.875***	-18.699***	-18.678***	-20.542***	-20.272***
	(6.120)	(6.141)	(6.129)	(6.185)	(6.173)
state32	-3.654	-3.271	-3.281	-5.859	-5.427
	(6.972)	(7.025)	(7.000)	(7.052)	(7.036)
state33	-1.523	-1.388	-1.367	-3.199	-3.017
	(6.752)	(6.751)	(6.750)	(6.789)	(6.789)
state35	-10.231*	-10.320*	-10.274*	-10.931*	-11.175*
	(5.776)	(5.763)	(5.769)	(5.790)	(5.796)
state41	-12.145**	-12.003**	-11.969**	-14.127**	-13.929**
	(5.923)	(5.936)	(5.936)	(5.982)	(5.988)
state42	-16.007**	-16.018**	-15.950**	-17.174***	-17.305***
	(6.319)	(6.314)	(6.319)	(6.348)	(6.361)
state43	-15.442**	-15.528**	-15.421**	-16.649**	-16.823***
	(6.461)	(6.447)	(6.458)	(6.478)	(6.499)
state50	-5.568	-5.427	-5.377	-6.862	-6.536
	(5.726)	(5.740)	(5.739)	(5.796)	(5.787)
state51	7.221	7.318	7.310	6.440	6.662
	(5.433)	(5.443)	(5.436)	(5.496)	(5.484)

Continued next page

Table 2 (Continued)

state52	-8.204	-8.133	-8.092	-9.486*	-9.205*
	(5.382)	(5.389)	(5.382)	(5.423)	(5.411)
state53	9.574*	9.730*	9.739*	8.062	8.715
	(5.752)	(5.765)	(5.753)	(5.756)	(5.722)
*ypc*gini*				0.075***	
				(0.008)	
*ypc*theil*					0.018
					(0.005)
Intercept	6314	9.294	7.043	27.942	13.500
	(15.158)	(15.253)	(15.198)	(15.858)	(15.283)
Observations	2,699	2,699	2,699	2,699	2,699
AIC	173,225	174,980	174,931	177,148	176,494
Pseudo-R^2	0.243	0.243	0.243	0.249	0.247
Pseudo-R^2 Adj.	0.233	0.233	0.233	0.239	0.236

Note: $^{*}p<0{,}1$; $^{**}p<0{,}05$; $^{***}p<0{,}01$

APPENDIX B

Table 5 – Model 1: Balance after Weightening

Control	Type	Pearson Corr. after Adjust		Result
Ypc	Continuous	0.1124	>0.1	Not Balanced
Pd	Continuous	0.0505	<0.1	Balanced
Unmp	Continuous	0.1486	>0.1	Not Balanced
Be	Continuous	0.1647	>0.1	Not Balanced
Ym	Continuous	0.0262	<0.1	Balanced
Cps	Continuous	0.2134	>0.1	Not Balanced
Long	Continuous	0.1094	>0.1	Not Balanced
Lat	Continuous	0.0014	<0.1	Balanced
state_11	Binary	0.0279	<0.1	Balanced
state_12	Binary	0.0181	<0.1	Balanced
state_13	Binary	0.0270	<0.1	Balanced
state_14	Binary	0.0149	<0.1	Balanced
state_15	Binary	0.0000	<0.1	Balanced
state_16	Binary	0.0034	<0.1	Balanced
state_17	Binary	0.0460	<0.1	Balanced
state_21	Binary	0.0875	<0.1	Balanced
state_22	Binary	0.0226	<0.1	Balanced
state_23	Binary	0.0128	<0.1	Balanced
state_24	Binary	0.0264	<0.1	Balanced
state_25	Binary	0.0354	<0.1	Balanced
state_26	Binary	0.0209	<0.1	Balanced
state_27	Binary	0.0134	<0.1	Balanced
state_28	Binary	0.0455	<0.1	Balanced
state_29	Binary	0.0373	<0.1	Balanced
state_31	Binary	0.0072	<0.1	Balanced
state_32	Binary	0.0516	<0.1	Balanced
state_33	Binary	0.2326	>0.1	Not Balanced
state_35	Binary	0.0443	<0.1	Balanced
state_41	Binary	0.0646	<0.1	Balanced
state_42	Binary	0.0531	<0.1	Balanced
state_43	Binary	0.0304	<0.1	Balanced
state_50	Binary	0.0277	<0.1	Balanced
state_51	Binary	0.0263	<0.1	Balanced
state_52	Binary	0.0377	<0.1	Balanced
state_53	Binary	0.0039	<0.1	Balanced
Observations before and after adjustment, respectively			5498	5488

Table 6 – Model 2: Balance after Weightning

Control	Type	Pearson Corr. after Adjust		Result
ypc	Continuous	0.1143	>0.1	Not Balanced
pd	Continuous	0.0513	<0.1	Balanced
unmp	Continuous	0.1484	>0.1	Not Balanced
be	Continuous	0.1662	>0.1	Not Balanced
ym	Continuous	0.0255	<0.1	Balanced
cps	Continuous	0.2157	>0.1	Not Balanced
long	Continuous	0.1088	>0.1	Not Balanced
lat	Continuous	0.0006	<0.1	Balanced
gini	Continuous	0.0804	<0.1	Balanced
state_11	Binary	0.0279	<0.1	Balanced
state_12	Binary	0.0181	<0.1	Balanced
state_13	Binary	0.0269	<0.1	Balanced
state_14	Binary	0.0149	<0.1	Balanced
state_15	Binary	0.0000	<0.1	Balanced
state_16	Binary	0.0034	<0.1	Balanced
state_17	Binary	0.0459	<0.1	Balanced
state_21	Binary	0.0855	<0.1	Balanced
state_22	Binary	0.0225	<0.1	Balanced
state_23	Binary	0.0126	<0.1	Balanced
state_24	Binary	0.0262	<0.1	Balanced
state_25	Binary	0.0353	<0.1	Balanced
state_26	Binary	0.0210	<0.1	Balanced
state_27	Binary	0.0134	<0.1	Balanced
state_28	Binary	0.0453	<0.1	Balanced
state_29	Binary	0.0363	<0.1	Balanced
state_31	Binary	0.0070	<0.1	Balanced
state_32	Binary	0.0517	<0.1	Balanced
state_33	Binary	0.2323	>0.1	Not Balanced
state_35	Binary	0.0452	<0.1	Balanced
state_41	Binary	0.0644	<0.1	Balanced
state_42	Binary	0.0529	<0.1	Balanced
state_43	Binary	0.0304	<0.1	Balanced
state_50	Binary	0.0276	<0.1	Balanced
state_51	Binary	0.0262	<0.1	Balanced
state_52	Binary	0.0376	<0.1	Balanced
state_53	Binary	0.0039	<0.1	Balanced
Observations before and after adjustment, respectively			5498	5489

Table 7 – Model 3: Balance after Weightening

Control	Type	Pearson Corr. after Adjust		Result
ypc	Continuous	0.1138	>0.1	Not Balanced
pd	Continuous	0.0511	<0.1	Balanced
unmp	Continuous	0.1484	>0.1	Not Balanced
be	Continuous	0.1658	>0.1	Not Balanced
ym	Continuous	0.0257	<0.1	Balanced
cps	Continuous	0.2150	>0.1	Not Balanced
long	Continuous	0.1089	>0.1	Not Balanced
lat	Continuous	0.0008	<0.1	Balanced
theil	Continuous	0.0878	<0.1	Balanced
state_11	Binary	0.0279	<0.1	Balanced
state_12	Binary	0.0181	<0.1	Balanced
state_13	Binary	0.0269	<0.1	Balanced
state_14	Binary	0.0149	<0.1	Balanced
state_15	Binary	0.0000	<0.1	Balanced
state_16	Binary	0.0034	<0.1	Balanced
state_17	Binary	0.0459	<0.1	Balanced
state_21	Binary	0.0861	<0.1	Balanced
state_22	Binary	0.0225	<0.1	Balanced
state_23	Binary	0.0127	<0.1	Balanced
state_24	Binary	0.0263	<0.1	Balanced
state_25	Binary	0.0353	<0.1	Balanced
state_26	Binary	0.0210	<0.1	Balanced
state_27	Binary	0.0134	<0.1	Balanced
state_28	Binary	0.0454	<0.1	Balanced
state_29	Binary	0.0366	<0.1	Balanced
state_31	Binary	0.0070	<0.1	Balanced
state_32	Binary	0.0517	<0.1	Balanced
state_33	Binary	0.2324	>0.1	Not Balanced
state_35	Binary	0.0450	<0.1	Balanced
state_41	Binary	0.0644	<0.1	Balanced
state_42	Binary	0.0530	<0.1	Balanced
state_43	Binary	0.0304	<0.1	Balanced
state_50	Binary	0.0276	<0.1	Balanced
state_51	Binary	0.0263	<0.1	Balanced
state_52	Binary	0.0376	<0.1	Balanced
state_53	Binary	0.0039	<0.1	Balanced
Observations before and after adjustment, respectively			5498	5489

Table 8 – Model 4: Balance after Weightening

Control	Type	Pearson Corr. after Adjust		Result
ypc	Continuous	0.1114	>0.1	Not Balanced
pd	Continuous	0.0487	<0.1	Balanced
unmp	Continuous	0.1476	>0.1	Not Balanced
be	Continuous	0.1637	>0.1	Not Balanced
ym	Continuous	0.0251	<0.1	Balanced
cps	Continuous	0.2158	>0.1	Not Balanced
long	Continuous	0.1107	>0.1	Not Balanced
lat	Continuous	0.0001	<0.1	Balanced
gini	Continuous	0.0802	<0.1	Balanced
state_11	Binary	0.0281	<0.1	Balanced
state_12	Binary	0.0182	<0.1	Balanced
state_13	Binary	0.0272	<0.1	Balanced
state_14	Binary	0.0150	<0.1	Balanced
state_15	Binary	0.0003	<0.1	Balanced
state_16	Binary	0.0035	<0.1	Balanced
state_17	Binary	0.0463	<0.1	Balanced
state_21	Binary	0.0876	<0.1	Balanced
state_22	Binary	0.0232	<0.1	Balanced
state_23	Binary	0.0132	<0.1	Balanced
state_24	Binary	0.0267	<0.1	Balanced
state_25	Binary	0.0359	<0.1	Balanced
state_26	Binary	0.0204	<0.1	Balanced
state_27	Binary	0.0131	<0.1	Balanced
state_28	Binary	0.0452	<0.1	Balanced
state_29	Binary	0.0370	<0.1	Balanced
state_31	Binary	0.0073	<0.1	Balanced
state_32	Binary	0.0493	<0.1	Balanced
state_33	Binary	0.2461	>0.1	Not Balanced
state_35	Binary	0.0438	<0.1	Balanced
state_41	Binary	0.0651	<0.1	Balanced
state_42	Binary	0.0535	<0.1	Balanced
state_43	Binary	0.0309	<0.1	Balanced
state_50	Binary	0.0279	<0.1	Balanced
state_51	Binary	0.0266	<0.1	Balanced
state_52	Binary	0.0381	<0.1	Balanced
state_53	Binary	0.0039	<0.1	Balanced
Observations before and after adjustment, respectively			5498	5487

Table 9 – Model 5: Balance after Weightening

Variável	Tipo	C. Pearson após Ajuste		Resultado
ypc	Continuous	0.1109	>0.1	Not Balanced
pd	Continuous	0.0494	<0.1	Balanced
unmp	Continuous	0.1472	>0.1	Not Balanced
be	Continuous	0.1629	>0.1	Not Balanced
ym	Continuous	0.0245	<0.1	Balanced
cps	Continuous	0.2144	>0.1	Not Balanced
long	Continuous	0.1099	>0.1	Not Balanced
lat	Continuous	0.0001	<0.1	Balanced
theil	Continuous	0.0858	<0.1	Balanced
state_11	Binary	0.0281	<0.1	Balanced
state_12	Binary	0.0182	<0.1	Balanced
state_13	Binary	0.0272	<0.1	Balanced
state_14	Binary	0.0150	<0.1	Balanced
state_15	Binary	0.0002	<0.1	Balanced
state_16	Binary	0.0035	<0.1	Balanced
state_17	Binary	0.0463	<0.1	Balanced
state_21	Binary	0.0878	<0.1	Balanced
state_22	Binary	0.0232	<0.1	Balanced
state_23	Binary	0.0131	<0.1	Balanced
state_24	Binary	0.0266	<0.1	Balanced
state_25	Binary	0.0358	<0.1	Balanced
state_26	Binary	0.0206	<0.1	Balanced
state_27	Binary	0.0132	<0.1	Balanced
state_28	Binary	0.0453	<0.1	Balanced
state_29	Binary	0.0372	<0.1	Balanced
state_31	Binary	0.0071	<0.1	Balanced
state_32	Binary	0.0436	<0.1	Balanced
state_33	Binary	0.2455	>0.1	Not Balanced
state_35	Binary	0.0448	<0.1	Balanced
state_41	Binary	0.0650	<0.1	Balanced
state_42	Binary	0.0534	<0.1	Balanced
state_43	Binary	0.0307	<0.1	Balanced
state_50	Binary	0.0279	<0.1	Balanced
state_51	Binary	0.0265	<0.1	Balanced
state_52	Binary	0.0380	<0.1	Balanced
state_53	Binary	0.0039	<0.1	Balanced
Observations before and after adjustment, respectively			5498	5486