Race, Gender and Poverty: Evidence from Brazilian Data

Yeutseyeva, Sasha and Deguilhem, Thibaud

Smith College, Université de Paris, LADYSS UMR 7533

31 August 2022

Race, Gender and Poverty: Evidence from Brazilian Data

Sasha Yeutseyeva* Thibaud Deguilhem†

August 30, 2022

Abstract

Race and gender are commonly considered as two of the most important structural factors associated with unequal socioeconomic systems. Previous research has found that these factors are significant for explaining the income inequality in Latin America and particularly in Brazil. This study aims to address whether both determinants predict an individual's chances of being in poverty in Brazil, using national dataset and articulating different econometric strategies. Overall, being a woman had a small positive impact on an individual's predicted chance of poverty and only in a probability linear specification. We think that this result does not align well with previous literature because of the selection bias affecting women labor market participation. However, evidence of strong and robust racial differenciation in Brazil was present. Discussing the representativeness of the sample, this study highlights the importance of data quality as well as the relevance of using various statistical methods.

Keywords: Brazil, poverty, race, gender, inequality

*Smith College, syeutseyeva@smith.edu

†Université Paris Cité – LADYSS UMR 7533, thibaud.deguilhem@u-paris.fr.

Introduction

In the late 1960s, Brazil was known as a “miracle economy” seeing double digit growth rates and having “relative political stability” because of military rule. This “allowed the country to expand its industries,” prompting people to relocate “from the countryside to the cities to find work. By the early 1980s, Brazil had become one of the leading industrial nations, boasting the tenth-largest gross national product in the world.” However, this was coupled with double-digit annual inflation in the 1970s, which “turned to triple digits by the 1980s (and would reach the 50 percent level on a monthly rate by 1994).” Because of the “high inflation of the 1980s”, Brazil enacted “a policy of increasing interest rates, which slowed economic investment and expansion.”

In the late 70s and early 80s, due to the oil crisis, Brazil and other countries in Latin America were impacted, seeing skyrocketing foreign debt and small growth in annual GDP.[1] As a result, the 80s are considered a lost decade for Latin America’s economy. Many of the countries “experienced a loss of wealth” and “many social indicators registered a sharp deterioration in the quality of life within the region throughout the decade.” As a result, lots of Latin American countries spent the 90s attempting to recover their former standard of living (Baumann, 2002).

For Brazil, the restructuring of its economic system began in 1991 with the Mercosul Agreement that established free trade between Brazil, Argentina, Paraguay, and Uruguay. Then, in 1994, Brazil established the Real Plan, “which sought to end years of high inflation by introducing a new currency (the real) pegged to the U.S. dollar” (Ferderer, 1997). Although the Real Plan was effective at creating price stability, “the Mexican crisis of 1994/5 led to a sudden stop in the flow of international finance to Latin American countries. Brazil, already suffering from a fastwidening current account deficit, endured a balance of payments crisis that forced the Central Bank to impose a steep rise in interest rates to try to stop capital flight.” Multiple crises like this occurred “until the balance of payments crisis of late 1998/early 1999, which led to a change in the macroeconomic policy regime.” These policies remained “fundamentally in force” for many years (Cardim de Carvalho & Pires de Souza, 2011).

In the 2000s, Brazil experienced limited crises “until the last quarter of 2008, when the international crisis” arrived there. Fortunately, the “impact of the crisis was strong but relatively short-lived,” and “after a two-quarter recession, the Brazilian economy resumed growth.” However, the rates of growth were low (Cardim de Carvalho & Pires de Souza, 2011). In 2015, the “Brazilian economy entered a deep recession … and since then has shown a sluggish recovery” (Mantoan et al., 2021). The government has implemented economic policy that is focused “heavily on the fiscal agenda” and “seeks to resolve problems through structural changes in the State apparatus, generation of new income, and reduced public spending.” In 2017, “after two consecutive years of declines in GDP… the prolonged recession came to an end”(ECLAC, 2018).[2] Brazil’s weak growth continued into 2019, when the GDP

[1] Ito, T. (1999, January). Brazil: A History of Political and Economic Turmoil. *Washington Post*. https://www.washingtonpost.com/wp-srv/inatl/longterm/brazil/overview.htm

[2] http://hdl.handle.net/11362/42652.

rose by just 1.1% (ECLAC, 2021).[3]

Brazil has been long known as a country with one of the highest inequality rates in the world. In 2021, Brazil's Gini coefficient of income inequality was 53.4 (OECD, 2021). Historically, there are many reasons why this is the case. A previous study has shown that slavery has played a negative effect on economic development for countries that have previously had slavery. Furthermore, it was found that there is a "strong significant negative relationship between past slave use and current income" when examining the US and other countries in the Americas (Nunn, 2007). Brazil, with three and a half centuries of slavery part of its history, is no outlier to this trend (Gradin, 2007). Although Brazil was able to decrease the rates of inequality from the early 2000s up until 2015, poverty rates increased again as a result of economic downturn (Rocha, 2019). This is inline with it's general economic successes from 2000 to 2015. There are numerous causes of the fall of inequality, "such as increasing education levels, falling education premiums, the diffusion of social programmes such as conditional cash transfers (CCTs), and the expansion of non-contributory social security benefits. However, most of the fall in inequality was driven by changes in labour earnings" (Neri, 2019).

One of the inequalities present in Brazil is race inequality. In 2006, the percentage of people below the poverty line was 43.1% for blacks and browns while only 21.6% for whites (Lima & Prates, 2018). Although the racial divide of the income ratio has decreased between 1987-2012, it has done so slowly and the racial divide is still considerable. According to a study, if the trend of poverty reduction from 2001-2012 continues, Brazil would need another 41 years to equalize the racial poverty divide. Although Brazil's increase in public cash transfers with programs like Bolsa Familia and efforts to equalize educational attainment have been effective, this shows their strategies need to be adjusted accordingly to speed up the reduction of racial inequality (Pereira, 2016).

It is worth nothing that there has been empirical research conducted that has shown that racial discrimination is a significant factor that explains income inequality in Brazil. While places of residence and education attainment are factors, whites receive higher incomes than blacks due to discrimination as well. A 2020 study has also found that race directly reduces the income of blacks by 7.4% and indirectly affects their education. Therefore, when accounting for direct and indirect effects, it was found that the income of blacks was reduced by 16.8% (Salata, 2020). The same overall pattern of higher poverty rates for Afro-descendants holds true for other countries in Latin America. In 2014, the poverty rates in Ecuador were 42% for Afro-descendants and 27% for non-Afro-descendants. For Peru, the respective values were 21 and 14, and for Uruguay it was 11 and 4 (Lais Wendel, 2016).[4]

Another one of the major inequalities in Brazil is gender inequality. One of the ways that gender inequality throughout the world is measured is the United Nation Development Programme's Gender Development Index (GDI), which accounts for health, education, and income dimensions of human development, only using estimates of the female and male

[3] https://repositorio.cepal.org/handle/11362/44675; https://repositorio.cepal.org/bitstream/handle/11362/44675/EEI2019_Brazil_en.pdf

[4] Abramo Lais Wendel. The Social Inequality Matrix in Latin America. United Nations ECLAC, 2016. https://repositorio.cepal.org/bitstream/handle/11362/40710/1/S1600945_en.pdf

income per capita relative to the gross national/domestic income, as well as labor force participation and average earnings. Indicators like the GDI can change depending on the weight associated with dimensions like health and education, giving the women in countries like Brazil higher scores than men (Bradshaw, 2018). As a result, official metrics like GDI can sometimes be misleading.

Brazil has made efforts to reduce poverty among women through Bolsa Familia, which directly transfers cash to women, with an intent to "compensate mothers for their traditional domestic and care work role, to ensure that programme co-responsibilities are met and in recognition of the fact that they are most likely to ensure that increased household income benefits children." Assigning women as the beneficiaries of cash also promotes the control of women over household resources (de Brauw et al., 2014). Along with the short term benefits for women, the intent of Bolsa Familia is to stop intergenerational poverty by emphasizing the importance of education and health. Bolsa Familia is "seen as a more cost-effective means of reaching the poor directly through efficient targeting while minimizing resource wastage" (Hall, 2008). However, the effectiveness of Bolsa Familia is debated. Although it "appears to have been effective in providing short-term relief to some of the most deprived groups in Brazil", "there is a risk that, due to its popularity among both the poor and Brazil's politicians, Bolsa Familia could greatly increase patronage in the distribution of economic and social benefits and induce a strong dependence on government handouts" (Hall, 2008). Furthermore, one study had found that "no improvement in the nutritional status of the families entitled to the program was observed, nor the interruption of the intergenerational cycle of poverty was ensured" (Neves et al., 2020). This goes to show that conditional cash transfer programmes like Bolsa Familia may not have the desired impact on poverty eradication.

The intersection of race and gender has also been shown to have a negative effect on Brazilians. This is important to note because despite having "a national ideology of 'racial democracy,' members of disadvantaged groups [in Brazil], particularly citizens with intersectional identities, may often fail to perceive their experiences as linked to those of others with similar characteristics." A study had found that the intersection of race and gender does in fact matter, and that "dark-skinned women are substantially more likely to perceive themselves as targets of gender discrimination than their light-skinned peers." It is also found that "Afro-Brazilian women are disadvantaged in a great variety of outcomes and can be targets of police violence" (Layton & Smith, 2017). Although this study focused on inequalities different from economic inequality, it still highlights the presence of a clear interaction between the variables of race and gender and it's impact on Brazilians. More recently, the decreasing amount of aid provided by the Brazilian government due to the COVID-19 pandemic had increased the total poverty of the population and had widened the gender and racial gaps, "mainly due to the greater economic vulnerability of black women". A study states "Before the pandemic, poverty reached 33 percent of black women, 32 percent of black men, and 15 percent of white women and white men in Brazil. The scenario for 2021 leads to poverty rates of 38 percent, 36 percent, 19 percent, and 19 percent, respectively" (Nassif-Pires, et al., 2021). With findings like these, it is important to inspect the intersection of race and gender and it's impact on an individual's chances of being in poverty.

Although studies have shown that gender and race inequalities in Brazil exist, there are also other factors that could contribute to the poverty of individuals, such as education, occupation, and location. Taking Paraiba, Brazil as a case study, one paper has found that "all levels of education from primary to tertiary are strongly significant and negatively associated with the probability of being poor" (Verner, 2004). However, a different study has found that for improved education to create a significant reduction in poverty it would take decades and would require a "significant scale-up of tertiary education" (Medeiros & al. 2020). There is evidence that programs like Bolsa Escola, later renamed Bolsa Familia, were able to increase childhood education and reduce income inequality and poverty. Furthermore, it has been shown that for youth aged 15-29, higher educational levels came with higher hourly earnings. This effect was particularly strong for women, whose earnings were 51% higher for those who completed high school compared to those who didn't (Garcias & al., 2021).

As for occupation, a "prevalence of poverty among agricultural workers" has been found (Ferreira et al., 2003). Research into the inequality in Britain has shown that the wage gap by race is small within occupations, but is higher across. This indicates that non-white minorities in Britain "tend to concentrate in low-paying occupations." For the British, it may be that the occupation an individual pursues is more significant to inequality than wage discrimination (Brynin & Longhi, 2015). Inequality is also present in urban and rural areas of Brazil. In rural areas, a large percentage of the land is owned by a small number of people, leaving millions of "small landowners, landless workers and rural workers living in precarious conditions." In urban areas, a significant percentage of living conditions are inadequate, where people only have "precarious access to housing, infrastructure, public equipment and public health, education, culture, information, recreation, sports and transportation services, among others" (Beghin, 2008).

The main question that will be answered in this paper is the intersection of race and gender on poverty in a Brazilian context. This paper further aims to determine the major factors responsible for the discrepancy in poverty levels between Brazilian racial and gender groups. Based on the findings of previous literature described above, my hypothesis is that being a woman has a positive impact on the predicted odds of an individual being in poverty, as well as the same for being non-white. I suspect that the intersection of being a woman and non-white will have an even greater positive impact on the predicted odds of an individual being in poverty.

The main contributions of the paper are twofold. First, we argue this empirical paper is one of the few papers to study quantitatively the gender and racial inequality in Brazil with a large dataset. Second, unlike most herding literature, the paper employs three kind of models (linear, logistic and matching) to show robust results. Looking for a variety of methods is necessary to overcome some of the empirical limitations.

Data and Empirical Strategy

Data

The Brazilian dataset used in this paper is available on Kaggle with more than 20,000 observations.[5] It contains defining characteristics such as whether the individual is a woman, their age, and whether they are non-white. It also has variables for the level of education, the work, and work permit status of the individual. This dataset includes information about whether the person lives in a metropolitan and/or urban area. Lastly, the poverty indicator is a binary variable that shows whether the individual received less than 457 Brazilian Reals in 2020, equivalent to roughly 88 USD as of December 31 2020.[6] Due to the postponement of Brazil's census from 2020 to 2022, official census data is unavailable for comparison with this data,[7] so other sources will be used accordingly.

The UN Department of Economic and Social Affairs estimated that in 2020, there would be 96.6 males for every 100 females, an equivalent to 50.9% of the population being female.[8] This is evidently very different than the values observed in this data set, where females make up only 39.4% of the sample (Table 1). Due to the lack of information about how this data was collected, it is impossible to say exactly why this isn't representative of the population. A possible selection bias could include the survey being administrated in a way that unemployed women who stay at home do not get selected, or being administered in spaces with low female presence. As a result, possible implications of this sample makeup include empirical effects on the values of certain coefficients and inconsistent results. The same selection bias that led to an under representation of women might also create a bias for representing woman with certain characteristics, such as a larger representation of highly educated or non-white women. In turn, this would change the estimated impact that being a woman would have on an individuals chance of being in poverty.

In the sample, 55% of the population is non-white (Table 1). Due to the absence of a recent census conducted in Brazil, the latest reliable figures about the race make-up of the Brazillian population come from the 2010 census. In 2010 it was estimated that 48% of the population was white.[9] Although there is likely an over representation of non-white individuals in this sample, it is not possible to conclude due to the lack of data. However, an over representation of non-white individuals could signal a selection bias, which would affect the estimated poverty risk for this sample.

Regarding the distribution in Table 1 and the histogram in the appendix, age appears normally distributed. However, the data is a not a good representation of Brazil's age

[5] "This dataset contains data from 2020 on more than 20,000 Brazilian individuals and their characteristic[s]"https://www.kaggle.com/patrickgomes/determinants-of-poverty-in-brazil

[6] "Exchange-Rates.org." Brazilian Reals (BRL) to US Dollars (USD) Rates for 12/31/2020 - Exchange Rates, https://www.exchange-rates.org/Rate/BRL/USD/12-31-2020.

[7] Conteudo, Estadao. "IBGE Adia Início Do Censo 2022 Para 1º De Agosto." Exame, 25 Jan. 2022, https://exame.com/brasil/ibge-adia-inicio-do-censo-2022-para-1o-de-agosto/.

[8] "World Population Prospects - Population Division." United Nations, United Nations, https://population.un.org/wpp/DataQuery/.

[9] "Brazil Destination Guide." Diversity Abroad, https://www.diversityabroad.com/articles/travel-guide/brazil.

distribution. In Brazil, 15 to 24 year olds make up 20.36% of the population that is 15 and over, however in this sample the same group only makes up 10.66%. Furthermore, there is an over representation of people aged 25-54, with 73.90% of the sample belonging to that age group, meanwhile in Brazil it's only 55.56%. The 55-64 age group makes up 12.26% of the sample and 12.40% of the Brazilian population which makes it the most accurately represented age group. Lastly, the share of the sample over 64 years old is only 3.18%, whereas in Brazil it is 11.68%.[10].

An Age vs Poverty distribution graph of the age variable based on whether the individual is in poverty or not is created in the appendix. As can be seen, the distribution of age for individuals in poverty is not spread towards the young and the old. Instead, the cases of poverty are concentrated towards the middle of the age range, around 40 years. To account for this nonlinear relationship, the age variable is squared. This may be a result of a selection bias that excludes unemployed individuals. Although younger and older people might face more poverty due to age discrimination, this would not be evident in the data because unemployed individuals aren't represented. A complementing graph, Age Squared vs Poverty, is included in the appendix.

The most common levels of education for this population is incomplete elementary, complete audio, or complete superior. For the purposes of this analysis, education has been transformed into three dummy variables: some or complete elementary, some or complete audio, and some or complete superior. As can be seen in the bar graph in the appendix titled Education, the percentage of the sample belonging to each category is strikingly similar, with the smallest category being people attaining some or completing superior education. Compared to the statistics published by the World Bank in 2018 that state that the rate of people aged 25 and over in Brazil that have completed primary education is over 80 percent,[11] this sample produces a smaller percentage, at only 74.4 percent having completed elementary. Overall, the sample is fairly evenly distributed among the three categories of educational attainment. With an underrepresentation of people who have completed elementary education or above, this sample may overrepresent people in poverty, due to previous findings that higher education attainment correlates to the reduction of poverty (Verner, 2004).

The work variable has been refactored into the three sectors of industry based on defined classifications.[12] The breakdown is as follows:

Primary: agriculture, livestock, forestry, fisheries and aquaculture

Secondary: general industry; construction; ill-defined activities

Tertiary: transport, storage and mail; trade, repair of motor vehicles and motorcycles; accommodation and food; other services; home services; information, communication and

[10] "Field Listing-Age Structure." Central Intelligence Agency, Central Intelligence Agency, https://www.cia.gov/the-world-factbook/field/age-structure.

[11] "Educational Attainment, at Least Completed Primary, Population 25+ Years, Total (%) (Cumulative) - Brazil." The World Bank, Sept. 2021, https://data.worldbank.org/indicator/SE.PRM.CUAT.ZS?locations=BR.

[12] Pettinger, Tejvan. "Sectors of the Economy." Economics Help, 19 Dec. 2019, https://www.economicshelp.org/blog/12436/concepts/sectors-economy/.

financial, real estate, professional and administrative; public administration, defense and social security; education, human health and social services

Something important to note is that there is no representation of unemployed individuals in this sample. Although this makes sense because the determinant for the poverty variable is whether the individual had earned 457 Brazilian Reals in 2020, which would make unemployed people having an income of 0 automatically be considered in poverty, this leads to a selection bias. In 2020, the unemployment rates for different populations of Brazil was unequal. That year, the "unemployment rate was 12.8% for men, 16.8% for women and 19.8% for black women".[13] With this, there may be an under representation of women, and especially non-white women.

In the sample, 37% of the individuals live in a metropolitan area, and 79% live in an urban area (Table 1). In 2020, the World Bank estimated that only 13% of Brazilians live in a rural area,[14] which means rural people are over represented in this sample. A possible selection bias could be that the data collectors were interested in the causes of poverty, therefore surveying people in poverty-prone areas, such as rural ones.

Checking if urban and metropolitan_area are correlated with a chi-squared test, it is found that the two variables are correlated. As a result, only urban will be used in the model.

The work_permit variable has 3 possible values, which are: does not have a work permit; has a work permit; other situations, employer, civil servant.

According to the World Bank, in response to the global COVID-19 pandemic the Brazilian government "put forward a large, timely, targeted and time bound fiscal package focused on health spending, ... social assistance ... to 66 million individuals and the expansion of the Bolsa Familia Conditional Cash Transfer." This package "served as a swift and generous temporary relief, which helped poverty go down from 19.6 percent in 2019 to 12.8 percent in 2020 (poverty rate is rate based on the USD 5.5/day (PPP) line)".[15] This is significantly different than the poverty rate of 22.5% present in the sample. Furthermore, the sample uses a rate of 457 Brazilian Reals per year as the poverty determinant, which is significantly lower than the one of the World Bank, which equates to 7791 Brazilian Reals per year. Selection biases might involve the data collector being more interested in the causes of poverty, not in the prevalence of poverty, therefore selecting people in areas with high poverty rates and purposely seeking out people in poverty.

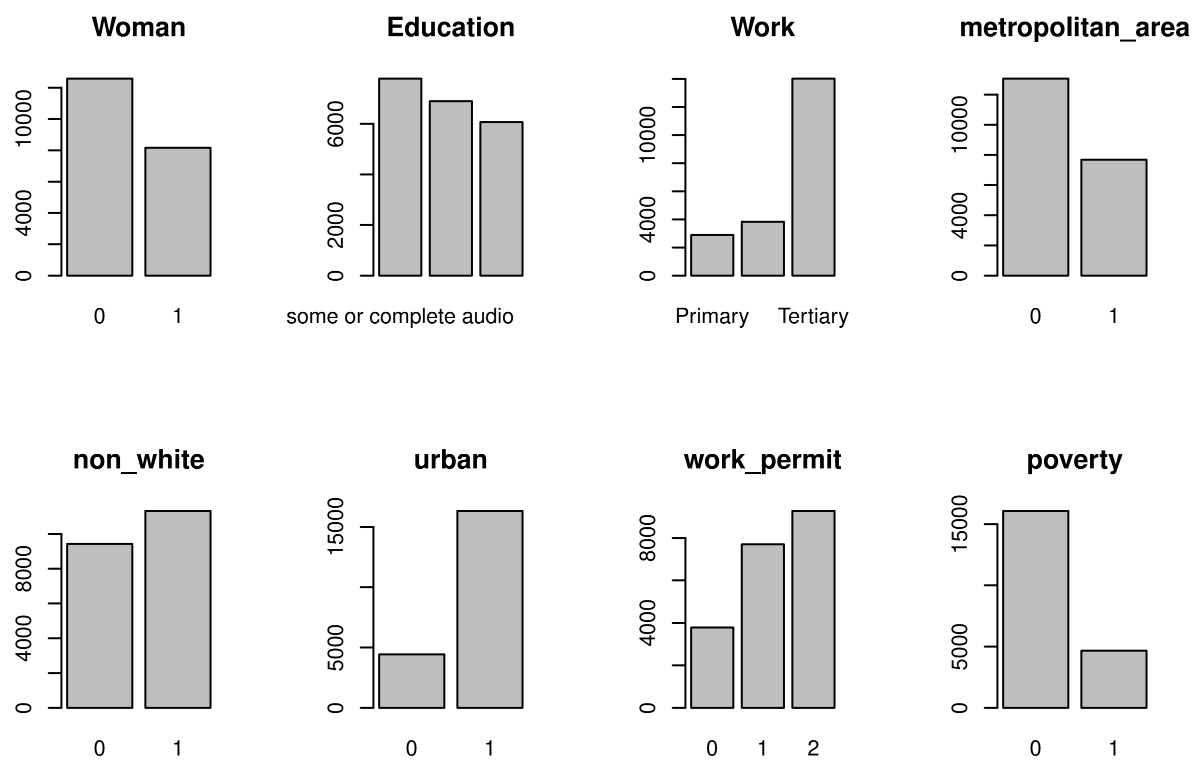
[13] Oliveira, Caroline. "Unemployed, Precarious and Outsourced: the State of Brazilian Women amid the Pandemic." Brasil De Fato, 8 Mar. 2021, https://www.brasildefato.com.br/2021/03/08/unemployed-precarious-and-outsourced-the-state-of-brazilian-women-amid-the-pandemic.

[14] "Rural population (% of total population) - Brazil." The World Bank, 2018, https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?locations=BR.

[15] "Brazil Overview." World Bank, https://www.worldbank.org/en/country/brazil/overview#1.

Table 1: Descriptive statistics

variable	n	mean	sd	min	max
age	20752	40.86	12.60	14	91
age_squared	20752	1828.40	1087.87	196	8281
metropolitan_area	20752	0.37	0.48	0	1
no_work_permit	20752	0.18	0.39	0	1
non_white	20751	0.55	0.50	0	1
other_situation_work_permit	20752	0.45	0.50	0	1
poverty	20752	0.22	0.42	0	1
primary_work_sector	20752	0.14	0.35	0	1
secondary_work_sector	20752	0.18	0.39	0	1
some_or_complete_audio_education	20752	0.38	0.48	0	1
some_or_complete_elementary_education	20752	0.33	0.47	0	1
some_or_complete_tertiary_education	20752	0.29	0.46	0	1
tertiary_work_sector	20752	0.68	0.47	0	1
urban	20752	0.79	0.41	0	1
woman	20752	0.39	0.49	0	1
work_permit	20752	0.37	0.48	0	1



Overall, this sample is not an accurate representation of the population of Brazil. It appears that the "factors responsible for selection [were] not controlled for explicitly" and could be "correlated with the dependent variable." This results in the factors becoming "part of the error term" and a bias in the coefficient estimates. As a result, these findings are likely not applicable to the entire Brazilian population (Hughes, 1997). It is possible that due to this sample not representing unemployed individuals, other variables became misrepresented. For example, a possible reason for women having higher unemployment rates is staying at home to care for children. Unemployment may also be higher for non-white groups due to discrimination or other factors. Lastly, age could be misrepresented because of smaller employment rates in younger people and older people.

Empirical Strategy

The methods of analysis used are a probability linear model (PLM), a logistic regression model, and propensity score matching. The key coefficients of interest that will be explored are whether the individual is a woman and whether they are non white.

The linear regression model is first done with poverty as the dependent variable and woman and non-white as independent variables. They are added one after the other, then an interaction is introduced. In this model the ordinary least squares method is used to determine the line of best fit. In order to estimate the linear model, poverty is treated as a numeric variable and define the dependent variable (Y). *Woman* and *Non-white* variables (X) are our explanatory variables, introduced individualy and interactivly in the models to measure clearly individual and multiplicated effect. The linear regression is then made with the gradual addition of all the other explanatory variables as controls (Z).

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_i$$

In our case:

$$Poverty_i = \beta_0 + \beta_1 Non - white_i + \beta_2 Woman_i + \beta_3 (Woman * Non - white)_i + \beta_4 Z_i + \epsilon_i$$

The logistic regression is done similarly to the linear regression, with poverty as the dependent variable and woman and non-white as independent variables. To this, control variables are added one by one. The results are transformed into probabilities with relevant t-tests.

The third method of analysis included is propensity score matching (PSM). The main idea behind PSM is to find observations that are similar based on specified characteristics that are then automatically assigned to be in the treatment or control group. This method was particularly useful during this investigation because there are a lot of observations and a large number of covariates to take into account. Using PSM allowed us to mock a randomized control trial with this observational data without losing large numbers of observations, leading to more trusted results. With this, we are able to determine the predicted probability of being in poverty for individuals that have the same characteristics. In this analysis, a one-to-one nearest method for finding neighbors was used, with the points assigned being chosen in a random order.The results were then interpreted with relevant t-tests.

The methods used in this paper have been used by other publications in the past to analyze

Table 2: Probability matrix

	man	woman
white	0.160	0.110
non_white	0.321	0.254

common contributors to poverty of individuals. Saefuddin et al. (2011) applied ordinary linear regression to Indonesian poverty data. Similarly, Farah (2015) has used logistic regression to analyze characteristics of poverty on Bangladeshi. In 2013, there was also a study using propensity score matching to measure the impact of "Bolsa Familia on Women's decision making power" (de Brauw et al., 2014). Overall, we are using the common methods found in previous literature to test the hypothesis.

Results

Preliminary analysis One of the first things to examine is whether there are gaps in the poverty levels when accounting for whether an individual is non-white and their gender.

The first thing to note in Table 2 is that the poverty rates for women are actually lower than that of men. This goes against the literature discussed in the previous section. Although a counter-intuitive result, the data itself might show why there is a discrepancy. As discussed previously, women are considerably under-represented in this sample, which demonstrates the presence of a selection bias.

As seen in Table 3, the predicted changes in the probability of an individual being affected by poverty whether they are a woman or non-white changes as more variables are controlled for. Controlling for education, the value associated with an individual being a woman becomes statistically insignificant and increases to 0.003. Once the work variable is introduced, the coefficient for being a woman becomes statistically significant and has the value 0.02, indicating an individual's predicted probability of being in poverty increases by 2 percent if they are a woman, when accounting for whether they are non_white, their level of education, and work sector. The same value occurs for when whether the individual lives in an urban area is accounted for. When work-permit and age is introduced, the estimated effect becomes insignificant with a value of 0.01.

An interesting observation is that the predicted effect on being non_white changes very little as more variables are introduced. With the introduction of education, the work sector, and urban variables, the non-white coefficient remains statistically significant with a value of 0.13 after all introductions. After work-permit is introduced, the value drops to 0.12 and remains significant. Finally, with the introduction of age the value becomes 0.11 and is still significant. All throughout the process of introducing and controlling for variables, the interaction term between gender and race remains insignificant.

These results differ from the findings of previous literature. Unlike previous literature, which had directly found that gender has a significant impact on an individual's economic status (Bradshaw et. al, 2018; de Brauw et al., 2014), this analysis concludes that being a woman is an insignificant variable when assessing poverty rates. As discussed above, the interaction of race and gender also plays a significant part in economic status (Layton & Smith, 2017). The data analyzed here, however, shows no trace of this interaction making an impact on the chances of an individual being in poverty, perhaps because of the insignificance of the gender variable on its own. One thing that adheres to the findings of previous literature is the significance of an individual's race on their odds of being in poverty.

When a logistic regression model is used to estimate the predicted probability for individuals to be in poverty while accounting for variables other than their gender and race, the same values are not obtained. This can be viewed in Table 4. As with the linear regression model, the interaction term between gender and race remains insignificant.

When accounting for education, it's predicted that being non-white will increase a man's probability of being in poverty by 17.5%, whereas the value is 19.4% for women. Meanwhile, being a woman decreases the predicted probability that a white individual will be in poverty by 1.7%. For non-white individuals, that value is a 0.2% increase.

When work is accounted for with education, the predicted value that being non-white will affect a man's probability of being in poverty drops to 17.1%, while value rises to 20.0% for women. Meanwhile, being a woman increases the predicted probability that a white individual will be in poverty by 1.3%. For non-white individuals, that value is a 4.3% increase.

After also accounting for whether the individual lives in an urban area, the predicted value that being non-white will affect a man's probability of being in poverty increases to 19.6%, while the value rises to 22.2% for women. Meanwhile, being a woman increases the predicted probability that a white individual will be in poverty by 1.5%. For non-white individuals, that value is a 4.0% increase.

Further including the status of whether the individual has a work permit, it's found that the predicted value that being non-white will affect a man's probability of being in poverty is an increase of 16.5%, while the value is 18.3% for women. Meanwhile, being a woman increases the predicted probability that a white individual will be in poverty by 0.6%. For non-white individuals, that value is a 2.4% increase.

Accounting for all the available variables, it's predicted that being non-white will increase a man's probability of being in poverty by 18.0%, whereas the value is 20.3% for women. Meanwhile, being a woman increases the predicted probability that a white individual will be in poverty by 0.8%. For non-white individuals, that value is a 3.0% increase.

As with linear regression, some of the results from this data are inconsistent with previous literature. Unlike in other studies, being a woman is not a significant variable when assessing poverty based on this analysis. Furthermore, the intersection of race and gender remains insignificant as well, which is unexpected given previous literature indicates this does in fact have a significant impact on poverty. What remains consistent with previous literature is the significant impact of race on poverty.

For the third part of the analysis (Tables 5 and 7), a propensity score matching technique is used to determine whether being non-white affects an individual's chance of being in poverty. Performing the propensity score matching technique using woman as the treatment and control variable, it was found that upon addition of control variables, the predicted chance of poverty decreased more and more, leading to insignificant results. Because of this, the PSM results that will be discussed in depth will instead be the analysis done when using the value of the non_white variable. The value of the non_white variable was used to assign individuals into the treatment and control groups. Then, the chance of poverty of the observations in the two groups are viewed with a t-test.

In the appendix, there is a distribution plot of the matched and unmatched treatment and control groups after they have been matched using the one-to-one nearest method. The points were taken in a random order, instead of highest to lowest propensity scores, to prevent all the matched treated units from concentrating in the upper half. As can be seen, the distribution of the matched control and treated units is very similar, so the results of PSM is applicable to the data at hand.

Overall, propensity score matching in produced results that showed that for this sample, being a woman doesn't have an effect on an individual's risk of being in poverty. On the other hand, being non-white is found to be a statistically significant factor with non-white individuals having a 15.8% higher chance of being in poverty. This can be seen in Table 5 of the appendix.

Estimations

The three analysis methods used yielded similar but slightly different results.

With linear regression (Table 3), it was found that being a woman decreases a person's chance of poverty when not accounting for other variables. After the addition of whether the individual is non-white, their education level, their work sector, whether they live in an urban area, whether they have a work permit, and their age, it was predicted that being a woman only slightly increased an individual's chance of being in poverty. An interesting observation is that the interaction term between race and gender is statistically insignificant. On the other hand, accounting for only whether an individual is non-white, it was found that non-white individuals have a 16% greater chance of being in poverty than whites. After accounting for all other variables, as with the woman variable, it was found that the effect of the non-white variable stayed relatively the same and in the end it was predicted that being non-white increases a person's chance of poverty by 11 percent.

Moving on to logistic regression (Table 4), it was found that being a woman increases a person's chance by 6.5% when not accounting for other factors, which is more than the estimate of the linear regression. After adding variables, being a woman only slightly increases an individuals estimated chance of being in poverty by 0.8% for white individuals and 3.0% for non-whites. This is very similar to the result of the linear regression analysis. Likewise, the interaction term between woman and non-white was also found statistically insignificant. Accounting for all the available variables, logistic regression predicted that being non-white will increase a man's probability of being in poverty by 18.0%, and for a woman the value was found to be 20.3%. These values are higher than those found by linear regression.

Finally, propensity score matching yielded no new results for the effect of the woman variable

Table 3: Probability Linear Models

	Dependent variable:				
	Chance of Poverty				
	(1)	(2)	(3)	(4)	(5)
Woman	0.003	0.02***	0.02**	0.02**	0.01*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Non-white	0.13***	0.13***	0.13***	0.12***	0.11***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Elementary education	0.16***	0.13***	0.11***	0.10***	0.12***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Superior Education	−0.14***	−0.13***	−0.12***	−0.12***	−0.12***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Secondary Sector		−0.15***	−0.06***	−0.05***	−0.05***
		(0.01)	(0.01)	(0.01)	(0.01)
Tertiary Sector		−0.17***	−0.08***	−0.08***	−0.08***
		(0.01)	(0.01)	(0.01)	(0.01)
Urban			−0.14***	−0.13***	−0.12***
			(0.01)	(0.01)	(0.01)
Work perm 1				−0.20***	−0.20***
				(0.01)	(0.01)
Work perm 2				−0.12***	−0.10***
				(0.01)	(0.01)
Age					0.001
					(0.001)
Age Squared					−0.0000***
					(0.0000)
Woman*Non-white	−0.01	−0.01	−0.01	−0.01	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	1.14***	1.29***	1.32***	1.45***	1.48***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
Observations	20,751	20,751	20,751	20,751	20,751
R^2	0.11	0.13	0.14	0.17	0.18
Adjusted R^2	0.11	0.13	0.14	0.17	0.18
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 4: Logistic Models

	Dependent variable:				
	Chance of Poverty				
	(1)	(2)	(3)	(4)	(5)
Woman	−0.09	0.07	0.07	0.04	0.03
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Non-white	0.77***	0.81***	0.81***	0.77***	0.73***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Elementary education	0.80***	0.63***	0.57***	0.49***	0.67***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Superior Education	−1.37***	−1.32***	−1.28***	−1.31***	−1.31***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Secondary Sector		−0.70***	−0.25***	−0.15**	−0.21***
		(0.06)	(0.06)	(0.07)	(0.07)
Tertiary Sector		−0.84***	−0.36***	−0.38***	−0.41***
		(0.05)	(0.06)	(0.06)	(0.06)
Urban			−0.74***	−0.67***	−0.63***
			(0.05)	(0.05)	(0.05)
Work perm 1				−1.25***	−1.24***
				(0.05)	(0.05)
Work perm 2				−0.64***	−0.51***
				(0.05)	(0.05)
Age					0.01
					(0.01)
Age Squared					−0.0004***
					(0.0001)
Woman*Non-white	0.10	0.10	0.09	0.06	0.09
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Constant	−1.77***	−1.13***	−0.96***	−0.24***	−0.09
	(0.04)	(0.06)	(0.06)	(0.07)	(0.17)
Observations	20,751	20,751	20,751	20,751	20,751
Log Likelihood	-9,765.49	-9,622.09	-9,506.13	-9,200.71	-9,109.34
Akaike Inf. Crit.	19,542.97	19,260.18	19,030.25	18,423.43	18,244.67
Note:					*p<0.1; **p<0.05; ***p<0.01

on an individual's predicted chance of poverty: it had very little impact (Table 5). It was found that after accounting for all variables, being non-white leads to a 15.8% higher chance of being in poverty. This places the results of PSM in the middle of those of linear regression (11%) and logistic regression (18.0% - 20.3%). It may be deduced that linear regression is an underestimation of the results, while logistic regression may lead to an overestimation.

Table 5: PSM for non_white: woman, education, work, urban, age, work permit (continued below)

Test statistic	df	P value	Alternative hypothesis
26.61	9428	1.636e-150 * * *	two.sided

mean of the differences
0.1544

Table 7: PSM for woman: non_white, education, work, urban, age, work permit (continued below)

Test statistic	df	P value	Alternative hypothesis
-0.8588	8166	0.3905	two.sided

mean of the differences
-0.00502

Overall, no results were obtained that being a woman had a large positive impact on an individual's predicted chance of poverty after accounting for other relative variables. This does not conform to the results of previous literature, and can be traced back to the data set used. It is evident that there is a strong selection bias for "super women", those who have values for other variables such as work sector and education level that contribute to the reduction of poverty. It is interesting to note that when performing logistic regression, the value of effect of the woman variable did not decrease periodically, but instead slightly changed positively or negatively after the addition of each variable. Overall, the findings of these methods cannot be applied to the Brazilian population at large.

An important finding that corresponds to those of previous literature is the presence of racial differentiation in Brazil (Table 7). All models estimated that being non-white increases an individual's chance of poverty, even after accounting for other variables. The three methods estimated three different values: 11% for linear regression, 18-20.3% for logistic regression (depending on whether the individual is a woman or man) and 15.8% for propensity score matching. This study highlights the fact that if only able to use one model, researchers need to carefully examine which one would best fit the situation and the data available. With such a drastic difference between the results of these models, this could lead to very different interpretations.

Conclusion

Brazil is a country with one of the highest rates of inequality in the world. With a long and complicated history of economical hardship and prosperity, it continues to make inequality reduction one of it's goals in policy. Two types of inequality previously observed in Brazil had been gender and racial inequalities. Although Brazil has been attempting to reverse these inequalities with programs like Bolsa Familia, the effectiveness of their approach is questioned. Furthermore, with the 2020 COVID-19 pandemic, Brazil was once again thrown into uncertainty. This paper had used a data set found on Kaggle to explore the relationship between race, gender, and poverty in Brazil. The data set included over 20000 observations and had information on the individuals race, gender, education level, work sector, work permit holding, age, whether they lived in an urban area, and whether they were measured to be in poverty. Using this data, an analysis was conducted including linear regression, logistic regression, and propensity score matching analyses.

It was found that being a woman had a very little increase on the predicted chance of an individual being in poverty, which does not go along with previous literature findings. When accounting for the gender of the individual, their education level, the sector of their work, whether they live in an urban area, whether they have a work permit, and their age, a clear presence of racial inequality was observed. The predicted chance of being in poverty of individuals rose by 11 percent with linear regression, 18% for women and 20.3% for men with logistic regression, and 15.8% with propensity score matching. This data provides empirical proof that there is evidence of racial discrimination in Brazil.

Unfortunately, the results of this study cannot be applied to the Brazilian population, because there is an evidence selection bias in the data. With an over representation of men and individuals in poverty, as well as the exclusion of unemployed individuals, the data set is not a true representation of Brazil's population. Once all factors are accounted for, such as an individual's education and work sector, the importance of whether the individual is a woman or not is diminished, so it can be concluded that the women observed are part of the "super women" group in Brazil, women who's characteristics naturally decline their chances of poverty, such as education level. As a result, this paper serves as an important study of data quality and inspection.

References

Baumann, R. (2002). *Brazil in the 1990s: An Economy in Transition.* (2001st ed., Vol. 73). Palgrave Macmillan. https://repositorio.cepal.org/bitstream/handle/11362/10795/73147169I_en.pdf?sequence=1&isAllowed=y

Beghin, N. (2008). Notes on Inequality and Poverty in Brazil: Current Situation and Challenges. Oxfam International. https://oxfamilibrary.openrepository.com/bitstream/han dle/10546/112516/fp2p-bp-notes-inequality-poverty-brazil-current-140608-en.pdf

Bradshaw, S., Chant, S., & Linneker, B. (2018). Challenges and Changes in Gendered Poverty: The Feminization, De-Feminization, and Re-Feminization of Poverty in Latin America. *Feminist Economics, 25*(1), 119–144. https://doi.org/10.1080/13545701.2018.1529417

Brynin, M., & Longhi, S. (2015). The Effect of Occupation on Poverty among Ethnic Minority Groups. Joseph Rowntree Foundation. https://www.jrf.org.uk/sites/default/files/jrf/migrated/files/occupation-poverty-ethnic-minority-full.pdf

Cardim De Carvalho, F. J., & Pires De Souza, F. E. (2011). Brazil in the 2000's: Financial Regulation and Macroeconomic Stability._ITF_. http://www.itf.org.ar/pdf/documentos/75-2011.pdf

de Brauw, A., Gilligan, D. O., Hoddinott, J., & Roy, S. (2014). The Impact of Bolsa Familia on Women's Decision-Making Power. *World Development, 59*, 487–504. https://doi.org/10.1016/j.worlddev.2013.02.003

Farah, Nusrat, "Impact of Household and Demographic Characteristics on Poverty in Bangladesh: A Logistic Regression Analysis" (2015). 2015 *Awards for Excellence in Student Research and Creative Activity - Documents.* 3. http://thekeep.eiu.edu/lib_awards_2015_docs/3

Ferderer, J. Pete (1997). *Rules and Outcomes: Brazil in Transition during the 1990s.* Macalester International: Vol. 5, Article 14.

Ferreira, F. H. G., Lanjouw, P., & Neri, M. (2003). A robust poverty profile for Brazil using multiple data sources. *Revista Brasileira de Economia, 57*(1), 59–92. https://doi.org/10.1590/s0034-71402003000100003

Garcias, M. O., & Kassouf, A. L. (2021). Intergenerational mobility in education and occupation and the effect of schooling on youth's earnings in Brazil. *EconomiA, 22*(2), 100–113. https://doi.org/10.1016/j.econ.2021.05.001

Gradin, C. (2009). Why is Poverty So High Among Afro-Brazilians? A Decomposition Analysis of the Racial Poverty Gap. *Journal of Development Studies, 45*(9), 1426–1452. https://doi.org/10.1080/00220380902890235

Hall, A. (2008). Brazil's Bolsa Familia: A Double-Edged Sword? *Development and Change, 39*(5), 799–822. https://doi.org/10.1111/j.1467-7660.2008.00506.x

Hughes, M. R. (1997). Sample Selection Bias in Analyses of the Political Democracy and Income Inequality Relationship. *Social Forces, 75*(3), 1101. https://doi.org/10.2307/2580533

Layton, M., & Smith, A. (2017). Is It Race, Class, or Gender? The Sources of Perceived Discrimination in Brazil. *Latin American Politics and Society, 59*(1), 52-73. doi:10.1111/laps.12010

Lima, M., & Prates, I. (2018). Racial Inequalities in Brazil: A Persistent Challenge. *Paths of Inequality in Brazil*, 113–134. https://doi.org/10.1007/978-3-319-78184-6_6

Mantoan, E., Centeno, V., & Feijo, C. (2021). Why has the Brazilian economy stagnated in the 2010s? A Minskyan analysis of the behavior of non-financial companies in a financialized economy. *Review of Evolutionary Political Economy*, 2(3), 529–550. https://doi.org/10.1007/s43253-021-00051-6

Medeiros, M., Barbosa, R. J., & Carvalhaes, F. (2020). Educational expansion, inequality and poverty reduction in Brazil: A simulation study. *Research in Social Stratification and Mobility, 66*, 100458. https://doi.org/10.1016/j.rssm.2019.100458

Nassif-Pires, L., Cardoso, L., & Matos De Oliveira, L. (2021). Gender and Race in the Spotlight During the COVID-19 Pandemic: The Impact of the Emergency Benefit on Poverty and Extreme Poverty in Brazil. Levy Economics Institute of Bard College. https://www.levyinstitute.org/pubs/pn_21_2.pdf

Neri, M. (2019) Inequality in Brazil: Inclusive Growth Trend of this Millennium is Over. WIDER Policy Brief 2019/1. https://www.wider.unu.edu/sites/default/files/Publications/Policy-brief/PDF/PB2019-1-Inequality-in-Brazil.pdf

Neves, J. A., Vasconcelos, F. D. A. G. D., Machado, M. L., Recine, E., Garcia, G. S., & Medeiros, M. A. T. D. (2020). The Brazilian cash transfer program (Bolsa Família): A tool for reducing inequalities and achieving social rights in Brazil. *Global Public Health, 17*(1), 26–42. https://doi.org/10.1080/17441692.2020.1850828

Nunn N. Slavery, Inequality, and Economic Development in the Americas: An Examination of the Engerman-Sokoloff Hypothesis. In: Helpman E Institutions and Economic Performance. Cambridge: Harvard University Press ; 2008. pp. 148-180.

OECD. (2021). Economic Policy Reforms 2021: Going for Growth. https://www.oecd.org/economy/growth/Brazil-country-note-going-for-growth-2021.pdf

Pereira, C. (2016). Ethno-Racial Poverty and Income Inequality in Brazil. *Working Paper No. 60.* https://www.commitmentoequity.org/publications_files/Brazil/CEQ_WP60_Pereira_Nov23_2016.pdf

ECLAC, 2018. *Preliminary Overview of the Economies of Latin America and the Caribbean 2017.*

ECLAC, 2021. *Preliminary Overview of the Economies of Latin America and the Caribbean: 2020.*

Rocha, S. (2019). Poverty upsurge in 2015 and the rising trend in regional and age inequality among the poor in Brazil. *Nova Economia, 29*(1), 249–275. https://doi.org/10.1590/0103-6351/3992

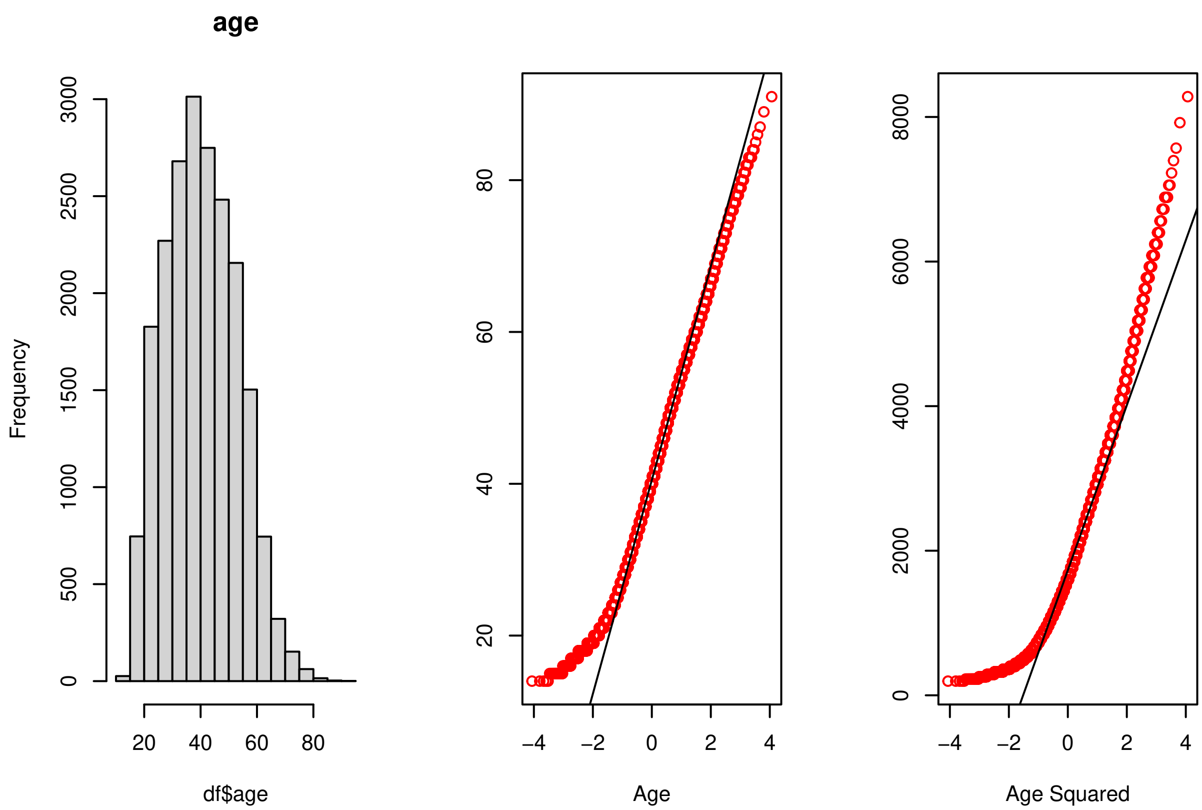
Saefuddin, A., Setiabudi, N. A., & Achsani, N. A. (2011). On Comparisson between Ordinary Linear Regression and Geographically Weighted Regression: With Application to Indonesian Poverty Data. *European Journal of Scientific Research, 75*(2). http://achsani.blog.mb.ipb.ac.id/files/2011/08/mb.ipb_.ac_.id_EJSR_57_2_10.pdf

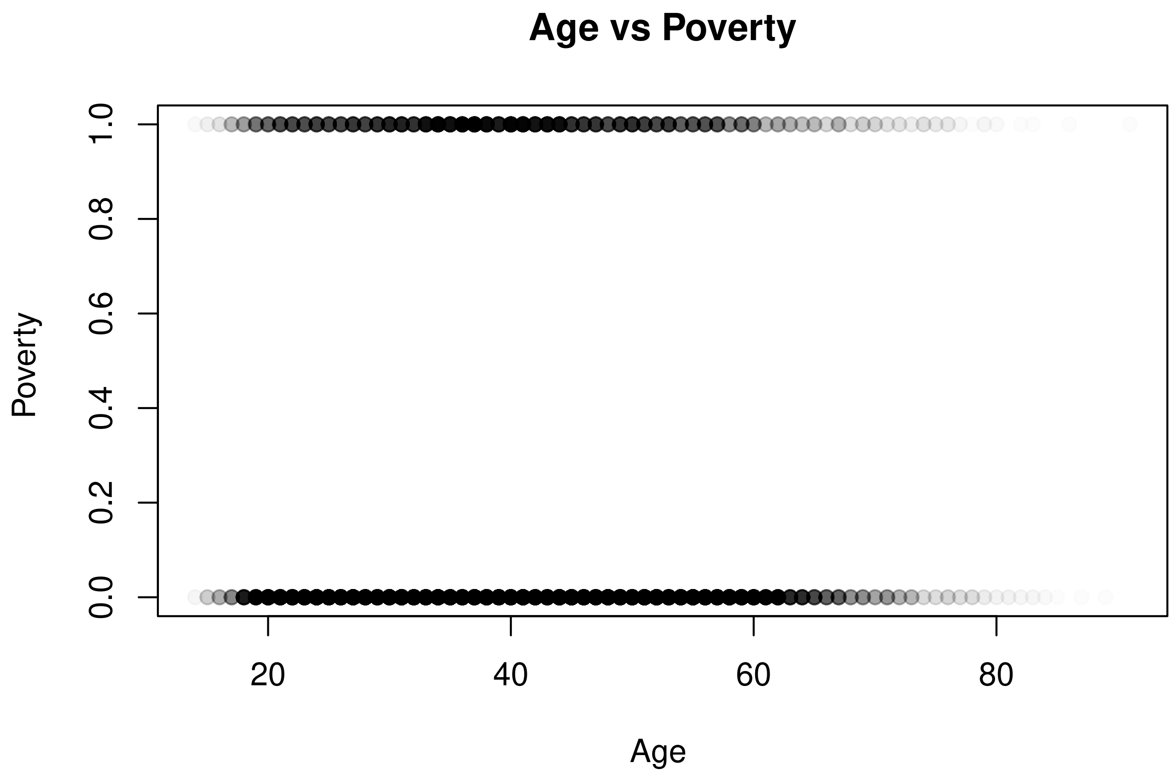
Salata, A. (2020). Race, Class and Income Inequality in Brazil: A Social Trajectory Analysis. *Dados, 63*(3). https://doi.org/10.1590/dados.2020.63.3.213

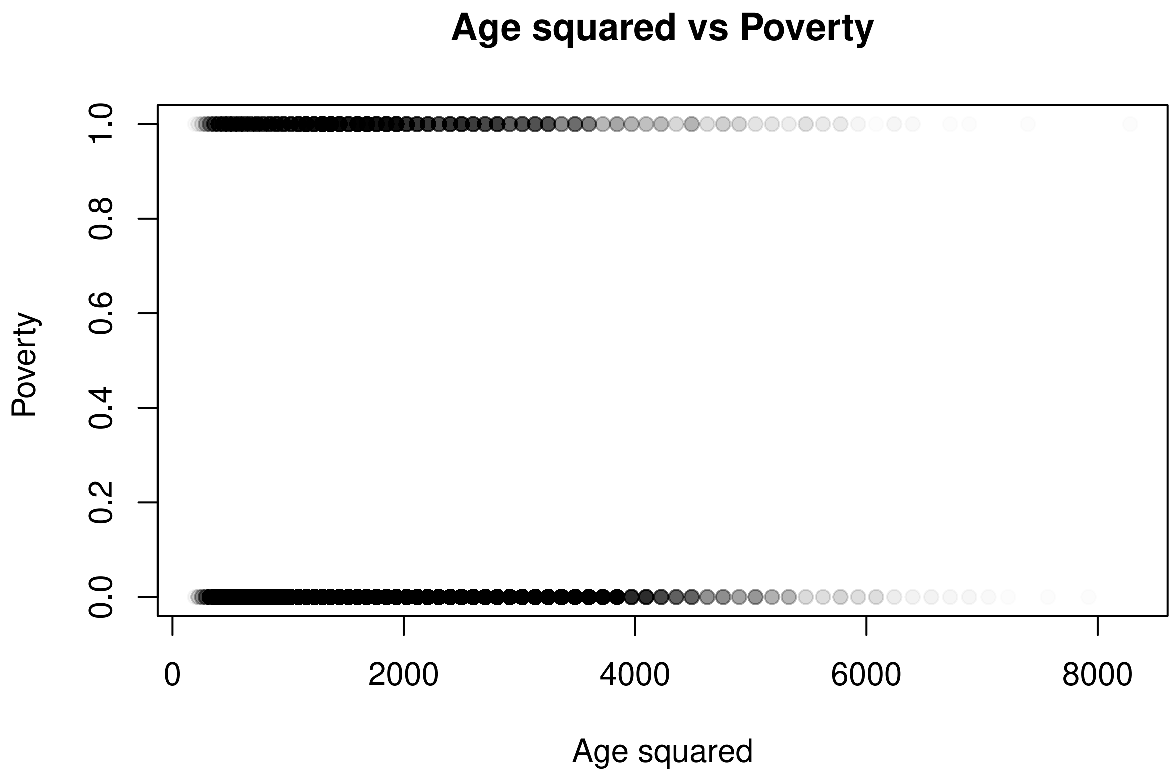
United Nations Publications & United Nations Publications. (2019). *Preliminary Overview of the Economies of Latin America and the Caribbean 2017*. UN.

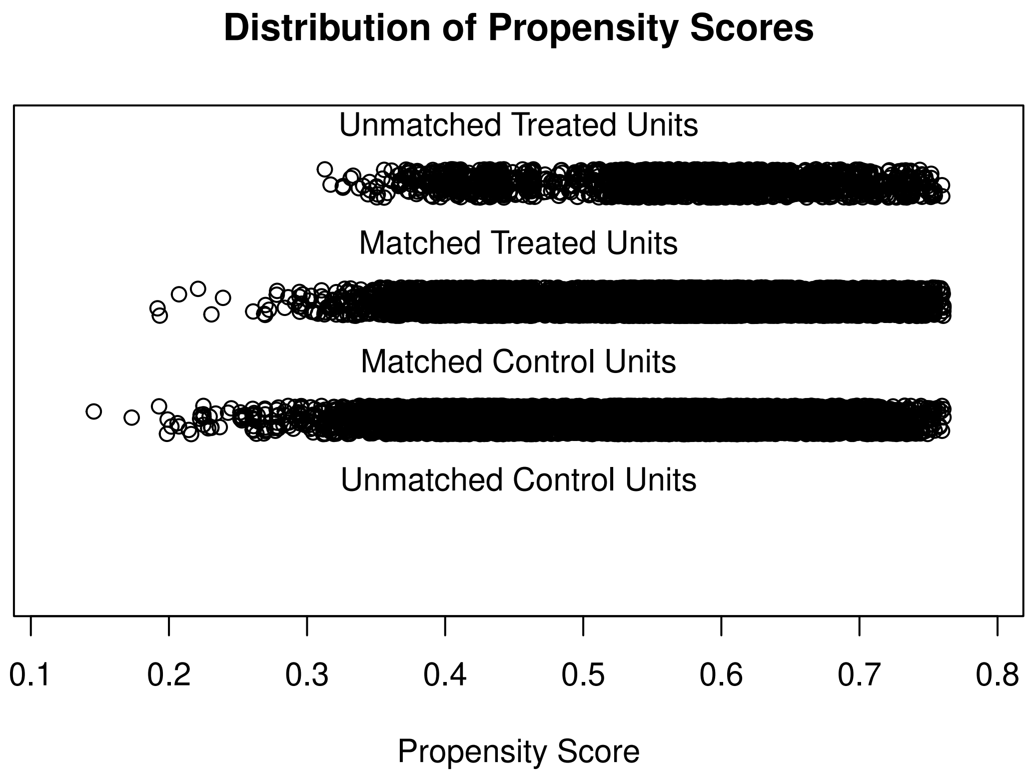
Verner, D. (2004). Education and its Poverty-Reducing Effects: The Case of Paraíba, Brazil. *Policy Research Working Papers*. https://doi.org/10.1596/1813-9450-3321

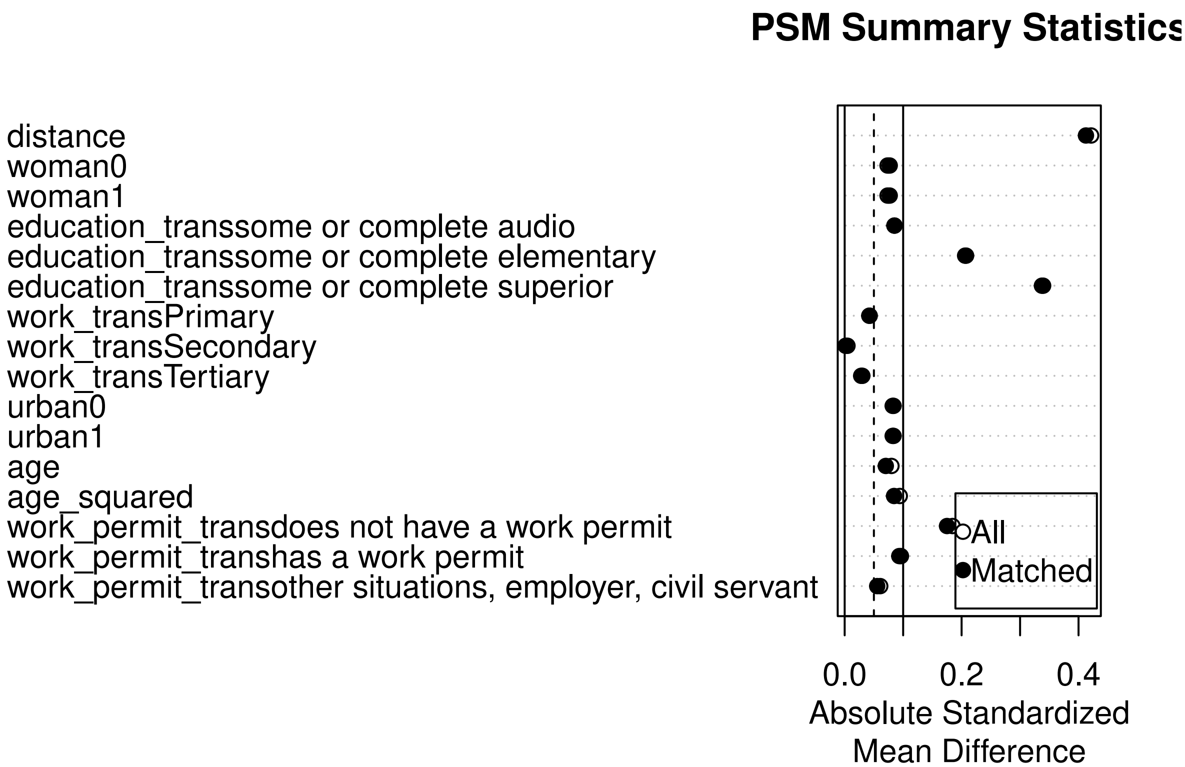
Appendix

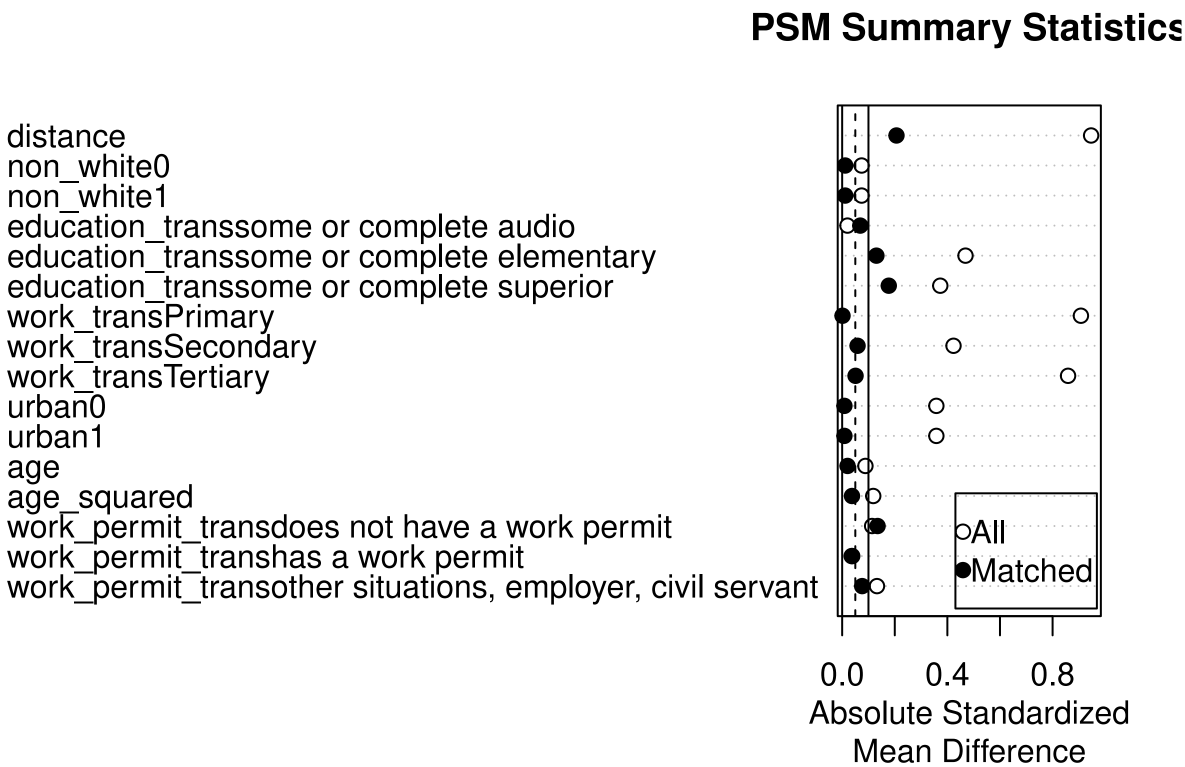



Age vs Poverty
1.0
0.8
0.6
0.4
0.2
0.0
Poverty
20
40
60
80
Age


Age squared vs Poverty
Poverty
0.0
0.2
0.4
0.6
0.8
1.0
0
2000
4000
6000
8000
Age squared


Distribution of Propensity Scores
Unmatched Treated Units
Matched Treated Units
Matched Control Units
Unmatched Control Units
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Propensity Score


PSM Summary Statistics
distance
woman0
woman1
education_transsome or complete audio
education_transsome or complete elementary
education_transsome or complete superior
work_transPrimary
work_transSecondary
work_transTertiary
urban0
urban1
age
age_squared
work_permit_transdoes not have a work permit
work_permit_transhas a work permit
work_permit_transother situations, employer, civil servant
All
Matched
0.0
0.2
0.4
Absolute Standardized
Mean Difference


PSM Summary Statistics
distance
non_white0
non_white1
education_transsome or complete audio
education_transsome or complete elementary
education_transsome or complete superior
work_transPrimary
work_transSecondary
work_transTertiary
urban0
urban1
age
age_squared
work_permit_transdoes not have a work permit
work_permit_transhas a work permit
work_permit_transother situations, employer, civil servant
All
Matched
0.0
0.4
0.8
Absolute Standardized
Mean Difference