

# Voting Patterns and Socio-Economic Tides: A Twenty-Year Analysis of Brazil's Developmental Dynamics\*

Hao Wang / Siqi Fei / Xuecheng Gao

February 12, 2024

## Abstract

This paper presents a comprehensive analysis of socio-economic data for Brazil spanning from 1996 to 2016. The data, representing over 33,000 Brazilians across six electoral cycles, encompasses economic indicators such as income per capita, low-income rates, and unemployment rates, as well as political variables like electoral years and district codes. Through meticulous data visualization and statistical analysis, we examine the distributions of these key indicators and their interplay with electoral processes.

Our findings reveal diverse economic conditions as indicated by the spread of income per capita and a bimodal distribution in low-income rates, suggesting the existence of two distinct economic subgroups within the population. Furthermore, the unemployment rate exhibits a concentration around lower rates with a tail extending towards higher values, highlighting the prevailing low unemployment with instances of higher rates.

By integrating these economic measures with educational data like dropout rates, we uncover complex correlations and the varying impact of voter numbers on educational outcomes. We discuss the implications of these relationships, exploring how electoral districts and voter numbers correlate with socio-economic conditions.

This study provides valuable insights into the socio-economic and political dynamics of Brazil, emphasizing the importance of electoral systems in shaping economic and educational landscapes. It contributes to a deeper understanding of how policy and governance can influence socio-economic development and underscores the need for nuanced policy interventions to address economic disparities and promote equitable growth.

---

\*Code and data are available at: <https://github.com/groundUofT888/Term-Paper-2.git> ; Replication on Social Science Reproduction platform available at: [www.aeaweb.org/articles?id=10.1257/app.20210529](http://www.aeaweb.org/articles?id=10.1257/app.20210529)

# **1 Introduction**

## **1.1 Two-round elections in Brazil**

Brazil's two-round electoral system is a relatively common electoral method. The mode of election is to select the winner after two votings. In the first-round voters can choose the candidate they support among multiple candidates. If one of the candidates gets more than half of the votes in the first round, then he can be directly elected. If no candidate receives more than half of the votes in the first round, a second round will take place. In the second round of voting, voters will choose between the two candidates who received the highest votes in the first round. Election regulations 1. Voters must be registered at least 151 days before the election. 2. Voting is compulsory for all literate Pakistani citizens between the ages of 18 and 69. 3. The election is held on the first Sunday in October.

## **1.2 Economic and social issues in Brazil**

Major social and economic issues in Brazil. First, Brazil's poverty gap. There are many depressed areas in Brazil. Many people live below the poverty line and lack basic living conditions such as housing, education, medical care, and jobs. This leads to social instability. Second, the high crime rate. Some cities in Brazil have high crime rates and violence problems, such as robberies, gun proliferation, and drug trafficking. These problems mainly stem from poverty caused by uneven distribution of wealth and the deterioration of local security. Third, Brazil has high unemployment. The number of unemployed people is increasing in Brazil, and many structural unemployment results in high unemployment rates for a large number of low-skilled laborers. Also lower wages for many people worsen poverty and inequality. Fourth, there are issues with the education system. Brazil's education system has many problems, such as uneven education quality, uneven allocation of government funds for education, shortage of professors, and high dropout rates.

## **1.3 Overview**

In 2022, Brazil's presidential election broke new grounds with an unprecedented turnout of over 150 million voters, capturing the attention of national and international media. The world had seen not just about the political tactics of the candidates, but also the public's vested interest in how these elections could shape their day-to-day lives. This paper aims to study the influence of Brazil's electoral system on three critical socio-economic indicators: income level, unemployment rate and educational resources. With a comprehensive look at recent data and trends, we will investigate how the details of voting and representation create a chain reaction throughout the economy, affecting everything from job opportunities to the quality of education. By collecting together statistical analysis with real-world implications, we'll provide a clearer picture of the electoral system's impact on the Brazilian populace and discuss the broader implications of these findings on social equality and economic stability.

# **2 Data**

## **2.1 Source**

This paper replicates the survey data that was originally collected by Moya Chin, AMERICAN ECONOMIC JOURNAL: APPLIED ECONOMICS VOL. 15, NO. 3, JULY 2023 (pp. 183-209).

It uses data sets constructed from four data sources: - 1) Brazil municipal election data (Tribunal Superior Eleitoral, 1996-2016); - 2) Brazil demographic censuses (Instituto Brasileiro de Geografia e Estatística, 1980, 1991, 2000, 2010); - 3) Brazil school census (Ministério da Educação, 1997-2016); - 4) nighttime lights (National Oceanic and Atmospheric Administration, 1997-2016a,-)

## **2.2 Methodology**

R(R Core Team 2023) is the language and environment used for this analysis,supported by the robust features of Posit Cloud as our primary integrated development environment (IDE). Library includes

`ggplot2`(Wickham 2016), `here`(Müller 2020).

Git(version 2.39.2) is used as version control.

## 2.3 Features

### `election.csv/election_clean.csv`

The data set offers a comprehensive longitudinal collection of socio-economic indicators for Brazil from 1996 to 2016. Over the course of these twenty years, the data set captures the socio-economic fluctuations and trends across six Brazilian electoral cycles, with a granular focus on a variety of indicators.

- **Time frame:** Spanning two decades, the data set provides insights into the socio-economic changes over time, coinciding with multiple electoral events in Brazil.
- **Volume:** Each election year is detailed with 5,567 data entries, resulting in a total of 33,188 entries over the entire period, allowing for in-depth analysis and trend identification.
- **Indicators Included:**
  - `cut_dist`: A geographical identifier for districts or electoral cuts, offering spatial analysis opportunities.
  - `elect_year`: The year of the election, anchoring the socio-economic data within a political timeline.
  - `inc_post`: Per capita income figures post-election, reflecting the general economic well-being.
  - `inc_0_50_post`: Income figures for the bottom 50% post-election, highlighting the economic status of the less affluent demographic.
  - `edu_univ`: A variable intended to capture educational attainment or resources, yet unspecified in the initial data extract.
  - `unempl_post`: Post-election unemployment rates, indicating labor market conditions.
  - `tse_code`: Potentially linked to electoral zones or districts, allowing for analysis of electoral influence on socio-economic aspects.

### `schools.csv`

The dataset includes an array of variables that shed light on the multifaceted nature of Brazil's socio-economic and political landscape:

- **Electoral Dynamics:** With variables like `elect_year` and `tse_code`, the dataset chronicles the electoral events, reflecting the evolving political scene at regular four-year intervals.
- **Geographic and Demographic Insights:** Indicators such as `ibge7_code`, `cut_dist`, and `pop_den_pre` provide detailed geographic and demographic information, highlighting regional disparities and population distribution.
- **Economic Measures:** Economic variables such as `dep_pca_ptile_mean`, `eq_pca_ptile_mean`, `inc_T0`, and `inc_0_50_T0` reveal insights into the income levels, economic stratification, and wealth distribution within Brazil.
- **Educational Outcomes:** Metrics like `MAT_DROP_mean`, `MAT_PASS_mean`, and `MAT_FAIL_mean` offer a glimpse into the educational sector's performance, a critical aspect of Brazil's socio-economic fabric.
- **Social Indicators:** Social factors such as illiteracy rates (`illit_T0`), unemployment figures (`unempl_T0`), and measures of inequality (`gini_T0`) provide a deeper understanding of the societal challenges that have shaped Brazil's development over the years.

### 3 Result

Our statistical analysis began with a computation of the sample size which represents the population of Brazilians in 2022 election, with a population size of 203.1 million. We determined the sample size as 33188 in this study, ensuring the analysis result would be within an acceptable margin of error. This sample size was substantiated by Quota sampling, ensuring the representations of our data set.

The sample we used were randomly selected every 4 year from 1996 to 2016. Each year, there are 5567 Brazilians were chose to be the sample set. we generated all the sample sets and run the data analysis to study the factors.

#### 3.1 Income per Capita

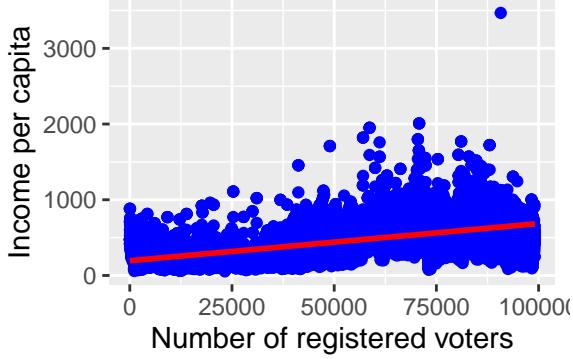


Figure 1: Income per capita

For the above Figure 1, from 1996 to 2016, the trend line indicates a slight positive correlation between the number of registered voters and income per capita. This could suggest that in areas with more registered voters, there may be a slightly higher income per capita.

#### 3.2 Low Income Rate

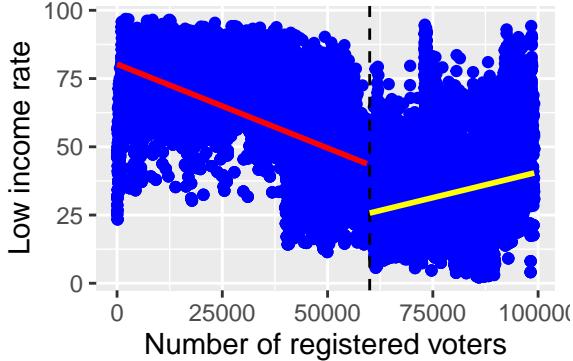


Figure 2: Low income rate

We defined the lowest 50% income as low income. The plot Figure 2 appears to show a slight negative correlation, as indicated by the downward trend line. This suggests that as the number of registered voters increases, the low-income rate decreases slightly. However, the wide spread of the data points indicates that many other factors likely influence the low-income rate beyond just the number of registered voters.

### 3.3 Unemployment rate

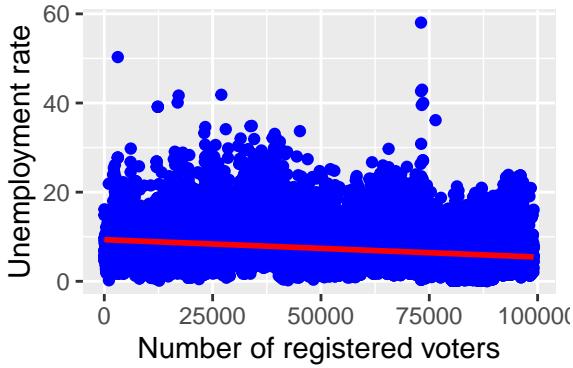


Figure 3: Unemployment rate

Analyzing the Figure 3, there are noticeable differences. The plot shows a vast spread of unemployment rates across varying numbers of registered voters, with a trend line that remains relatively flat. This suggests there isn't a strong or clear correlation between the number of registered voters and the unemployment rate within the represented data set.

### 3.4 Educational Resources

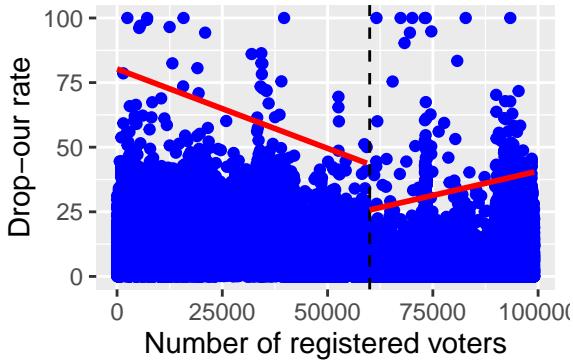


Figure 4: education

Analyzing the plot presented in Figure 4, titled "Educational Resources," we see a substantial variation in dropout rates across a range of registered voter numbers. The dual trend lines, both descending, suggest that areas with more registered voters tend to have lower dropout rates, yet the change in the slope at around the 50,000 voter mark could imply different dynamics in larger districts. Despite the overall negative correlation, the broad spread of the data points, especially in regions with fewer voters, indicates that voter numbers alone may not be a definitive predictor of dropout rates. The relationship is not uniform across the dataset, highlighting the complexity of factors that influence educational outcomes.

### 3.5 Overview

Distribution of Economic Indicators

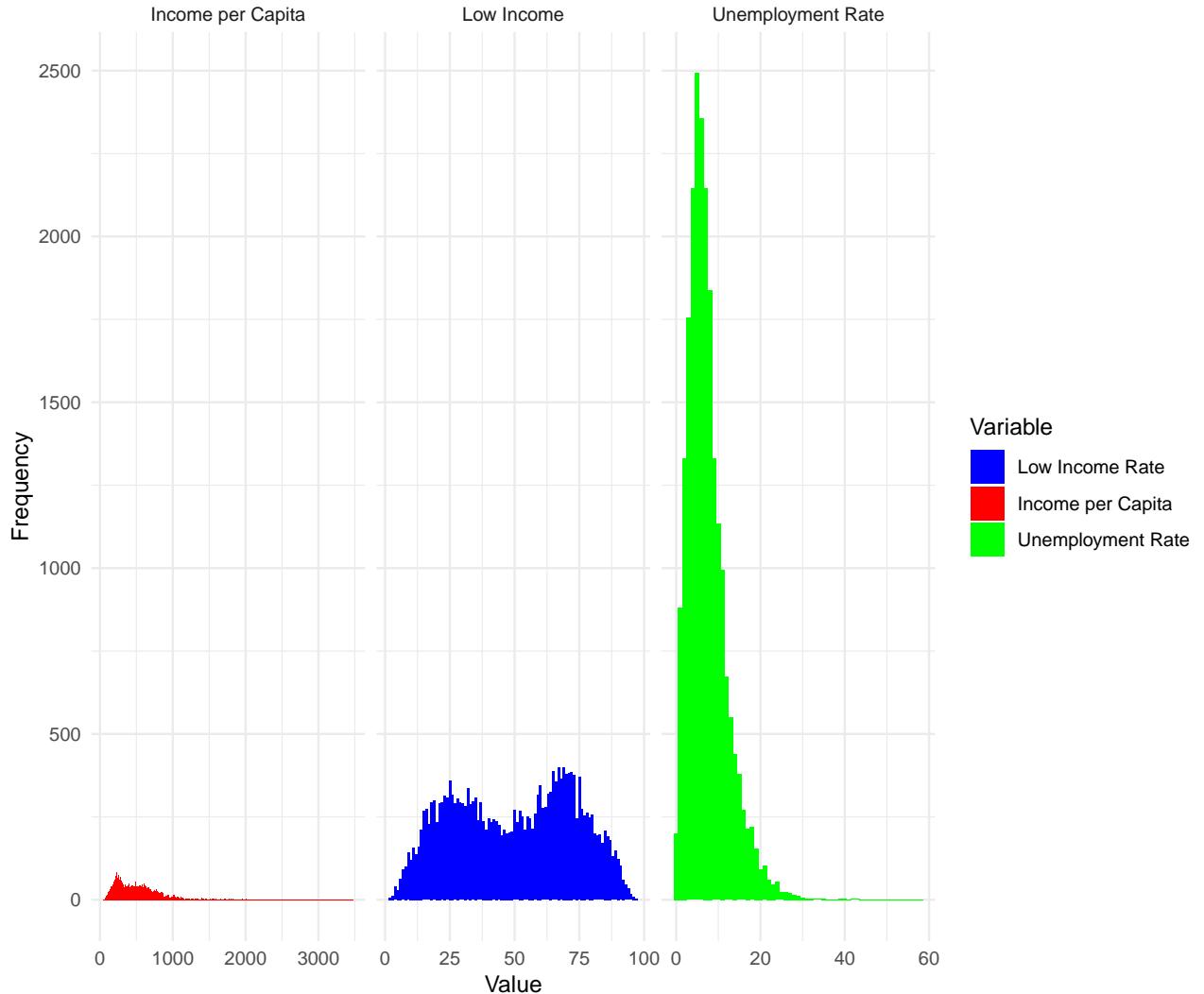


Figure 5: histogram

Analyzing the histogram in Figure 5, we see distinct distributions for three key economic indicators. The Income per Capita is broadly spread, suggesting varied economic conditions. In contrast, the Low Income Rate displays a bimodal distribution, hinting at two prevalent income groups within the population. The Unemployment Rate is predominantly concentrated around a lower value but with a long tail toward higher rates, indicating that most of the population experiences lower unemployment, with fewer instances of high unemployment. This visual representation underscores the complexity of economic health and inequality within the dataset.

## 4 Discussion

### 4.1 Findings

In the paper, Moya Chin uses the regression discontinuity design to compare single-round and two-round elections in Brazil. Our paper makes two findings based on Moya Chin's data and theory.

1. The two rounds of elections in Brazil have a good impact on Brazil's downstream economy.
2. Candidates prefer to attract low-income people and people with low education during the election.

We found from Moya Chin's paper that politicians' behavior during elections affects the provision of local public facilities, but the allocation of these resources changes when politicians are elected.

We find that two-round elections can incentivize candidates to provide resources for downstream economic and public education. In terms of education, the dropout rate has been significantly reduced and the literacy rate has increased. According to census data in 2000 and 2010, the primary school literacy rate has been on an upward trend, while the dropout rate has been on a downward trend. According to the school census, the school's infrastructure has been significantly improved during the campaign, and public facilities such as the school's library, gymnasium, and classrooms have received more funding. Resource levels in public schools have increased and differences between schools have become smaller.

In terms of the provision of public goods, we found that in the two-round election of urban school resource allocation compared with the single-round municipal government, these schools receive more resources in terms of equipment and infrastructure, and these resources are more distributed among schools. Politicians provide a large number of products such as computers, photocopiers, projectors, etc. These products reduce dropout rates and increase literacy rates. These grants have improved local education standards and have a positive impact on local education. This shows that two rounds of elections will bring about more public goods provision, which will have a positive impact on education. These improvements in public education are important for long-term socioeconomic development because better education can create a more highly skilled workforce in the future and may reduce structural unemployment in the future.

However, we find that two rounds of elections have limited economic impact in the mid-term. According to the Figure, it can be seen that per capita income increased during the election period, and the number of low-income people decreased compared with before. But these results were not evident in the statistics. This suggests that while two rounds of elections will improve the downstream economy, changes in income levels and unemployment will take longer to materialize.

We also found politicians in Brazil using political manipulation and tactics in elections. In Brazil, low-income groups make up a larger share of the population, so attracting votes from this group could have a significant impact on election results. Therefore, even if high-income groups pay more attention to elections than low-income groups, politicians will allocate more resources to the majority of voters.

We believe that Brazil's two rounds of elections will bring short-term benefits to the downstream economy and education. However, long-term changes in Brazil's economic and social problems require long-term observation.

## **4.2 Ethical Implications**

The datasets and original sources used in their paper come from Brazil demographic censuses and Brazil school census. Demography refers to summarizing the demographic characteristics of a community, such as gender, age, race, education level, socioeconomic status, etc. Ethical issues may arise if used improperly during the collection, analysis, interpretation, and dissemination of demographic data.

Demographics impact minority communities, which can lead to structural inequalities. When analyzing demographic variables, the author selected specific variables for analysis based on his research questions and hypotheses. However, these highly characteristic data will have a negative impact on some social groups. For example, high income, low income, dropout rate, etc. This may reinforce social hierarchies in which social groups are compared, exacerbating prejudice and stereotypes and causing harm to minority populations. In the paper, the author compares the interest of high-income people and low-income people in voting. This may cause certain social conflicts and encourage the idea of inequality among different groups. When explaining differences between social groups in the conclusion of the author's paper, it can lead to increased social conflict.

## **4.3 Accounting for Bias**

In the paper, Moya Chin proposes a stylized model that explores the motivations of politicians to appeal to a wider range of voters in elections and to provide public goods in different ways. The author uses this model to observe the Economic Consequences of Electoral Rules in Brazil. The data used by the author mainly comes from Brazil demographic censuses, Brazil school Census, and Brazil municipal election data. Bias can occur when collecting quantitative and qualitative data.

The problem with these data sets and original sources is that the variables used by the authors to construct the equipment index come from the Brazilian School Census and the Brazilian demographic censuses. Considering Brazil's national conditions and the government's influence on the region, Response Bias may exist in the data set of Brazil demographic censuses. Census respondents may choose not to participate in the census or provide inaccurate information because of privacy concerns or distrust of the government. The school census data comes from students in schools. Considering the lack of educational resources in poor areas of Brazil, there will be a sampling frame bias. The earliest census data used by the author is from 1980. At that time, in some remote or poor areas, it may be more difficult for enumerators to reach or collect data due to lack of infrastructure and information technology. This may lead to underestimation or inaccurate population data in these areas, and the sample may not be representative of the characteristics of the entire population.

## **4.4 Limitations**

Our analysis is based on a data set from the Moya Chin literature, so most of our analysis is limited to the original survey situation.

Because the data are not comprehensive enough, we cannot compare with many variables in the original data, such as whether political intervention and media propaganda affect voters' thoughts and voting results. Excessive propaganda of policies and political interference are suspected of manipulating voters and leading to distorted election results. Also the political contributions a politician receives during an election can influence the outcome of the election.

The model created by Moya Chin cannot be fully applied during our analysis, which results in our analysis not being as complete as Moya Chin described. So we have to give up some variables.

## **4.5 Future Research**

We believe that this research direction can receive more investment in the future. We hope that the relationship between electoral systems, education, and socioeconomic outcomes can be further explored in the future. We hope to observe and predict the impacts of elections over the long term to fully understand the

impact of the electoral system on economic indicators, and to reveal the possible long-term effects of changes in the electoral system.

## References

- Chin, Moya. 2023. “When Do Politicians Appeal Broadly? The Economic Consequences of Electoral Rules in Brazil.” *American Economic Journal: Applied Economics* 15 (3): 183–209. <https://doi.org/10.1257/app.20210529>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://here.r-lib.org/>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org>.