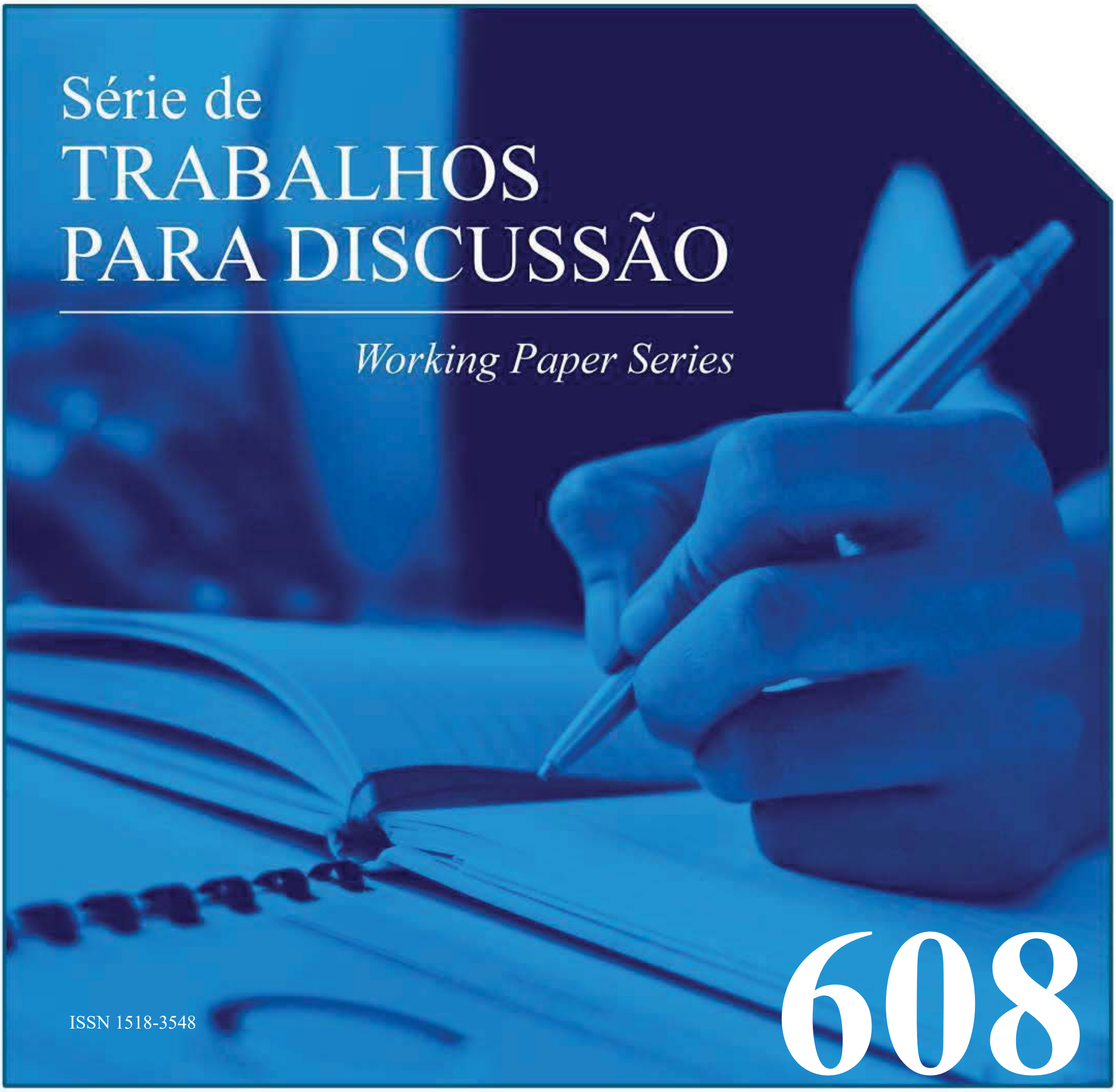


Série de
TRABALHOS PARA DISCUSSÃO

Working Paper Series

ISSN 1518-3548

608

Novembro 2024

Measuring Inequality Using Electronic Payment Data

Carlos A. Piccioni, Saulo B. Bastos, Daniel O. Cajueiro

BANCO CENTRAL DO BRASIL

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	no. 608	Novembro	2024	p. 3-44

Working Paper Series

Edited by the Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Rodrigo Barbone Gonzalez

Co-editor: Eurilton Alves Araujo Jr

Head of the Research Department: André Minella

Deputy Governor for Economic Policy: Diogo Abry Guillen

The Banco Central do Brasil Working Papers are evaluated in double-blind referee process.

Citizen Service Division

Banco Central do Brasil

Deati/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Toll Free: 0800 9792345

Fax: +55 (61) 3414-2553

Internet: http//www.bcb.gov.br/?CONTACTUS

Non-Technical Summary

Economic inequality, encompassing differences in income, wealth, and consumption, is a widely discussed issue in economics due to its significant impact on social welfare and political stability. Measuring this inequality poses major challenges, especially at the municipal level, where traditional data are scarce and often outdated. Our study proposes a new approach to measure consumption inequality using electronic payment data, such as credit card and Pix transactions, providing a more current and accurate analysis. Our research aims at answering how consumption inequality, measured through electronic payment data, relates to the economic complexity of Brazilian municipalities. We found that consumption inequality, as measured by electronic payment data, is moderately correlated with income inequality calculated from census data, but generally indicates higher levels of inequality. Additionally, our study reveals a negative and non-linear relationship between consumption inequality and economic complexity: municipalities with higher economic complexity tend to exhibit lower consumption inequality. Our innovative approach, using electronic payment data, offers a valuable tool for real-time monitoring and policy formulation.

Sumário Não Técnico

A desigualdade econômica, que inclui diferenças de renda, riqueza e consumo, é um tema amplamente discutido na economia devido ao seu impacto significativo no bem-estar social e na estabilidade política. Medir essa desigualdade apresenta grandes desafios, especialmente no nível municipal, onde dados tradicionais são escassos e frequentemente desatualizados. Nosso estudo propõe uma nova abordagem para medir a desigualdade de consumo usando dados de pagamentos eletrônicos, como cartões de crédito e Pix, fornecendo uma análise mais atual e precisa. Nossa pesquisa busca responder como a desigualdade de consumo, medida através de dados de pagamentos eletrônicos, se relaciona com a complexidade econômica dos municípios brasileiros. Descobrimos que a desigualdade de consumo está moderadamente correlacionada com a desigualdade de renda calculada com base nos dados do censo, mas geralmente indica níveis mais altos de desigualdade. Além disso, nosso estudo revela uma relação negativa e não linear entre a desigualdade de consumo e a complexidade econômica: municípios com maior complexidade econômica tendem a apresentar menor desigualdade de consumo. Nossa abordagem inovadora, utilizando dados de pagamentos eletrônicos, oferece uma ferramenta valiosa para monitoramento e formulação de políticas públicas em tempo real.

Measuring Inequality Using Electronic Payment Data

Carlos A. Piccioni*[1], Saulo B. Bastos†[1], and Daniel O. Cajueiro‡[2,3,4]

[1]Department of Banking Operations and Payments System (Deban), Banco Central do Brasil (BCB), Setor Bancário Sul Q. 3 Bloco B - Asa Sul, Brasília - DF, 70074-900
[2]Department of Economics - Universidade de Brasília (UnB), Campus Universitario Darcy Ribeiro - FACE, 70910-900, Brasília - DF, Brazil.
[3]National Institute of Science and Technology for Complex Systems (INCT-SC), Universidade de Brasília (UnB), 70910-900, Brasília - DF, Brazil
[4]Machine Learning Laboratory in Finance and Organizations (LAMFO), Universidade de Brasília (UnB), 70910-900, Brasília - DF, Brazil

Abstract: We measure municipal-level inequality based on electronic payment data, specifically credit card and Pix payments, which we consider as a proxy for consumption. Our consumption inequality measure is correlated with income inequality calculated using census data, and it exhibits similar regional behavior, although it indicates higher inequality on average, given the nature of the data used. As an application, we assess the relationship between our inequality measure and the Economic Complexity Index (ECI) at the municipal level. We find a negative relationship, indicating that higher economic complexity is associated with lower consumption inequality. Additionally, the relationship is non-linear: with increasing ECI, the influence on consumption inequality becomes more significant.

Keywords: Consumption Inequality. Electronic Payment Methods. Economic Complexity Index (ECI). Brazilian Municipalities.

JEL Classification: D31. E21. O12.

*carlos.piccioni@bcb.gov.br
†saulo.benchimol@bcb.gov.br
‡danielcajueiro@gmail.com

1 Introduction

Economic inequality, which includes differences in income, wealth, and consumption, is a highly discussed issue in economics because of its significant impact on social welfare and political stability (Alesina et al., 2004; Roe and Siegel, 2011). However, the task of measuring it presents significant challenges. Traditionally, inequality is quantified through sample surveys or census data, generally at national, regional, or state levels. These approaches, however, are not without problems: they are prone to errors, consume considerable public resources, and may underestimate inequality (Medeiros et al., 2015). Moreover, they are carried out infrequently. For instance, Brazil's demographic census, the sole survey capable of providing population data for municipal-level inequality calculations, is conducted only once every ten years, on average. This infrequent data collection can limit the effectiveness of proposing and evaluating public policies aimed at promoting equity[1]. It becomes impractical to assess the impact of a political cycle (4 years) on the inequality of Brazilian municipalities. Additionally, available data predominantly pertain to the income dimension. Data on consumption, which might offer a better measure of well-being (Hassett and Mathur, 2012; Meyer and Sullivan, 2009; Trapeznikova, 2019), are scarcer (Attanasio and Pistaferri, 2016).

In our research, we propose to measure municipal-level consumption inequality using a new database: the electronic payment methods from the Brazilian Central Bank's Payment System. This dataset includes credit card data and Pix transactions, the instant transfer and payment instrument that now exceeds more than 3.5 billion transactions per month. With this database, we were able to calculate the Gini index, a commonly used measure of inequality, which we consider a suitable metric for assessing consumption inequality in Brazilian municipalities. This approach allows for a timely analysis with up-to-date data[2]. The most recent available municipal inequality data, for example, are from the 2010 census. We can also mention as an advantage that the data are neither sample-based nor declarative, hence not subject to common survey errors [3]. For these reasons, we believe that our measure of consumption inequality has the potential to support public policies, especially at the municipal level, that require more timely diagnoses and monitoring.

Despite the Electronic Payment data serving as a proxy for consumption, we compare the inequality calculated from our database with the income inequality calculated using data from the 2010 Brazilian Institute of Geography and Statistics (IBGE) census. Even with the substantial time lag, as inequality tends to exhibit some degree of persistence, we demonstrate a moderate correlation between our inequality index and income inequality measured from census data.

We also present an application for our inequality index: we explore the relationship between inequality and economic complexity, a concept that assesses the sophistication of economic activities in a given economy (Hidalgo and Hausmann, 2009), adapted to the municipal level in our case.

[1]Inequality is positively correlated with criminality and negatively correlated with income growth at the municipal level, and people living in more unequal municipalities classify themselves as less happy than those living in more egalitarian places (Glaeser et al., 2009).

[2]We chose the Gini index as a measure of inequality for comparative purposes with other studies and also because it is the most commonly used measure of inequality (De Maio, 2007).

[3]Moore et al. (2000); Bee and Rothbaum (2019) review the literature on measurement issues in surveys, showing that, generally, survey respondents underreport income. Hokayem et al. (2015) show that populations with the lowest and highest incomes are those most likely not to respond to income questionnaires. Burkhauser et al. (2018) use administrative data (income tax data) to demonstrate that, in the case of the United Kingdom, the increase in income inequality measured using survey data might be underestimated.

This relationship, particularly concerning income inequality, has been the subject of recent debate when studied at the country level (Hidalgo, 2021; Hartmann et al., 2017; Lee and Vu, 2020; Chu and Hoang, 2020; Lee and Wang, 2021; Pham et al., 2023; Amarante et al., 2023) and also at the regional level, albeit by a limited number of studies (Sbardella et al., 2017; Gao and Zhou, 2018; Török et al., 2022; Bandeira Morais et al., 2021). The literature's findings are still mixed, so we hope to contribute to the discussion. We calculated the Economic Complexity Index (ECI) for each Brazilian municipality and, through cross-sectional regressions, we show that ECI has a non-linear and negative relationship with consumption inequality: higher economic complexity is associated with lower inequality.

As contributions to the literature, to the best of our knowledge, we are the first to use an extensive electronic payment database to examine consumption inequality. We are also the first to measure consumption inequality at the municipal level in Brazil. Furthermore, while other works focus on investigating the relationship between income inequality and economic complexity, we are the first to investigate the relationship between consumption inequality and economic complexity. We are also the first to investigate the relationship between inequality-ECI at the municipal level in Brazil.

In addition to the availability of the database used, Brazil presents an intriguing case for such a study, as it represents a significant number of economies that can be classified as lower-middle to upper-middle income while simultaneously experiencing high levels of poverty and inequality (Bandeira Morais et al., 2021). Additionally, Brazil is characterized by substantial regional economic disparities.

The paper is structured as follows. In Section 2, we provide a brief introduction to the literature. In Section 3, we describe the data sources used for constructing the electronic payment Consumption Inequality Index, the Economic Complexity Index, and the other variables employed in our analysis. In Section 4, we present the methodologies for calculating the Gini index and the Economic Complexity Index. In Section 5, we explore the Consumption Inequality Index derived from electronic payment methods and its correlation with the Gini Income Inequality Index as reported by IBGE. In Section 6, we present our study of the Inequality-ECI relationship. Section 7 concludes.

2 Literature review

2.1 Electronic Payment Methods Data and Consumption Inequality

Our work aims at obtaining a measure of inequality based on electronic payments made by individuals, which we consider as a proxy for consumption. While it is not a measure of income, which is the variable with better availability in advanced economies (Trapeznikova, 2019; Attanasio and Pistaferri, 2016), Aguiar and Bils (2015) show consumption inequality has closely tracked income inequality in the United States. Consumption also can be considered a better measure of well-being than income, considering that savings and loans can be used to smooth consumption over time (Hassett and Mathur, 2012; Meyer and Sullivan, 2009; Trapeznikova, 2019). Furthermore, the joint assessment of income inequality and consumption inequality can be interesting, for example, in enabling the investigation of consumption smoothing mechanisms and the nature of income shocks (temporary or permanent) (Attanasio and Pistaferri, 2016).

In the United States, research on consumption inequality typically relies on data from the Consumer Expenditure Survey (CE), a microdata source that has been available since the 1980s, or the Panel Study of Income Dynamics (PSID), which since 1999 has covered approximately 70% to 90% of the expenditures collected by the CE (Attanasio and Pistaferri, 2014, 2016). Another data source explored by researchers in the United States is the Residential Energy Consumption Survey (RECS), which enables the assessment of consumption inequality in durable goods (Hassett and Mathur, 2012). Similar data sources are available in other countries, such as the Chinese Residential Energy Consumption Survey (CRECS) used by Wu et al. (2017) to evaluate inequality in rural areas of China.

The literature using administrative data to study income inequality is well-established. Examples include Piketty and Saez (2003); Piketty et al. (2018); Larrimore et al. (2021). Regarding the use of administrative data to study consumption, studies are more scarce, and household consumption is generally determined indirectly. For example, Browning and Leth-Petersen (2003); Kolsrud et al. (2017); Eika et al. (2020) use extensive administrative databases of income, taxes, and wealth and the following accounting identity to determine household consumption: the total household spending is equal to income plus capital gains minus the change in wealth over a certain period.

Regarding the use of electronic payment data specifically, studies have used credit card data (Gross and Souleles, 2002; Aydin, 2015) and financial aggregator data[4] (Gelman et al., 2014, 2020; Baker and Yannelis, 2017; Baker, 2018; Olafsson and Pagel, 2018) to investigate consumption, although without assessing economic inequality. The database we used in our study has been explored by other authors in different contexts, such as by Gonçalves et al. (2022) in the area of nowcasting economic activity.

2.2 Economic Complexity Index

Since Kuznets (1955)[5], various works seek to establish a relationship between economic growth and inequality (Barro, 2008; Thomas, 2015; Galbraith, 2007; Palma, 2011; Deininger and Squire, 1996; Perera and Lee, 2013). However, conclusions appear to depend on theoretical preferences, the econometric methods employed, the economies under consideration, and the type of income distribution used (De Dominicis et al., 2008). Furthermore, economic growth may only reflect a portion of economic development (Hartmann et al., 2017; Le Caous and Huarng, 2020), and the determinants of inequality are broader, encompassing a range of economic, social, institutional, historical trajectories, technological changes, and rates of return to capital (Chu and Hoang, 2020; Hartmann and Pinheiro, 2022). In this context, new measures of economic development are needed to capture some of these factors, and Hidalgo and Hausmann (2009)'s Economic Complexity Index may be one such measure (Hartmann et al., 2017; Hartmann and Pinheiro, 2022).

[4]Web or mobile applications where users can link virtually any financial account, such as bank accounts and credit card accounts (Baker, 2018; Gelman et al., 2014, 2020).

[5]Kuznets (1955) suggested that economic development, measured by the income level of an economy, is related to income inequality through an inverted U-shaped curve. The hypothesis is that in the early stages of development, there would be an increase in inequality as a transition from a rural to an industrial structure occurs. Urban-rural inequality increases in a scenario where the productivity of the agricultural sector is lower than that of the industrial sector, while entrepreneurial wages grow more rapidly than those of workers in urban centers, as these wages are pushed downward due to the injection of cheap labor from rural areas. At a certain stage of development, there is a significant movement of part of the workforce into new, higher-paying sectors, and there is also an increase in agricultural sector productivity, along with institutional transformations such as democratization, redistribution policies, and the establishment of a welfare state, which exert pressure for reduced inequality (De Dominicis et al., 2008; Hartmann and Pinheiro, 2022; Sbardella et al., 2017; Soave et al., 2019).

Hausmann et al. (2014) define economic complexity based on the distribution and utilization of knowledge within a society. Products and services serve as means of transferring and integrating knowledge (Hidalgo and Hausmann, 2009). However, tacit knowledge, which is challenging to transfer, limits growth and development. The challenge of incorporating tacit knowledge leads to training in specific occupations and the specialization of organizations so that they can perform specific functions or tasks efficiently. Adapting to the expanding realm of knowledge involves distributing parts of that knowledge to individuals, and harnessing the diversity of this knowledge requires society to form organizations connected by intricate networks. In this way, the amount of productive knowledge utilized by an economy is mirrored in the diversity of firms, the range of necessary occupations, and the extent of interactions between them. The Economic Complexity Index (ECI) proposed by Hidalgo and Hausmann (2009) is a measure of how entwined this network of interactions is, i.e., how much productive knowledge is allocated by an economy (Hausmann et al., 2014).

The ECI measures the sophistication of a country, region, or municipality's productive structure by combining information about the diversity of products it exports or productive sectors it possesses and the ubiquity of these products or sectors (the number of countries, regions, or municipalities that export the product or possess a certain productive sector — Hidalgo, 2021). Complex economies are those with high diversity and export products or possess productive sectors with low ubiquity, meaning they are more exclusive (only a few diverse economies are capable of producing these products or possessing sophisticated sectors — Hartmann et al., 2017). Less complex economies are those capable of producing only a few products that are highly prevalent in the market. The Economic Complexity Index explores the interaction between entity diversity and product/sector ubiquity to measure the productive structure of an economy, incorporating information about the sophistication of its products/sectors[6].

2.3 Economic Complexity Index and Inequality

How would economic complexity affect inequality? Advocates of a negative relationship between the ECI and inequality argue that economies with a greater variety and sophistication of products or sectors tend to offer better occupational opportunities and upward mobility in social stratification, more inclusive institutions, a more equitable distribution of political power, and greater bargaining power for workers — forces capable of reducing economic inequality (Hartmann et al., 2017; Hartmann, 2014; Constantine and Khemraj, 2019). Arif (2021) demonstrates that an economy's sophistication leads to a higher labor share, which serves as a mechanism for inequality reduction. Given the need for physical capital, human capital, and technological advancements, the production of complex products results in increased demand for skilled workers, labor productivity, and wages proportional to labor efficiency. As labor inherently embodies tacit productive knowledge, workers' bargaining power increases, subsequently allowing them to secure better wages and thereby increase their share in total product.

Furthermore, as an economy becomes more sophisticated, a broader spectrum of densely in-

[6]Products are distinguished by the amount of resources required for their production. The more diverse capabilities needed to manufacture a good, the more sophisticated it is considered. This complexity also applies to the economy as a whole, which becomes more advanced as it produces a greater variety of sophisticated goods. In other words, product sophistication and economic complexity are driven by the variety and quantity of capabilities available in a given locality or country (Hausmann and Hidalgo, 2011).

terconnected products promotes an increase in productive interactions across various sectors. This demands a more diversified labor force with broader skills and multiple levels of expertise. From a flatter occupational structure characterized by a higher number of job positions and learning opportunities, less specialized workers may gain advantages over more specialized workers, contributing to a reduction in inequality (Pham et al., 2023). More diversified economies would also ensure better long-term business sustainability in the face of volatility or crises, maintaining employability and wages at all levels, thus preventing an increase in inequality (Chu and Hoang, 2020).

On the other hand, in less sophisticated and diversified economies that heavily rely on natural resources, the income of the majority of workers depends on economic activities with diminishing returns to scale and low productivity. These individuals also face learning constraints and occupational limitations. Only a small portion of the population ends up enjoying higher income from more productive (yet limited) activities, as well as the knowledge and skills that remain confined within these groups (Lee and Vu, 2020).

Hartmann et al. (2017) are the first to study the relationship between economic complexity and inequality, providing support for a negative relationship between ECI and income inequality. They utilize economic complexity indices from the Observatory of Economic Complexity[7] at MIT to explain inequality, measured by the Gini index, for over 70 countries in a cross-sectional regression. They control for per capita GDP (and its square), education, population, and variables representing country institutions. In all models tested, the Economic Complexity Index was a negative and significant predictor of inequality. This result holds when they perform a fixed-effects panel estimation using data from 1962 to 2012: economic complexity reduces inequality. Lee and Vu (2020) arrive at a similar result when conducting cross-sectional OLS regression estimates for 96 countries, using data averages from 1980 to 2014 and similar controls to those used by Hartmann et al. (2017).

There are also hypotheses on how increased complexity could lead to greater inequality. Greater complexity creates a higher demand for skilled workers as new sectors emerge, replacing or rendering traditional sectors obsolete. While retraining is possible for low-income or low-skilled workers, it is easier and less costly for skilled workers to advance, as they have a greater capacity to adapt to changes, thereby widening income inequality (Lee and Vu, 2020; Chu and Hoang, 2020; Violante, 2008; Pham et al., 2023). Automation can also play a significant role, making medium or low-skilled jobs obsolete (Sebastian and Biagi, 2018). A process of "deindustrialization" may also occur as the economy becomes more complex: it specializes in sophisticated products and replaces in-house manufacturing with imports for resource-intensive or medium- to low-skilled labor-intensive products. Thus, with part of the workforce unable to qualify for higher-skilled jobs, they end up being reallocated to lower-income positions in other sectors, thereby increasing inequality (Pham et al., 2023; Violante, 2008).

Lee and Vu (2020) find a positive relationship between economic complexity and inequality when estimating the relationship using system-GMM. Chu and Hoang (2020) also find a positive relationship between economic complexity and inequality using 2SLS and system-GMM for 88 countries from 2002 to 2017. However, when interacting the ECI with socioeconomic variables, the authors show that, at certain levels of education, more efficient government spending, and trade openness, increased complexity can act to reduce inequality. Lee and Wang (2021) also find a positive relationship between economic complexity and inequality using fixed-effects panel

[7] `atlas.media.mit.edu`

strategies for their complete sample of 43 countries from 1991 to 2016. However, when dividing the sample into two subgroups, high-income countries and others, they demonstrate economic complexity reduces inequality in the former group and increases it in the latter.

Given the favorable and unfavorable hypotheses about the impact of economic complexity on reducing inequality, other studies find nonlinear relationships between these two variables. Pham et al. (2023) identify a U-shaped relationship between inequality and complexity in a system-GMM estimation for 99 countries from 2002 to 2016: initially, an increase in complexity would reduce inequality, but with an inversion beyond a certain level of complexity. The authors' hypothesis for the increase in inequality beyond a certain complexity threshold is the effect of deindustrialization (destruction of low and medium complexity jobs in manufacturing). This conclusion contradicts that of Amarante et al. (2023), who, in a fixed-effects panel estimation for 126 countries from 1995 to 2018, find an inverted U-shaped relationship: when economic complexity is low, increases in the sophistication of an economy's productive structure would increase inequality, and beyond a certain threshold, an increase in complexity would reduce inequality. The authors note it is possible that the negative relationship between complexity and inequality evidenced in previous studies may be driven by the group of high-income countries that have already reached this threshold.

In general, disparities among the results of empirical studies may depend on the country sample, the periods considered, and the estimation methods employed. Thus, at the country level, the question of the relationship between economic complexity and inequality remains open.

2.4 Economic Complexity Index and Regional Income Inequality

Does the relationship between economic complexity and inequality at the regional level follow the same dynamics as observed at the country level? Sbardella et al. (2017) assess the relationship between economic complexity and wage inequality among countries and also among counties within the United States. They use an alternative measure of Economic Complexity, called Fitness, developed by Tacchella et al. (2012). To construct the economic complexity index at the county level, they use employment data by economic activity sector instead of exported products. They find a positive relationship between economic complexity and wage inequality for U.S. counties in a cross-sectional assessment, in contrast to an inverted U-shaped curve when assessing countries, showing the relationship is not scale-invariant. According to the authors, the fact that institutions are relatively homogeneous in the United States explains the difference in the complexity-inequality relationship between the two approaches.

In their study, Gao and Zhou (2018) used data from 2690 firms to calculate the ECI for 31 Chinese provinces, based on a "Province-Industry" network[8], where the number of firms in each of 70 categories for each province is considered instead of exported products. In a bivariate analysis, they found a negative relationship between ECI and income inequality for the provinces analyzed. Török et al. (2022), in cross-sectional and fixed-effects panel estimations for counties in Romania, using data from 2008 to 2018, also found a negative relationship between ECI and inequality.

In their study, Bandeira Morais et al. (2021) analyze the relationship between ECI and inequality using panel data for Brazilian states, employing Pooled OLS and Random Effects, with data spanning from 2002 to 2014. They find an inverted U-shaped relationship. One hypothesis for this relationship is that in the early stages, an increase in economic complexity benefits capital owners and high-skilled

[8]As will be seen in section 4.2, for regional ECI calculation, each element of the matrix $\mathbf{M}$, $M_{p,i}$, receives the value of 1 if province p has revealed comparative advantage in industry i and 0 otherwise.

workers, leading to an increase in inequality. Beyond a certain level of economic complexity, other components of economic complexity become more important, such as institutions, labor unions, job opportunities, among others, which act to reduce inequality.

3 Data

3.1 Consumption Inequality

To measure consumption inequality, we utilized electronic payment data from the Brazilian Central Bank's Payment System[9], specifically data from the Pix and Credit Card payment instruments. The Instant Payment System, launched in November 2020, enables 24/7 settlement of instant payments (Pix) and has spurred the creation of numerous new financial applications, such as QR code-payable invoices. The credit card data refer to the outstanding balance of individuals over the month, which includes spot purchases made during the month, plus installments of on-credit purchases due in the period[10].

Credit card payments are already used as a proxy for consumption, while Pix has become a standard for spot purchases in Brazil. For this reason, we believe that by using payments made through both instruments, we construct a good proxy for consumption[11].

We extracted payment data over the year 2022 for each payment instrument at the individual level. We excluded all payments sent by companies since we want to measure the inequality of people (and not companies). Additionally, we excluded transactions made from and to the same person since they represent a simple fund transfer between financial institutions. We also excluded Pix payments made to institutions belonging to activity groups that we consider unlikely to receive transactions related to consumption. The list of CNAE codes of these institutions is in C. Next, we aggregated payments sent to other individuals (outgoing payments) since it reflects a consumption intent, as opposed to payments received from other individuals (income payments). Finally, we divided the values by 12 to obtain a consumption monthly average for each individual. Using data from the Brazilian Federal Revenue to determine the municipality of residence for each, we calculated the Gini inequality index of consumption as described in Section 4.1.

3.2 Economic Complexity Index

To calculate the municipal Economic Complexity Index (ECI), we used employment data by municipality, classified according to the National Classification of Economic Activities (CNAE) 2.0, as

[9] A Payment System is a set of instruments, rules, procedures, and technologies used to settle money transfers between economic agents (Aprigliano et al., 2019).

[10] It's important to note these figures may not exactly match an individual's monthly credit card statement because the due date for the statement may differ, and payments may occur in the same month or the following month. As a result, within the population, some captured values will be higher than the actual monthly statement amount, while others will be lower. On average, though, we consider these data a reasonable representation of credit card consumption.

[11] In A, we show the correlation between the average annual consumption calculated based on electronic payment methods, base year 2021, and the average annual income estimated by Neri and Hecksher (2023), base year 2020, by municipality. We found a correlation that can be considered moderate to high (0.75), which would be expected if we have a good proxy for consumption. It is important to note, however, that the income calculated by Neri and Hecksher (2023) only considers tax return data, meaning it does not capture information from those who do not file tax returns, which is a considerable portion of the population.

presented in the Annual Social Information List (RAIS). The data were obtained from the Ministry of Labor and Employment[12].

3.3 Control Variables of the Application

As regression controls for assessing the relationship between consumption inequality and the ECI, we obtained the GDP per capita (base year 2021) and the population of each municipality (base year 2022) from the IBGE, as well as the distance from each municipality to the nearest hub city[13]. As a measure of human capital, we used a proxy for the quality of education based on the age-grade distortion rate, i.e., the percentage of students in the correct grade for their age in high school. These data are obtained from the Institute of Applied Economic Research (IPEA), base year 2021. As a measure related to institutions, we created a variable with the percentage of coverage of legislation and planning instruments for each municipality[14], based on information provided by the IBGE for Brazilian municipalities, for the year 2021[15].

4 Methods

4.1 Calculation of the Inequality Index

To measure consumption inequality in each municipality using electronic payment methods, we employ the Gini index[16]. Let the population percentages be ordered from the poorest (or lowest consumption) to the richest (or highest consumption) on the horizontal axis of the graph in Figure 1, and on the vertical axis, the cumulative proportion of income (or consumption) held by the population. Let the line of perfect equality be the diagonal line where each percentage of the population has an equal share of income or consumption, and the Lorenz curve be the actual income or consumption accumulation curve. The Gini index is given by the ratio of the area between the line of perfect equality and the Lorenz curve (area A in Figure 1) to the total area below the line of perfect equality (area $A + B$ in Figure 1).

To compute the Gini index in a direct and efficient manner (Cowell, 2011), we order all incomes or consumptions from lowest to highest, $y_{(1)}, y_{(2)}, ..., y_{(n)}$ (meaning $y_{(1)}$ is the lowest income/consumption, $y_{(2)}$ is the next, and so on, up to person n), and we calculate:

$$G = \frac{2}{n^2\overline{y}} \left[y_{(1)} + 2y_{(2)} + 3y_{(3)} + ... + ny_{(n)}\right] - \frac{(n+1)}{n} \tag{1}$$

[12] Available at http://pdet.mte.gov.br/acesso-online-as-bases-de-dados.

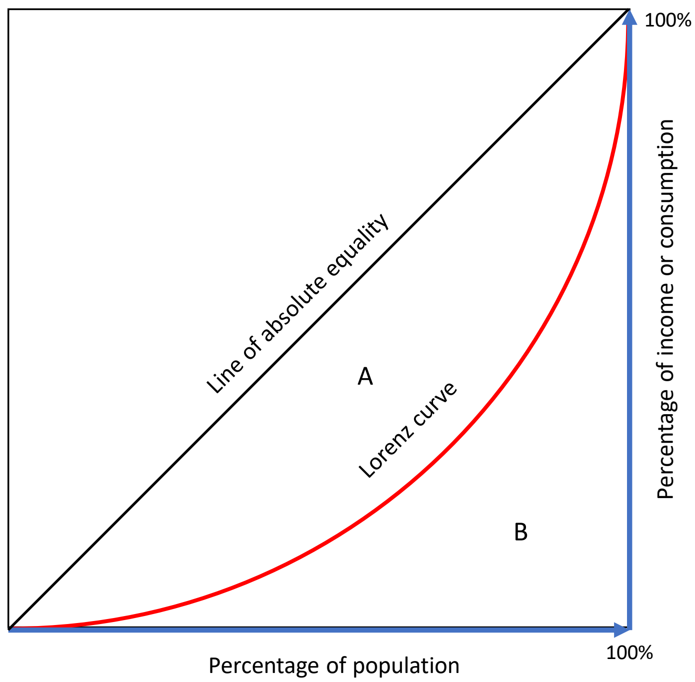
[13] For hub city, we refer to municipalities classified as metropolises or regional capitals by the IBGE. The list of these municipalities is available at https://www.ibge.gov.br/geociencias/cartas-e-mapas/redes-geograficas/15798-regioes-de-influencia-das-cidades.html.

[14] Quantity of legislation and planning instruments existing in relation to the total expected by the IBGE Basic Municipal Information Survey, listed in D.

[15] Given the limitation of available data, we faced difficulties in creating measures of institutions for Brazilian municipalities. The measure we propose here is an attempt to partially measure Formal Institutions based on the concept from Pereira et al. (2016). However, our measure presents a number of limitations, as we have outlined in the section 6.4.

[16] For the origin of the Gini index, see Ceriani and Verme (2012).

Figure 1: **Lorenz curve.** On the horizontal axis, the population percentages are ordered from the poorest (or lowest consumption) to the richest (or highest consumption). The vertical axis shows the cumulative proportion of income (or consumption) held by the population. The diagonal line represents perfect equality, where each percentage of the population has an equal share of income or consumption. The Lorenz curve represents the actual distribution of income or consumption. The Gini index is calculated as the ratio of the area between the line of perfect equality and the Lorenz curve (area A) to the total area below the line of perfect equality (area $A + B$).



where:

$$\overline{y} = \sum_{i=1}^{n} \frac{y_{(i)}}{n} \tag{2}$$

This Gini index calculation method was implemented in SQL, directly in the database query, due to efficiency and computational cost considerations.

4.2 Calculation of the Economic Complexity Index

Following Hidalgo and Hausmann (2009), Hausmann et al. (2014), and Kemp-Benedict (2014), we define the Revealed Comparative Advantage (RCA) of a country c in a product p as:

$$RCA_{cp} = \frac{X_{cp} / \sum_{p'} X_{cp'}}{\sum_{c'} X_{c'p} / \sum_{c'p'} X_{c'p'}} \tag{3}$$

in which X_{cp} represents the total exports of product p from country c, $\sum_{p'} X_{cp'}$ denotes the total export portfolio of country c, $\sum_{c'} X_{c'p}$ signifies the total exports of product p by all countries, and

$\sum_{c'p'} X_{c'p'}$ represents the total exports of all products by all countries. RCA_{cp} will be greater than 1 if the export of product p is higher than expected given the size of country c's export economy and the global market for the product, indicating the country has a comparative advantage in the product (Hartmann et al., 2017).

Let $\mathbf{M}$ be a country-product matrix with elements M_{cp} indexed by country c and product p. We define each element of M_{cp} as:

$$M_{cp} = 1 \; if \; RCA_{cp} >= 1 \tag{4}$$

$$M_{cp} = 0 \; if \; RCA_{cp} < 1 \tag{5}$$

We then define the diversity of a country c and the ubiquity of a product p as:

$$Diversity = k_{c,0} = \sum_p M_{cp} \tag{6}$$

$$Ubiquity = k_{p,0} = \sum_c M_{cp} \tag{7}$$

That is, we measure diversity and ubiquity by summing over the rows or columns of the matrix $\mathbf{M}$ while fixing the country or product, respectively. However, these indicators separately are imprecise measures of economic complexity. For countries, we need to calculate the average ubiquity of the products they export, the average diversity of the countries producing these products, and so on. For products, we must calculate the average diversity of the countries producing them, the average ubiquity of the other products these countries produce, and so forth[17]. Therefore, the economic complexity proposed by Hidalgo and Hausmann (2009) jointly and interactively computes the average values of diversity and ubiquity using the so-called method of reflections, where:

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_p M_{cp} \cdot k_{p,N-1} \tag{8}$$

$$k_{p,N} = \frac{1}{k_{p,0}} \sum_c M_{cp} \cdot k_{c,N-1} \tag{9}$$

Substituting 9 into 8, we arrive at:

$$k_{c,N} = \sum_{c'} \left(\frac{1}{k_{c,0}} \sum_p M_{cp} \frac{1}{k_{p,0}} M_{c'p} \right) k_{c',N-2} \tag{10}$$

which can be written in matrix form as:

[17]For example, diamonds have low ubiquity but are generally produced by countries with low diversification, indicating a low requirement for productive knowledge. On the other hand, medical imaging devices have low ubiquity and are generally produced by countries with high diversification, indicating a high requirement for productive knowledge. Thus, a correction is necessary to ensure uncommon products are considered truly complex only if they are produced by diversified countries. Similarly, relatively common products can also be considered complex if their production is limited to a group of diversified countries (Sousa, 2018).

Table 1: **Summary statistics of municipal consumption inequality Gini coefficients.** This table presents the summary statistics for the Gini coefficients of consumption inequality across municipalities. The columns include the number of observations (**N**), the minimum value (**Min**), the first quartile (q_1), the median value (**Median**), the mean value (**Mean**), the third quartile (q_3), the maximum value (**Max**), and the standard deviation (**St. Dev**).

	N	**Min**	q_1	**Median**	**Mean**	q_3	**Max**	**St. Dev**
Gini Coefficients	5563	0.591	0.689	0.716	0.715	0.739	0.924	0.037

$$\vec{k}_N = \mathbf{W}\vec{k}_{N-2} \tag{11}$$

where $\vec{k}_N$ represents the set of values for countries $k_{c,N}$, and the matrix $\mathbf{W}$ has elements:

$$W_{cc'} = \frac{1}{k_{c,0}} \sum_p M_{cp} \frac{1}{k_{p,0}} M_{c'p} \tag{12}$$

The Economic Complexity Index for the set of countries is then defined as the eigenvector of $\mathbf{W}$ associated with the second largest eigenvalue (Kemp-Benedict, 2014; Hausmann et al., 2014).

In our work, we adopt the methodology adapted by Brandão (2023) and Sbardella et al. (2017) from Hidalgo and Hausmann (2009) to analyze municipalities instead of countries. Given the geographical focus of our study, we chose to use the number of employment links as a proxy for the size of each of the 285 activity groups in the CNAE 2.0, rather than relying on product export data. In this context, the number of employment links in each sector within a municipality serves a role analogous to the export value of products for each country in Hidalgo and Hausmann (2009)'s original model. Therefore, the economic complexity of a municipality is assessed based on the diversity and ubiquity of its productive sectors[18]. We used the R package *economiccomplexity* to calculate the ECI for each municipality.

5 Consumption Inequality Gini Index

Figure 2 shows the distribution of the Gini inequality index calculated using electronic payment methods across Brazilian municipalities. Table 1 displays the descriptive statistics.

Figure 3 displays the regional boxplots: the northern region has the highest municipal average

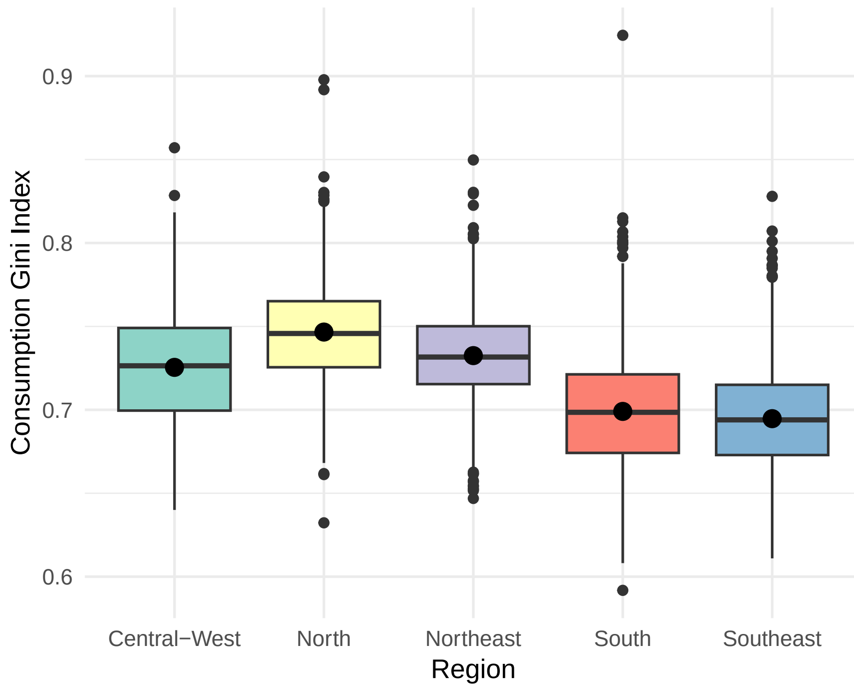
[18]According to Brandão (2023), the use of labor market data for regional ECI analysis is more appropriate because:

1. Labor market data encompass a broader range of economic activities than a municipality's export portfolio, including the services sector;
2. International trade data cover only external trade and do not account for the trade of products within municipalities themselves;
3. Due to bureaucratic or administrative reasons, a product's origin may be in one city but is often recorded in international trade data under another city.

Figure 2: **Consumption inequality Gini coefficients histogram.** This histogram displays the Gini coefficients of consumption inequality for the municipalities. The horizontal axis represents the Gini coefficient values, while the vertical axis shows the frequency of municipalities for each Gini coefficient range. The distribution of Gini coefficients approximates a normal distribution, indicating most municipalities have Gini coefficients around the central value with fewer municipalities having extreme values.

of the consumption Gini index, while the South and Southeast regions are similar and have the lowest averages. This ranking is similar to the one found in the evaluation of the Gini index boxplots calculated from income according to the 2010 IBGE census, shown in Figure 4. In B, we present the boxplots by state, both for our consumption inequality index and for the one calculated based on data from the 2010 IBGE census.

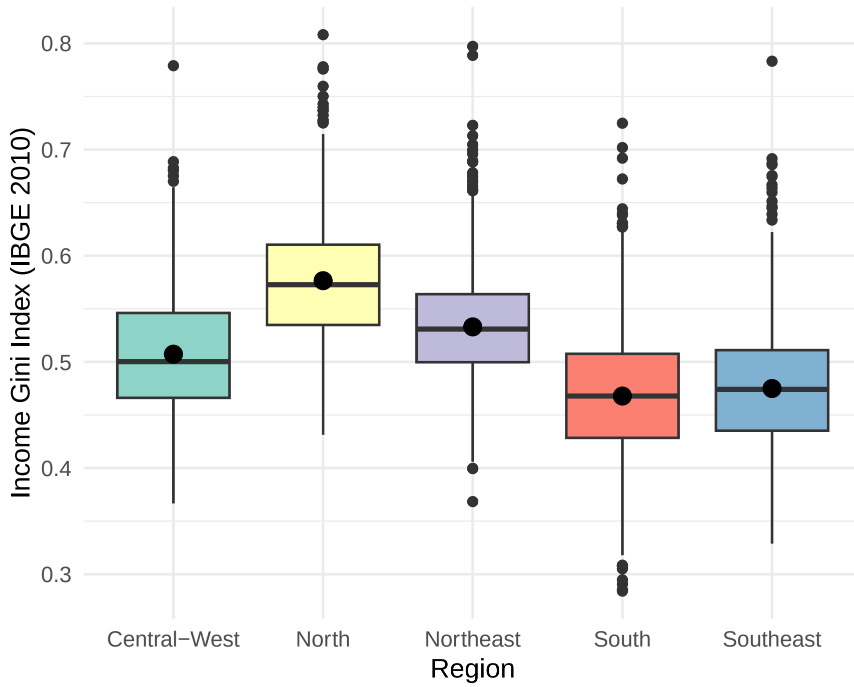
Figure 3: **Boxplots of municipal consumption Gini coefficients.** This figure displays boxplots for the consumption Gini coefficients across different regions. A boxplot is a graphical representation that shows the distribution of a dataset. Each boxplot displays the median (central line), the first quartile (bottom of the box), the third quartile (top of the box), and the potential outliers. In this figure, the northern region has the highest municipal average of the consumption Gini index, while the South and Southeast regions have the lowest averages. This regional ranking is similar to the one found in the evaluation of the Gini index boxplots calculated from income according to the 2010 IBGE census, shown in Figure 4.



Although our index is calculated based on data we consider a proxy for consumption, we expect it to be correlated with the income inequality measured by the 2010 IBGE census data, even with a lag of over a decade between the databases, given that inequality tends to show a certain persistence. Figure 5 shows the Pearson correlation between both inequality indices. There is also a distinction of municipalities by region, where we can see the patterns established in Figures 3 and 4: municipalities from the southern and southeastern regions are concentrated in the lower left corner of the chart, indicating lower inequality on average, while municipalities from the northern and northeastern regions are mostly in the upper right corner, indicating higher inequality.

We might find slightly higher correlations between the two indices if we exclude very small municipalities. Figure 6 shows the correlations for cases where we filter the municipalities by a minimum population threshold. For instance, if we consider only municipalities with more

Figure 4: **Boxplots of municipal IBGE 2010 income Gini coefficients.** This figure shows boxplots for the income Gini coefficients based on data from the 2010 IBGE census, highlighting variability among regions. As in the previous figure, the boxplots illustrate the distribution of Gini coefficients with the median, quartiles, and potential outliers. The northern region exhibits the highest municipal average of the Gini index, while the South and Southeast regions have the lowest averages. This pattern is consistent with the consumption Gini index shown in Figure 3.



than 30,000 inhabitants, the Pearson correlation between the indices could be higher than 0.55. However, as most Brazilian municipalities are small, if we exclude municipalities with less than 30,000 inhabitants, we reduce the number of municipalities considered by more than 3/4.

We also note the IBGE Gini index generally indicates less inequality for Brazilian municipalities. Our hypothesis regarding the difference in inequality is that, in the distribution of Electronic Payment transactions, the sum of transactions per person at the tail end of the poorest is lower, on average, than the household incomes of the poorest captured by the IBGE. It is possible that this population conducts more cash transactions than the higher-income population, which we do not capture in our database. Additionally, it is likely that higher incomes are under-reported in household surveys, but do not escape electronic transactions. Another factor is that the IBGE's Gini index is calculated based on per capita household income, which sums and then divides the income of all residents by the number of residents in each household. This likely eliminates some extremes in the income distribution, compared to our data distribution, where we consider the payments made by each individual, regardless of their residence. Given these factors, our index tends to show greater inequality than the Gini index calculated by the IBGE. Conversely, we have an inequality index that can be calculated promptly and as frequently as desired.

Figure 5: **Scatter plot comparing consumption and IBGE income Gini coefficients.** This figure shows a scatter plot comparing the Gini coefficients of consumption inequality (our index) with the Gini coefficients of income inequality from the 2010 IBGE census. The Pearson correlation coefficient (R) between the two indices is displayed. Each point represents a municipality, and different colors indicate different regions. Municipalities from the southern and southeastern regions are concentrated in the lower left corner of the chart, indicating lower inequality on average. In contrast, municipalities from the northern and northeastern regions are mostly in the upper right corner, indicating higher inequality.

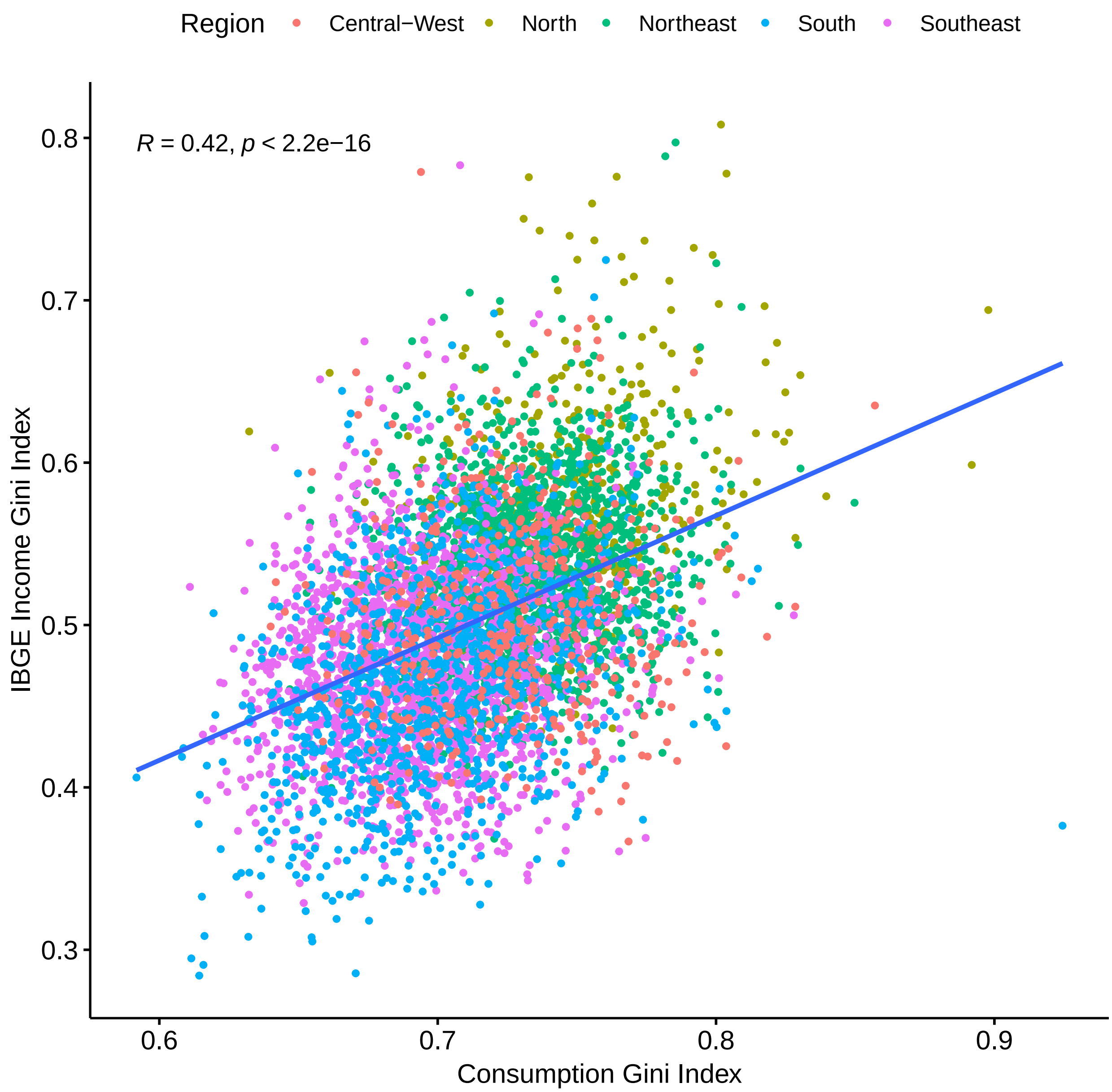
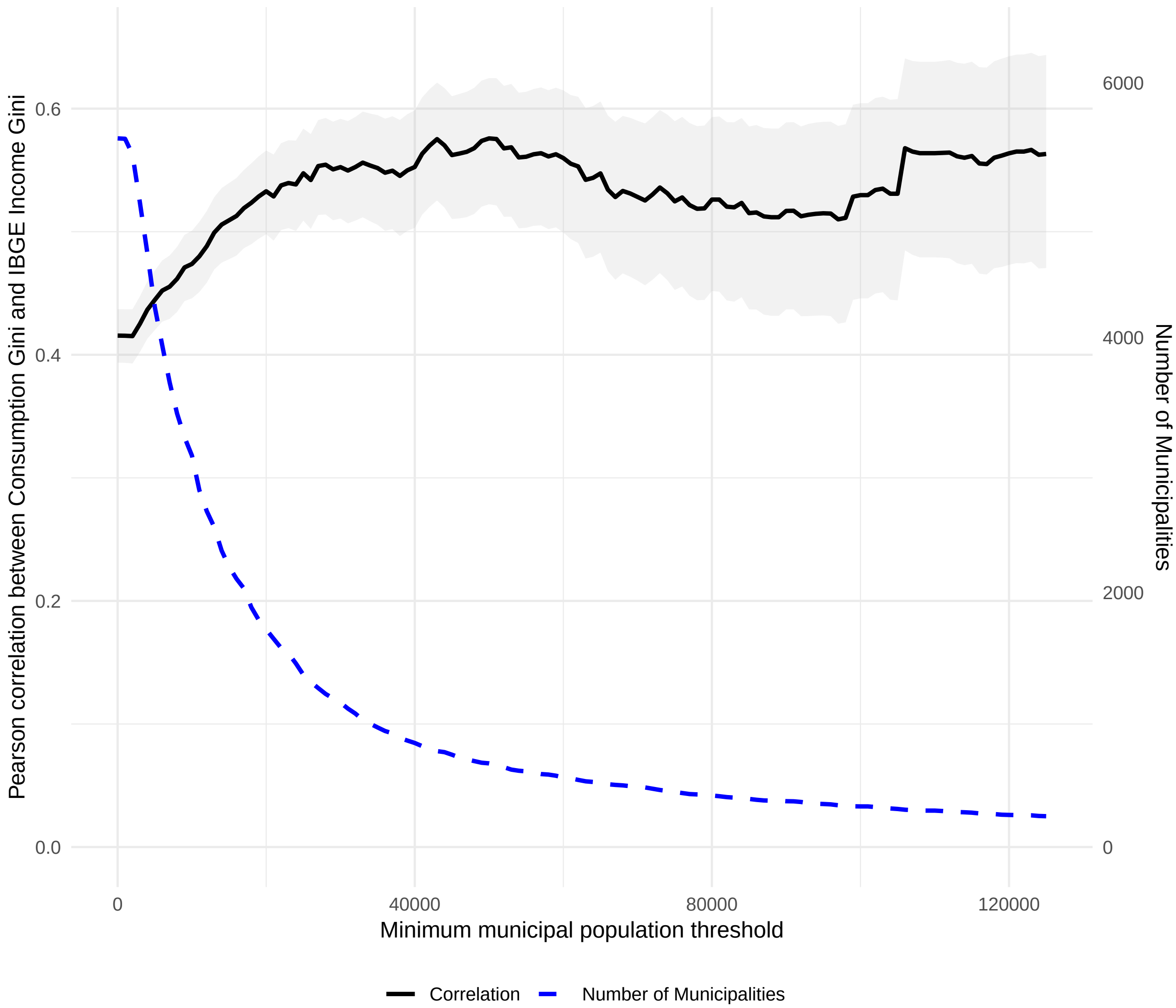


Figure 6: **Evolution of the Pearson Correlation between consumption and IBGE income Gini coefficients as a function of an increasing sample's minimum municipal population threshold.** This figure shows the evolution of the Pearson correlation between the consumption Gini coefficients (our index) and the income Gini coefficients from the 2010 IBGE census as a function of an increasing minimum municipal population threshold. The solid line represents the Pearson correlation coefficient with confidence intervals. The dashed blue line represents the number of municipalities considered in the sample for a given minimum municipal population threshold, as indicated by the secondary y-axis. This figure illustrates that higher correlations can be observed when very small municipalities are excluded from the sample. For instance, considering only municipalities with more than 30,000 inhabitants, the Pearson correlation between the indices can be higher than 0.55. However, excluding municipalities with fewer than 30,000 inhabitants reduces the number of municipalities considered by more than three-quarters.



6 Application - Inequality vs. Economic Complexity Index

In this section, we explore an application for our Gini index of consumption inequality based on electronic payment methods: similarly to Hartmann et al. (2017), we assess the relationship between inequality (in our case, consumption inequality) and the Economic Complexity Index at the municipal level.

6.1 Model

Figure 7 shows the scatter plot depicting the relationship between our Gini index of consumption inequality and the Economic Complexity Index for Brazilian municipalities. The relationship does not appear to be completely linear at first glance, with an initial indication that higher economic complexity might be associated with lower inequality. We can also observe larger municipalities tend to have a higher ECI, on average.

We then proceed to a multivariate analysis using cross-sectional data for Brazilian municipalities, where we estimate the coefficients of Equation 13 using Ordinary Least Squares (OLS):

$$Gini_i = \beta_0 + \beta_1 ECI_i + \beta_2 ECI\_sq_i + \boldsymbol{\beta}_3 \boldsymbol{X}_i + \varepsilon_i \tag{13}$$

For each municipality i, $Gini_i$ is the Gini index of consumption inequality calculated based on electronic payment methods; ECI_i the Economic Complexity Index and ECI_sq_i its square, aiming at capturing a possible nonlinear relationship, as in Pham et al. (2023), Amarante et al. (2023), and Bandeira Morais et al. (2021); $\boldsymbol{X}_i$ is a vector of control variables, including the natural logarithm of per capita GDP and its square; the natural logarithm of the population; the percentage of young population, i.e., the percentage of the population up to 19 years old[19]; an education measure which is the rate of students in the correct grade for their age; the natural logarithm of the distance from the municipality to the nearest hub city[20]; as an institutional measure the percentage of legislation implemented relative to that planned in the IBGE survey on Legislation and Planning Instruments.

We estimate Equation 13 with and without the squared term ECI_i for all Brazilian municipalities included in our database[21]. We also estimate Equation 13 for each Brazilian region, with the aim of verifying whether the results remain consistent in each region.

6.2 Results

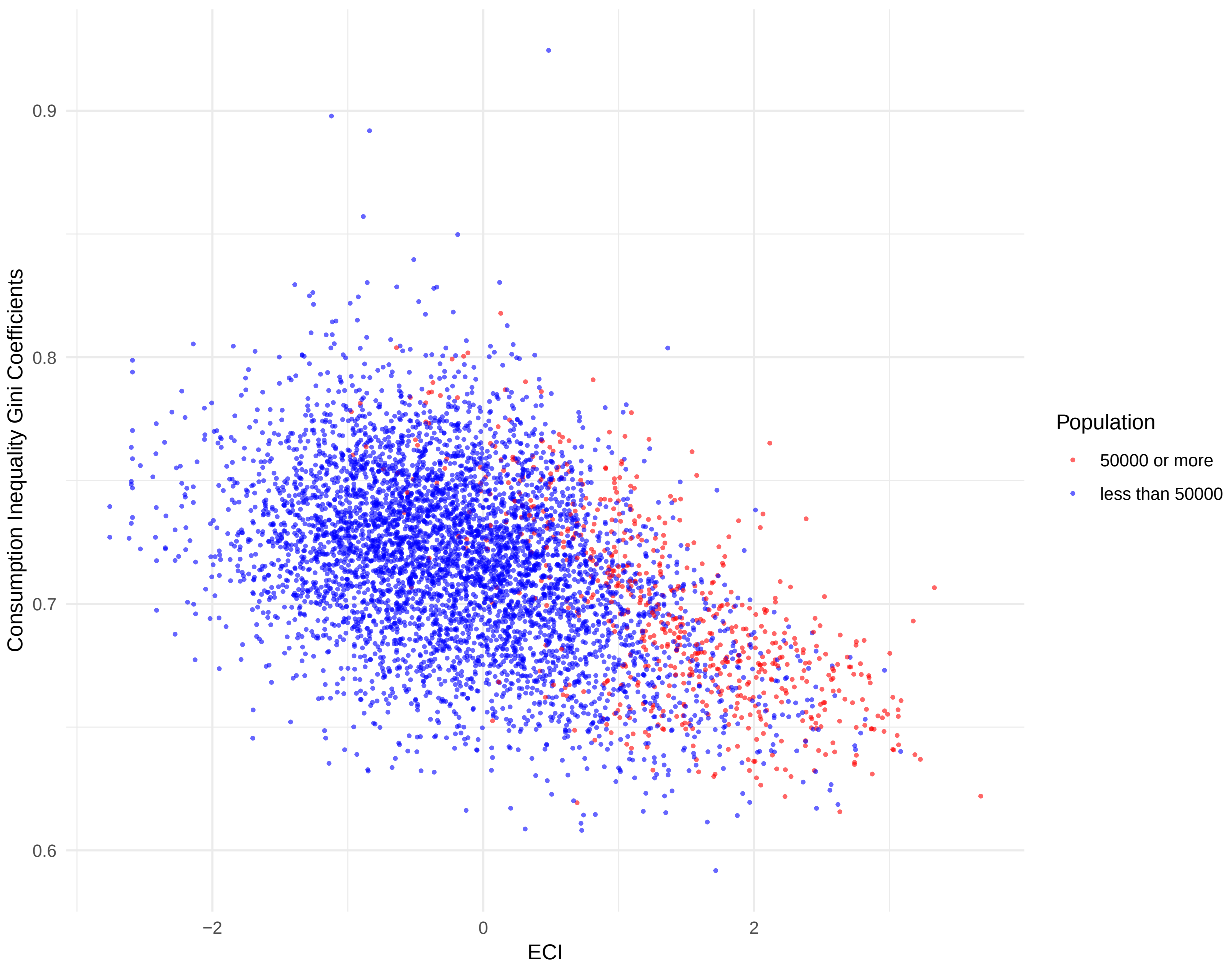
Table 2 shows the regressions results for specifications with and without the squared ECI term. We can observe the ECI coefficient is negative and statistically significant at 1% in both the first and second specifications, and its squared term is also negative and statistically significant at 1% in the second specification. This indicates a negative and nonlinear relationship between consumption

[19]This control aims at accounting for the usage of Pix, which tends to be higher among the younger population. As the younger population generally has lower income and wealth compared to the older population, a higher percentage of young population might indicate higher poverty, potentially influencing our inequality measure.

[20]For municipalities that are hubs, i.e., with zero distance, a distance of “1” Km was imputed to enable the use of the natural logarithm.

[21]We exclude municipalities that have missing data for any of the considered variables.

Figure 7: **Scatter plot of the relationship between the consumption inequality Gini index and the Economic Complexity Index.** This figure displays a scatter plot illustrating the relationship between the Gini index of consumption inequality (our index) and the Economic Complexity Index (ECI) for Brazilian municipalities. Each point represents a municipality. The horizontal axis represents the ECI values, while the vertical axis represents the Gini coefficients of consumption inequality. The scatter plot suggests the relationship is not completely linear; however, there is an indication that higher economic complexity might be associated with lower inequality. Additionally, it can be observed that larger municipalities tend to have a higher ECI, on average.



inequality and the ECI, meaning higher economic complexity is associated with lower inequality, and the greater the economic complexity, the larger the impact of ECI changes on inequality.

The negative relationship between inequality and the ECI is in line with the findings of Hartmann et al. (2017) and Lee and Vu (2020) — in OLS regressions — at the country level, and at the regional level, it aligns with the results of Gao and Zhou (2018) and Török et al. (2022), and is contrary to that of Sbardella et al. (2017).

If the relationship between economic inequality and the Economic Complexity Index is the same for both income inequality and consumption inequality, our result does not support the hypothesis of Hartmann and Pinheiro (2022) regarding the reversal of the relationship, from negative to positive, when scaling down the assessment to the regional level. Sbardella et al. (2017) posit that the difference in results when assessing the relationship at the country level versus at the municipal level is due to the fact that institutions are relatively homogeneous within the United States. It might be assumed that within Brazil, institutions are not as homogeneous as in the United States. However, this does not seem to explain the difference in the sign of the ECI coefficient between the studies.

We also do not support the finding of Bandeira Morais et al. (2021): although we find a nonlinear relationship between inequality and ECI, it is not in an inverted U-shape. Thus, we do not support the hypothesis that municipalities with low economic complexity would experience an increase in inequality at the initial levels of increased complexity.

We can assume that some of the hypotheses applicable to a negative relationship between ECI and inequality at the country level might also be applied at the municipal level: greater variety and sophistication of sectors would tend to offer better occupational opportunities, more inclusive institutions, more equitable distribution of power, greater bargaining power for workers, or a higher share of labor in the total product.

Regarding the control variables, as we can see in Table 2, we do not find empirical support for the Kuznets curve when evaluating the coefficients related to GDP per capita. Contrary to Lee and Vu (2020); Hartmann et al. (2017), the coefficient of per capita GDP is negative and statistically significant at 1% for the first specification, and the squared term is also negative and statistically significant at 1%.

Table 2 also shows a larger population is associated with greater inequality, as found in Hartmann et al. (2017), and a greater distance to a hub city is also related to higher inequality. Our education measure is negatively related to inequality, while our institutional measure is negatively related in the second specification. Additionally, when we include the control variable representing the percentage of the young population, we find a higher percentage of young population is associated with greater consumption inequality.

6.3 Robustness check

As a robustness check, we verified whether the main results remain invariant across different regions of the country. Table 3 displays specific regressions for each Brazilian region. The coefficients of ECI and its square are negative and statistically significant at the 1% level across all regressions, indicating a consistent relationship where higher complexity is associated with lower consumption inequality in all Brazilian regions.

The coefficients of control variables seem to depend on the regions analyzed. Population is a predictor of inequality in the South, Central-West, Northeast, and North regions, but not in the

Table 2: OLS regressions for all municipalities.

	Dependent variable:	
	Consumption Inequality Gini Index	
	(1)	(2)
ECI	−0.013***	−0.012***
	(0.001)	(0.001)
ECI^2		−0.004***
		(0.0003)
ln(GDP pc)	−0.014***	−0.005
	(0.004)	(0.005)
$\ln(\text{GDP pc})^2$	−0.001**	−0.0003
	(0.001)	(0.001)
ln(population)	0.004***	0.005***
	(0.001)	(0.001)
pct young pop	0.001***	0.002***
	(0.0001)	(0.0001)
ln(hub distance)	0.011***	0.010***
	(0.001)	(0.001)
adr high school	−0.009***	−0.007***
	(0.003)	(0.003)
n legis	0.0004	-0.003*
	(0.001)	(0.001)
Constant	0.713***	0.655***
	(0.023)	(0.027)
Observations	5,554	5,554
R^2	0.374	0.397
Adjusted R^2	0.373	0.396
F Statistic	413.339*** (df = 8; 5545)	406.049*** (df = 9; 5544)

ECI refers to the Economic Complexity Index, *GDP pc* to the municipal Gross Domestic Product per capita, *pct young pop* to the percentage of the population aged 19 and under, *hub distance* to the distance to the nearest hub municipality, *adr high school* to the age-grade distortion rate (the percentage of students in the correct grade for their age in high school), and *n legis* to an institutional measure - the quantity of legislation and planning instruments existing in relation to the total expected by the IBGE Basic Municipal Information Survey. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 3: OLS regressions for all municipalities by region.

	Dependent variable:				
	Gini				
	South	Southeast	Central-West	Northeast	North
ECI	−0.018***	−0.006***	−0.011***	−0.010***	−0.016***
	(0.002)	(0.001)	(0.004)	(0.001)	(0.004)
ECI^2	−0.005***	−0.001***	−0.009***	−0.005***	−0.006***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
ln(GDP pc)	−0.068***	−0.021***	0.001	−0.007	−0.008
	(0.014)	(0.007)	(0.020)	(0.014)	(0.013)
$\ln(\text{GDP pc})^2$	−0.013***	−0.002**	−0.0003	0.002	0.0001
	(0.002)	(0.001)	(0.003)	(0.002)	(0.002)
ln(population)	0.005***	−0.0003	0.006***	0.008***	0.009***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
pct young pop	0.002***	0.0004	−0.002***	0.001***	0.0005
	(0.0003)	(0.0003)	(0.001)	(0.0002)	(0.0003)
ln(hub distance)	0.010***	0.013***	0.005*	0.006***	0.011***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
adr high school	0.015	0.015**	0.026	0.010**	−0.021*
	(0.010)	(0.008)	(0.023)	(0.005)	(0.013)
n legis	−0.005	−0.005*	−0.003	−0.008***	0.001
	(0.004)	(0.003)	(0.007)	(0.002)	(0.005)
Constant	0.989***	0.794***	0.707***	0.557***	0.671***
	(0.077)	(0.037)	(0.117)	(0.070)	(0.074)
Observations	1,187	1,667	466	1,787	447
R^2	0.351	0.268	0.113	0.218	0.357
Adjusted R^2	0.346	0.264	0.096	0.214	0.344
F Statistic	70.598***	67.466***	6.465***	55.040***	26.948***

ECI refers to the Economic Complexity Index, *GDP pc* to the municipal Gross Domestic Product per capita, *pct young pop* to the percentage of the population aged 19 and under, *hub distance* to the distance to the nearest hub municipality, *adr high school* to the age-grade distortion rate (the percentage of students in the correct grade for their age in high school), and *n legis* to an institutional measure - the quantity of legislation and planning instruments existing in relation to the total expected by the IBGE Basic Municipal Information Survey. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Southeast. Distance to the nearest hub city relates to inequality in all regions. Better education is linked to lower inequality in the North, but the opposite is true in the Southeast and Northeast regions. Our institutional measure correlates negatively with inequality in the Southeast and Northeast, showing no relation in the other regions.

Part of our payment data comes from credit card expenses. However, the issue of indebtedness among the Brazilian population is a concerning fact, especially with credit cards. This aspect of indebtedness is not captured by our inequality index. Nevertheless, as a robustness check, we reanalyzed the data, considering only payments made through Pix, and found the main results are similar to those when we include credit card payments in our consumption proxy. The results of this test are presented in Table 6 in E.

To further ensure the robustness of our results, we conducted additional regressions by systematically removing one or two control variables at a time. The first specification includes only the variables of interest in the regression. We then performed five separate regressions, each excluding a different control: the natural logarithm of per capita GDP and its square, the natural logarithm of the population and the percentage of young people, the education measure, the distance from the municipality to the nearest hub city, and the institutional measure. The results showed the coefficients for ECI and ECI^2 remained negative and statistically significant at the 1% level in all cases, reaffirming the robustness of our findings. The results of this test are presented in Table 7 in E. These tests confirm our main results are not driven by any specific control variable and are robust to various model specifications.

6.4 Limitations

Due to data limitations, we do not have long enough time series to perform fixed effects panel estimations. Payment method data is recent, with the introduction of Pix, for example, occurring only in 2020. Its adoption continues to grow: around 1.4 billion transactions were made in December 2021, while over 3 billion were conducted in July 2023[22]. This trend is likely changing the distribution of payment method data over the past years and is expected to continue doing so in the future.

Another significant limitation is the scarcity and outdated nature of municipal data. For example, municipal GDP is computed with a two-year lag. This limitation affects the quality of the controls used in our regressions. Our measure of institutions, for example, does not capture the quality of the laws developed for each municipality, nor whether these laws are actually being applied or adopted. Additionally, we lack a measure for Informal Institutions. Unfortunately, we did not find more predictor variables for inequality with available data for all Brazilian municipalities, such as corruption, beyond those we already used as controls in our work.

One possible solution would be to use subgroups of municipalities, such as the larger ones, where more data is available. However, this approach could introduce selection bias, as larger municipalities may differ systematically from smaller ones in ways that are not captured by our model (e.g., larger municipalities may be more economically complex). Additionally, this would significantly reduce our sample size, limiting the generalizability of our findings. Therefore, as more data becomes available, future research can explore more robust estimation methods, such as fixed effects, to better address these limitations.

[22]For more details, see `https://www.bcb.gov.br/estabilidadefinanceira/estatisticaspix`.

7 Conclusion

Municipal-level inequality is positively correlated with crime and negatively with income growth, and people living in more unequal municipalities declare themselves more unhappy (Glaeser et al., 2009). However, measuring inequality at the municipal level depends on census surveys, which, in the case of Brazil, are conducted approximately once a decade. This compromises the proposition and evaluation of local public policies aimed at reducing inequality and its impacts.

In this context, our work proposes to calculate consumption inequality at the municipal level using electronic payment data, such as credit card and Pix payments, as a proxy. In this way, we can timely calculate a measure of municipal-level inequality at any desired frequency. We show our Gini index of consumption inequality is moderately correlated with the income inequality Gini index calculated using the 2010 IBGE Census data for Brazilian municipalities. Additionally, our consumption inequality index presents a similar regional distribution ordering to the income inequality index, although it generally indicates higher inequality, given the nature of the data used (electronic payments).

As an application of our index, we evaluate the relationship between consumption inequality and Economic Complexity at the municipal level. In a cross-sectional analysis, we show that Economic Complexity is negatively related to consumption inequality, meaning municipalities with more complex productive structures tend to have lower consumption inequality. This relationship is non-linear, indicating that the greater the complexity, the same variation in economic complexity is associated with a greater variation in inequality. This result remains quantitatively consistent when evaluated region by region.

Declaration of Competing Interest

The authors declare they do not have known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Disclaimer: The views expressed by the authors in this paper

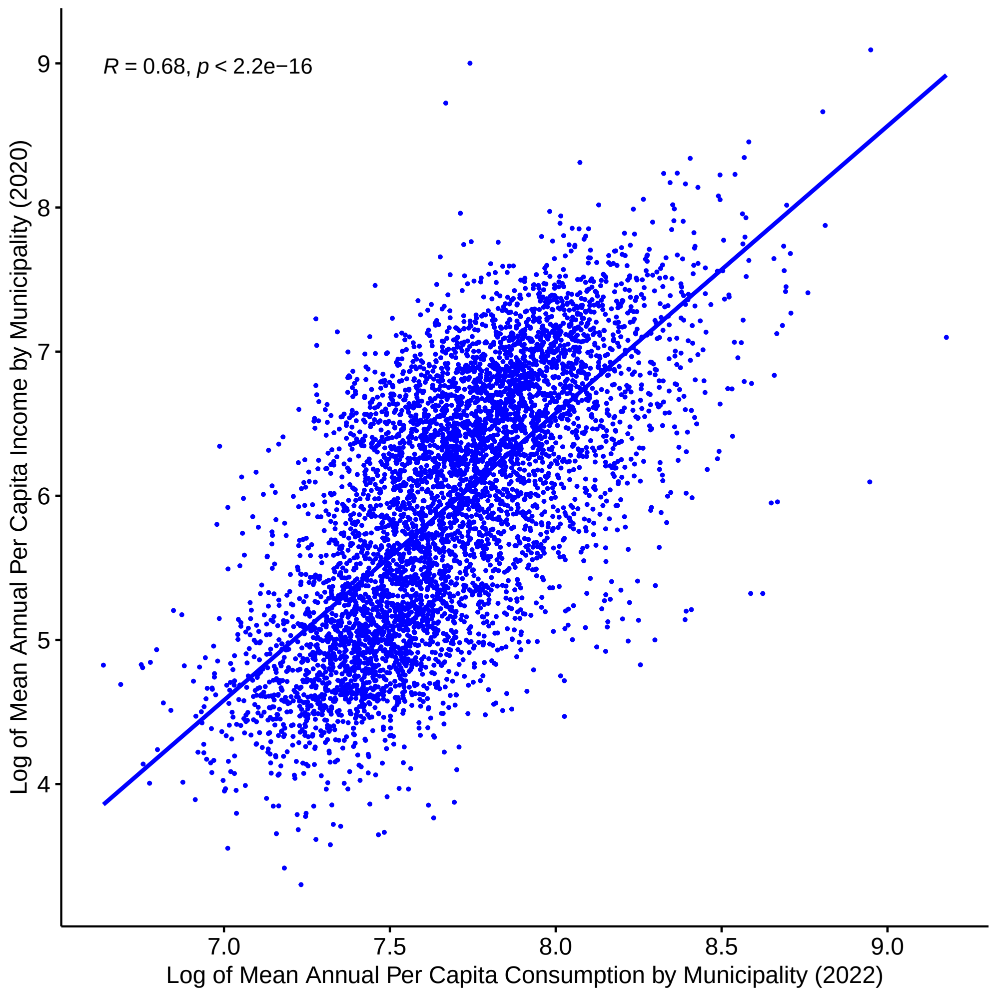
Acknowledgements

We would like to express our gratitude to José Luiz Rossi Jr, Bernardo Mueller, the Research and Consultancy Studies (Conep) team—especially Mário Bastos, Rogério Lucca, and Lorena Silva Brandão—all of whom contributed with their invaluable assistance, support, insights, or constructive feedback, which significantly improved this work.

DOC thanks to CNPq for partial financial support.

A Correlation between consumption and income data

Figure 8: **Correlation between per capita consumption and income.** This figure shows the correlation between the average annual per capita consumption of each municipality, based on electronic payment data from the base year 2022, and the average annual per capita income estimated by Neri and Hecksher (2023) for the base year 2020. The Pearson correlation coefficient (R) is 0.68, indicating a strong positive correlation. This suggests our proxy for consumption is closely related to the per capita income at the municipal level, demonstrating the validity of using electronic payment data to estimate consumption patterns.



B Boxplots of municipal consumption Gini coefficients, by state

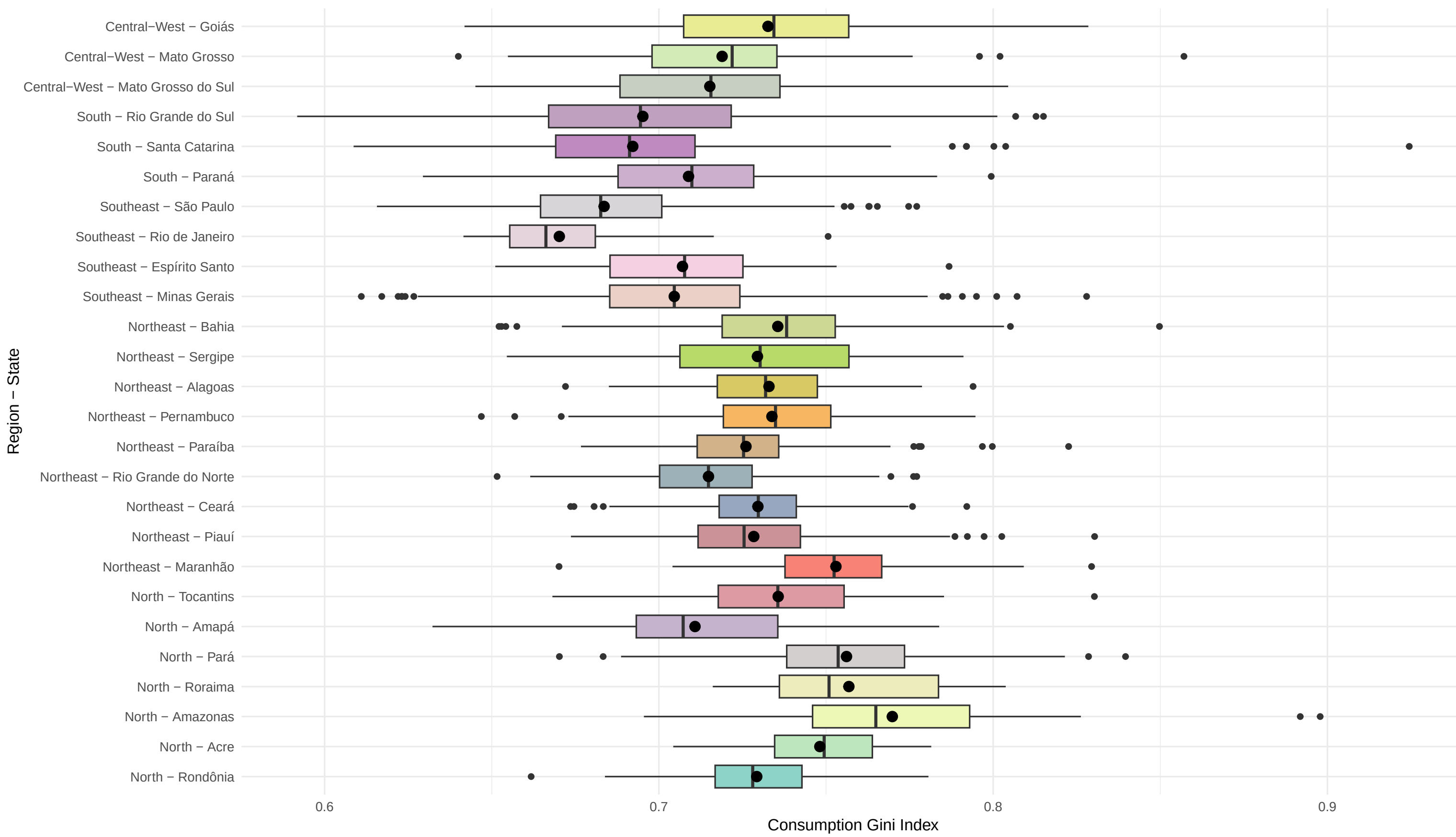


Figure 9: **Boxplots of municipal consumption Gini coefficients.** This figure displays boxplots for the consumption Gini coefficients across different states. A boxplot is a graphical representation that shows the distribution of a dataset. Each boxplot displays the median (central line), the first quartile (bottom of the box), the third quartile (top of the box), and the potential outliers. The states in the South and Southeast regions have the lowest averages of Gini coefficients, indicating lower consumption inequality, while some states in the North and Northeast regions show higher averages. This pattern is consistent with the regional analysis shown in Figure 3, which is based on income inequality data from the 2010 Brazilian Census.

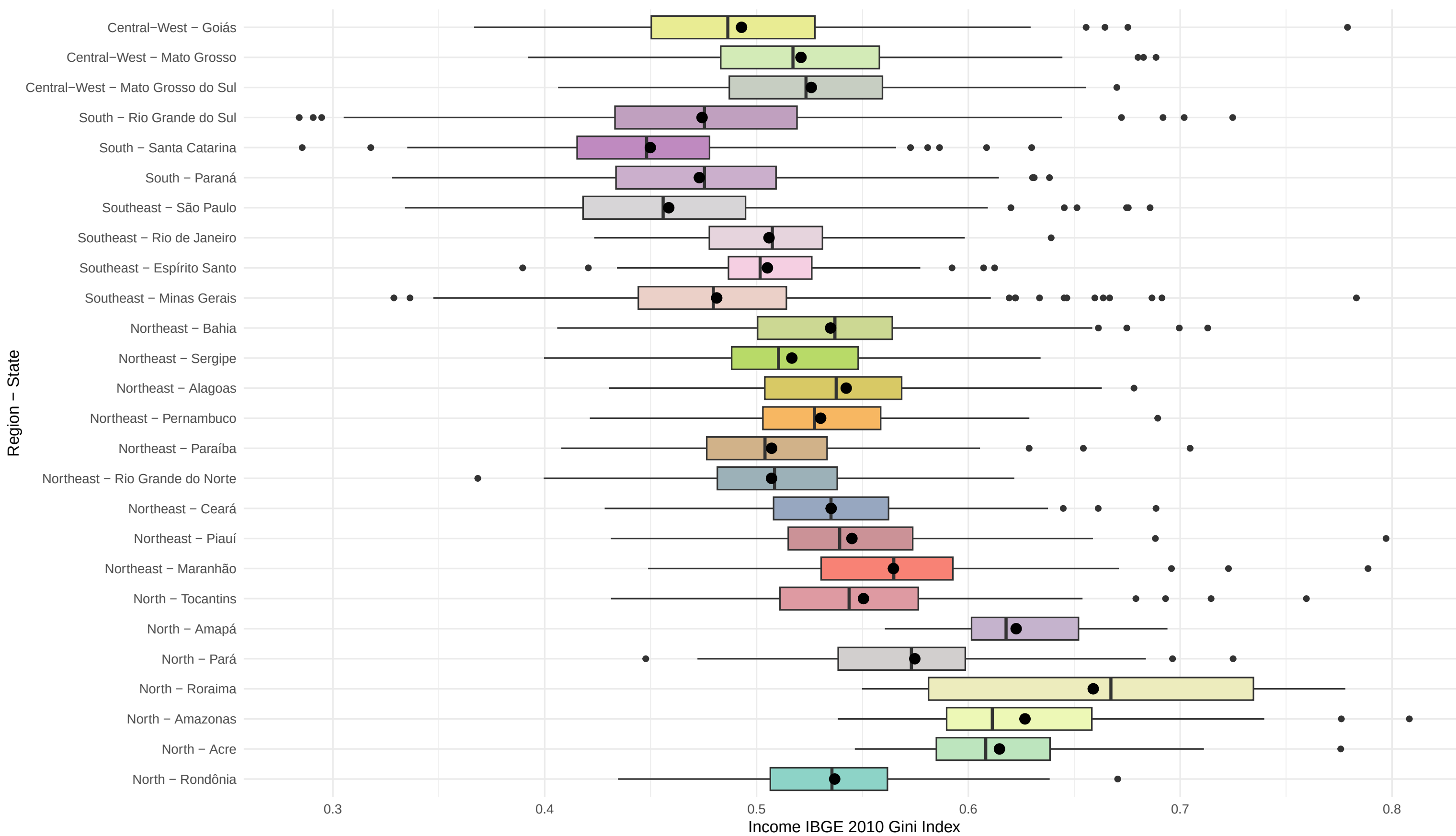


Figure 10: **Boxplots of municipal IBGE 2010 income Gini coefficients.** This figure displays boxplots for the income Gini coefficients based on data from the 2010 IBGE census, highlighting variability among states. As in the previous figure, the boxplots illustrate the distribution of Gini coefficients with the median, quartiles, and potential outliers. The states in the South and Southeast regions have the lowest averages of Gini coefficients, indicating lower income inequality. This pattern is similar to the consumption Gini index shown in Figure 9, indicating both consumption and income Gini indices follow a similar ranking across states.

C IBGE activity group codes not considered for consumption data

Table 4: **List of IBGE activity group codes excluded from payment data.** This table lists the IBGE activity group codes that identify institutions where payments made via Pix are excluded from being considered as consumption. These codes help differentiate between transactions that are part of regular consumption and those that fall outside this category.

Code	Activity
050	Coal Mining
060	Oil and Natural Gas Extraction
071	Iron Ore Mining
072	Extraction of Non-Ferrous Metal Ores
081	Extraction of Stone, Sand and Clay
089	Extraction of Other Non-Metallic Minerals
091	Support Activities for Oil and Natural Gas Extraction
099	Support Activities for Mining, Except Oil and Natural Gas
411	Real Estate Development
421	Construction of Roads, Railways, Urban Works and Special Works
422	Infrastructure Works for Electricity, Telecommunications, Water, Sewer and Pipeline Transport
641	Central Bank
646	Activities of Holding Companies
647	Investment Funds
649	Unspecified Financial Services Activities
653	Reinsurance
841	Administration of the State and Economic and Social Policy
842	Collective Services Provided by Public Administration
843	Compulsory Social Security
990	International Bodies and Other Extraterritorial Institutions

D Expected municipal legislation and planning instruments

Table 5: **Expected Legislation and Planning Instruments.** This table lists the various types of legislation and planning instruments expected for municipalities, based on information provided by the IBGE for the year 2021. We create a variable that measures the percentage of coverage of such legislation and planning instruments in each municipality.

Legislation/Instrument:

- City Master Plan
- Legislation on Special Social Interest Area/Zone
- Legislation on Special Interest Area/Zone
- Urban Perimeter Law
- Land Subdivision Legislation
- Zoning or Land Use and Occupation Legislation
- Created Soil Legislation or Onerous Granting of the Right to Build
- Improvement Contribution Legislation
- Consortium Urban Operation Legislation
- Neighborhood Impact Study Legislation
- Building Code
- Environmental Zoning or Ecological-Economic Zoning Legislation
- Administrative Servitude Legislation
- Heritage Preservation Legislation
- Conservation Unit Legislation
- Special Use Concession for Housing Legislation
- Special Urban Property Usucapion Legislation
- Surface Right Legislation
- Land Regularization Legislation
- Possession Legitimization Legislation
- Preliminary Environmental Impact Study Legislation
- Code of Postures

Source: IBGE, available at `https://www.ibge.gov.br/estatisticas/sociais/saude/10586-pesquisa-de-informacoes-basicas-municipais.html`.

E Robustness Check

E.1 Only Pix data as payment method

As a robustness check, the following results consider only Pix as the payment method:

Figure 11: **Histogram of Gini coefficients for consumption inequality using only Pix.** This figure presents a histogram of the Gini coefficients for consumption inequality across municipalities, calculated using only Pix as the payment method. The horizontal axis represents the Gini coefficient values, while the vertical axis shows the frequency of municipalities for each Gini coefficient range. The distribution of Gini coefficients using only Pix as payment method is similar to that in Figure 2.

Figure 12: **Boxplots of municipal consumption Gini coefficients using only Pix as the payment method.** This figure displays boxplots for the consumption Gini coefficients across different regions, calculated using only Pix as the payment method. A boxplot is a graphical representation that shows the distribution of a dataset. Each boxplot displays the median (central line), the first quartile (bottom of the box), the third quartile (top of the box), and the potential outliers. The regions in the South and Southeast continue to be the least unequal, while the North region has moved closer to the Northeast and Center-West regions in terms of inequality.

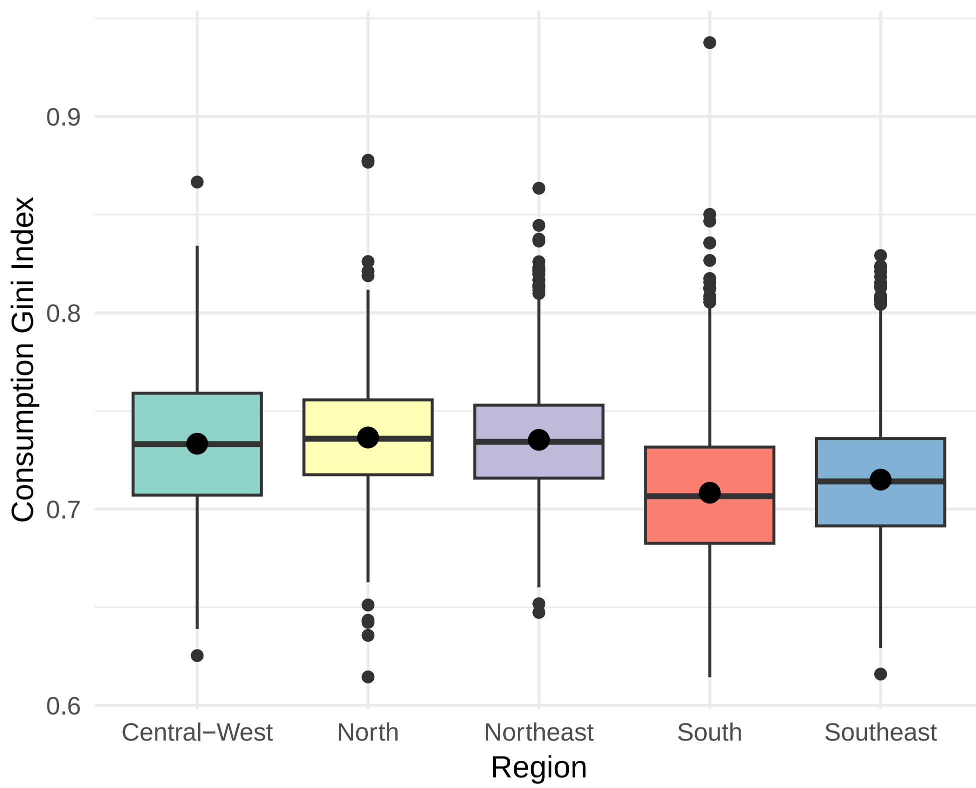


Figure 13: **Evolution of the Pearson correlation between consumption and IBGE income Gini coefficients as a function of an increasing sample's minimum municipal population threshold using only Pix as the payment method.** This figure shows the evolution of the Pearson correlation between the consumption Gini coefficients (our index) and the income Gini coefficients from the 2010 IBGE census as a function of an increasing minimum municipal population threshold, using only Pix as the payment method. The solid line represents the Pearson correlation coefficient with confidence intervals. The dashed blue line indicates the number of municipalities considered in the sample for a given minimum population threshold, as shown by the secondary y-axis. This figure illustrates that higher correlations can be observed when very small municipalities are excluded from the sample. For instance, considering only municipalities with more than 30,000 inhabitants, the Pearson correlation between the indices can be higher than 0.4. However, excluding municipalities with fewer than 30,000 inhabitants reduces the number of municipalities considered by more than three-quarters.

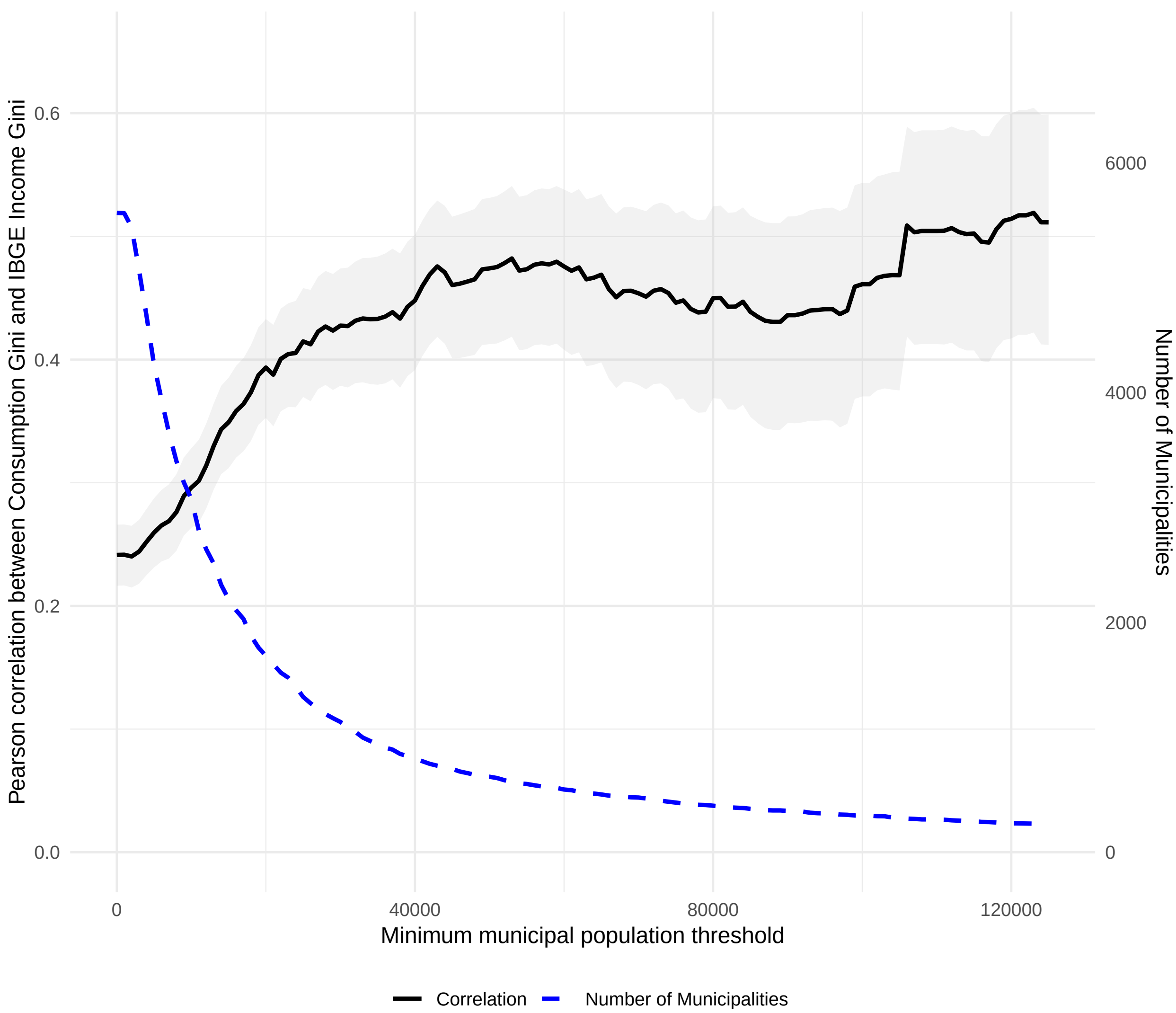


Figure 14: **Correlation between per capita consumption and income using only Pix.** This figure shows the correlation between the average annual per capita consumption of each municipality, based on Pix payment data from the base year 2022, and the average annual per capita income estimated by Neri and Hecksher (2023) for the base year 2020. The Pearson correlation coefficient (R) is 0.54, which is lower than when using all electronic payment methods (Figure 8). This suggests that credit card data, which are excluded in this exercise, may carry important information about consumption and/or income.

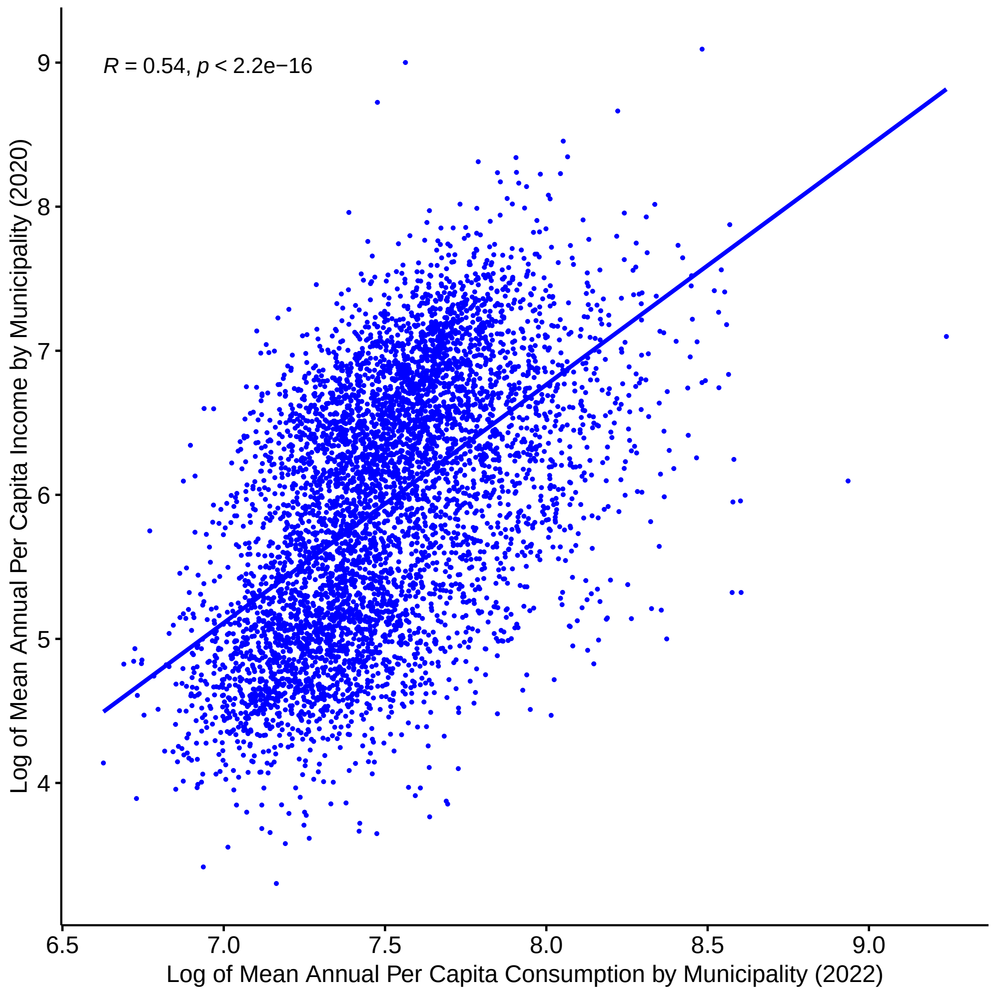


Table 6: OLS regressions for all municipalities, using only Pix as the payment method.

	Dependent variable:	
	Consumption Inequality Gini Index	
	(1)	(2)
ECI	-0.011^{***}	-0.010^{***}
	(0.001)	(0.001)
ECI^2		-0.006^{***}
		(0.0003)
ln(GDP pc)	-0.028^{***}	-0.016^{***}
	(0.005)	(0.005)
$\ln(\text{GDP pc})^2$	-0.004^{***}	-0.002^{***}
	(0.001)	(0.001)
ln(population)	0.006^{***}	0.009^{***}
	(0.001)	(0.001)
pct young pop	0.0004^{***}	0.0004^{***}
	(0.0001)	(0.0001)
ln(hub distance)	0.006^{***}	0.004^{***}
	(0.001)	(0.001)
adr high school	0.004	0.006^{*}
	(0.004)	(0.004)
n legis	-0.004^{**}	-0.007^{***}
	(0.002)	(0.002)
Constant	0.793^{***}	0.717^{***}
	(0.024)	(0.027)
Observations	5,554	5,554
R^2	0.132	0.177
Adjusted R^2	0.131	0.176
F Statistic	105.207^{***} (df = 8; 5545)	132.576^{***} (df = 9; 5544)

ECI refers to the Economic Complexity Index, *GDP pc* to the municipal Gross Domestic Product per capita, *pct young pop* to the percentage of the population aged 19 and under, *hub distance* to the distance to the nearest hub municipality, *adr high school* to the age-grade distortion rate (the percentage of students in the correct grade for their age in high school), and *n legis* to an institutional measure - the quantity of legislation and planning instruments existing in relation to the total expected by the IBGE Basic Municipal Information Survey. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

E.2 Additional regressions

In this section, we conduct additional regressions by systematically removing one or two control variables at a time. The results show the coefficients for ECI and ECI^2 remain negative and statistically significant at the 1% level in all cases, reaffirming the robustness of our findings.

Table 7: OLS regressions for all municipalities with different combinations of controls.

	Dependent variable: Gini					
	(1)	(2)	(3)	(4)	(5)	(6)
ECI	-0.016^{***}	-0.012^{***}	-0.012^{***}	-0.016^{***}	-0.008^{***}	-0.014^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ECI^2	-0.003^{***}	-0.004^{***}	-0.004^{***}	-0.005^{***}	-0.003^{***}	-0.004^{***}
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
ln(GDP pc)		−0.005	−0.004	−0.004	0.005	
		(0.005)	(0.005)	(0.005)	(0.005)	
$\ln(\text{GDP pc})^2$		−0.0001	−0.0001	−0.0004	0.002^{***}	
		(0.001)	(0.001)	(0.001)	(0.001)	
ln(population)		0.005^{***}	0.006^{***}	0.007^{***}		0.006^{***}
		(0.001)	(0.001)	(0.001)		(0.001)
pct young pop		0.001^{***}	0.002^{***}	0.002^{***}		0.002^{***}
		(0.0001)	(0.0001)	(0.0001)		(0.0001)
ln(hub distance)		0.010^{***}	0.010^{***}		0.014^{***}	0.010^{***}
		(0.001)	(0.001)		(0.001)	(0.001)
adr high school		-0.007^{***}		-0.007^{***}	-0.033^{***}	-0.009^{***}
		(0.003)		(0.003)	(0.003)	(0.003)
n legis			-0.003^{***}	-0.003^{***}	-0.005^{***}	-0.004^{***}
			(0.001)	(0.002)	(0.001)	(0.001)
Constant	0.718^{***}	0.653^{***}	0.642^{***}	0.629^{***}	0.697^{***}	0.628^{***}
	(0.001)	(0.027)	(0.026)	(0.028)	(0.028)	(0.006)
Observations	5,563	5,554	5,563	5,554	5,554	5,554
R^2	0.237	0.397	0.397	0.366	0.348	0.395
Adjusted R^2	0.236	0.396	0.396	0.365	0.347	0.394
F Statistic	861.646^{***}	456.132^{***}	456.132^{***}	400.625^{***}	422.459^{***}	516.542^{***}

ECI refers to the Economic Complexity Index, *GDP pc* to the municipal Gross Domestic Product per capita, *pct young pop* to the percentage of the population aged 19 and under, *hub distance* to the distance to the nearest hub municipality, *adr high school* to the age-grade distortion rate (the percentage of students in the correct grade for their age in high school), and *n legis* to an institutional measure - the quantity of legislation and planning instruments existing in relation to the total expected by the IBGE Basic Municipal Information Survey. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

References

Aguiar, M. and Bils, M. (2015). Has consumption inequality mirrored income inequality? *American Economic Review*, 105(9):2725–2756.

Alesina, A., Di Tella, R., and MacCulloch, R. (2004). Inequality and happiness: are europeans and americans different? *Journal of Public Economics*, 88(9-10):2009–2042.

Amarante, V., Lanzilotta, B., and Torres, J. (2023). Inequality and productive structure. *United Nations University Working Paper*, (2023/9).

Aprigliano, V., Ardizzi, G., and Monteforte, L. (2019). Using payment system data to forecast economic activity. *60th issue (October 2019) of the International Journal of Central Banking*.

Arif, I. (2021). Productive knowledge, economic sophistication, and labor share. *World Development*, 139:105303.

Attanasio, O. and Pistaferri, L. (2014). Consumption inequality over the last half century: some evidence using the new psid consumption measure. *American Economic Review*, 104(5):122–126.

Attanasio, O. P. and Pistaferri, L. (2016). Consumption inequality. *Journal of Economic Perspectives*, 30(2):3–28.

Aydin, D. (2015). The marginal propensity to consume out of liquidity: Evidence from a randomized controlled trial. 2015 Meeting Papers 270, Society for Economic Dynamics.

Baker, S. R. (2018). Debt and the response to household income shocks: Validation and application of linked financial account data. *Journal of Political Economy*, 126(4):1504–1557.

Baker, S. R. and Yannelis, C. (2017). Income changes and consumption: Evidence from the 2013 federal government shutdown. *Review of Economic Dynamics*, 23:99–124.

Bandeira Morais, M., Swart, J., and Jordaan, J. A. (2021). Economic complexity and inequality: Does regional productive structure affect income inequality in Brazilian states? *Sustainability*, 13(2):1006.

Barro, R. J. (2008). Inequality and growth revisited. Technical Report 11, ADB Working Paper Series on Regional Economic Integration.

Bee, A. and Rothbaum, J. (2019). The administrative income statistics (ais) project: Research on the use of administrative records to improve income and resource estimates. *US Census Bureau SEHSD working paper*, 36(2019):2017–39.

Brandão, L. S. (2023). *Complexidade Econômica: uma análise de como as desigualdades se materializam no território*. PhD thesis, Universidade de Brasília.

Browning, M. and Leth-Petersen, S. (2003). Imputing consumption from income and wealth information. *The Economic Journal*, 113(488):F282–F301.

Burkhauser, R. V., Hérault, N., Jenkins, S. P., and Wilkins, R. (2018). Top incomes and inequality in the UK: reconciling estimates from household survey and tax return data. *Oxford Economic Papers*, 70(2):301–326.

Ceriani, L. and Verme, P. (2012). The origins of the gini index: extracts from variabilità e mutabilità (1912) by Corrado Gini. *The Journal of Economic Inequality*, 10:421–443.

Chu, L. K. and Hoang, D. P. (2020). How does economic complexity influence income inequality? New evidence from international data. *Economic Analysis and Policy*, 68:44–57.

Constantine, C. and Khemraj, T. (2019). Geography, economic structures and institutions: A synthesis. *Structural Change and Economic Dynamics*, 51:371–379.

Cowell, F. A. (2011). *Measuring inequality*. Oxford University Press, 3rd edition.

De Dominicis, L., Florax, R. J., and De Groot, H. L. (2008). A meta-analysis on the relationship between income inequality and economic growth. *Scottish Journal of Political Economy*, 55(5):654–682.

De Maio, F. G. (2007). Income inequality measures. *Journal of Epidemiology & Community Health*, 61(10):849–852.

Deininger, K. and Squire, L. (1996). A new data set measuring income inequality. *The World Bank Economic Review*, 10(3):565–591.

Eika, L., Mogstad, M., and Vestad, O. L. (2020). What can we learn about household consumption expenditure from data on income and assets? *Journal of Public Economics*, 189:104163.

Galbraith, J. K. (2007). Global inequality and global macroeconomics. *Journal of Policy modeling*, 29(4):587–607.

Gao, J. and Zhou, T. (2018). Quantifying China's regional economic complexity. *Physica A: Statistical Mechanics and its Applications*, 492:1591–1603.

Gelman, M., Kariv, S., Shapiro, M. D., Silverman, D., and Tadelis, S. (2014). Harnessing naturally occurring data to measure the response of spending to income. *Science*, 345(6193):212–215.

Gelman, M., Kariv, S., Shapiro, M. D., Silverman, D., and Tadelis, S. (2020). How individuals respond to a liquidity shock: Evidence from the 2013 government shutdown. *Journal of Public Economics*, 189:103917.

Glaeser, E. L., Resseger, M., and Tobio, K. (2009). Inequality in cities. *Journal of Regional Science*, 49(4):617–646.

Gonçalves, R. N. C. et al. (2022). Nowcasting Brazilian GDP with eletronic payments data. Working Paper 564, Banco Central do Brasil.

Gross, D. B. and Souleles, N. S. (2002). Do liquidity constraints and interest rates matter for consumer behavior? Evidence from credit card data. *The Quarterly Journal of Economics*, 117(1):149–185.

Hartmann, D. (2014). *Economic complexity and human development: How economic diversification and social networks affect human agency and welfare*. Number 110. Routledge Studies in Development Economics.

Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., and Hidalgo, C. A. (2017). Linking economic complexity, institutions, and income inequality. *World Development*, 93:75–93.

Hartmann, D. and Pinheiro, F. L. (2022). Economic complexity and inequality at the national and regional level. *arXiv preprint arXiv:2206.00818*. Preprint.

Hassett, K. A. and Mathur, A. (2012). A new measure of consumption inequality. *AEI Economic Studies*, (2).

Hausmann, R. and Hidalgo, C. A. (2011). The network structure of economic output. *Journal of economic growth*, 16:309–342.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., and Simoes, A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press, Cambridge, MA, USA.

Hidalgo, C. A. (2021). Economic complexity theory and applications. *Nature Reviews Physics*, 3(2):92–113.

Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the national academy of sciences*, 106(26):10570–10575.

Hokayem, C., Bollinger, C., and Ziliak, J. P. (2015). The role of cps nonresponse in the measurement of poverty. *Journal of the American Statistical Association*, 110(511):935–945.

Kemp-Benedict, E. (2014). An interpretation and critique of the method of reflections. *Munich Personal RePEc Archive*, (60705).

Kolsrud, J., Landais, C., and Spinnewijn, J. (2017). Studying consumption patterns using registry data: lessons from Swedish administrative data. LSE Research Online Documents on Economics 87777, London School of Economics and Political Science.

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review*, 45:1–28.

Larrimore, J., Mortenson, J., and Splinter, D. (2021). Household incomes in tax data: Using addresses to move from tax-unit to household income distributions. *Journal of Human Resources*, 56(2):600–631.

Le Caous, E. and Huarng, F. (2020). Economic complexity and the mediating effects of income inequality: Reaching sustainable development in developing countries. *Sustainability*, 12(5):2089.

Lee, C.-C. and Wang, E.-Z. (2021). Economic complexity and income inequality: Does country risk matter? *Social Indicators Research*, 154:35–60.

Lee, K.-K. and Vu, T. V. (2020). Economic complexity, human capital and income inequality: a cross-country analysis. *The Japanese Economic Review*, 71:695–718.

Medeiros, M., Souza, P. H. G. F. d., and Castro, F. Á. d. (2015). A estabilidade da desigualdade de renda no Brasil, 2006 a 2012: estimativa com dados do imposto de renda e pesquisas domiciliares. *Ciência & Saúde Coletiva*, 20:971–986.

Meyer, B. D. and Sullivan, J. X. (2009). Five decades of consumption and income poverty. NBER Working Paper 14827, National Bureau of Economic Research.

Moore, J. C., Stinson, L. L., Welniak, E. J., et al. (2000). Income measurement error in surveys: A review. *Journal of Official Statistics*, 16(4):331–362.

Neri, M. and Hecksher, M. (2023). Mapa da riqueza. `https://cps.fgv.br/riqueza`. Accessed: 2023-11-27.

Olafsson, A. and Pagel, M. (2018). The liquid hand-to-mouth: Evidence from personal finance management software. *The Review of Financial Studies*, 31(11):4398–4446.

Palma, J. G. (2011). Homogeneous middles vs. heterogeneous tails, and the end of the 'inverted-u': It's all about the share of the rich. *development and Change*, 42(1):87–153.

Pereira, C., Alston, L. J., Melo, M. A., and Mueller, B. (2016). *Brazil in transition: Beliefs, leadership, and institutional change*. Princeton University Press.

Perera, L. D. H. and Lee, G. H. (2013). Have economic growth and institutional quality contributed to poverty and inequality reduction in Asia? *Journal of Asian Economics*, 27:71–86.

Pham, M. H., Truong, H. D. H., and Hoang, D. P. (2023). Economic complexity, shadow economy, and income inequality: fresh evidence from panel data. Preprint v1 from Research Square.

Piketty, T. and Saez, E. (2003). Income inequality in the United States, 1913–1998. *The Quarterly Journal of Economics*, 118(1):1–41.

Piketty, T., Saez, E., and Zucman, G. (2018). Distributional national accounts: methods and estimates for the United States. *The Quarterly Journal of Economics*, 133(2):553–609.

Roe, M. J. and Siegel, J. I. (2011). Political instability: Effects on financial development, roots in the severity of economic inequality. *Journal of Comparative Economics*, 39(3):279–309.

Sbardella, A., Pugliese, E., and Pietronero, L. (2017). Economic development and wage inequality: A complex system analysis. *PloS one*, 12(9):e0182774.

Sebastian, R. and Biagi, F. (2018). The routine biased technical change hypothesis: a critical review. Working Paper 2018/11, Joint Research Centre (Seville site).

Soave, G. P., Gomes, F. A. R., and Junior, F. B. (2019). Desigualdade e desenvolvimento: revisitando a hipótese de Kuznets após a redução da desigualdade nos municípios brasileiros. *Revista Brasileira de Estudos Regionais e Urbanos*, 13(4):581–605.

Sousa, R. A. d. (2018). A teoria da complexidade reencontra o desenvolvimento econômico: uma análise de insumo-produto. Master's thesis, Universidade de Brasília.

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., and Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Scientific Reports*, 2(1):723.

Thomas, C. (2015). Income inequality and economic development in Latin America: A test for the kuznets inverted-u curve. *Indian Journal of Economics and Business*, 14(1).

Török, I., Benedek, J., and Gómez-Zaldívar, M. (2022). Quantifying subnational economic complexity: Evidence from Romania. *Sustainability*, 14(17):10586.

Trapeznikova, I. (2019). Measuring income inequality. *IZA World of Labor*, (462).

Violante, G. L. (2008). Skill-biased technical change. *The new Palgrave dictionary of economics*, 2:1–6.

Wu, S., Zheng, X., and Wei, C. (2017). Measurement of inequality using household energy consumption data in rural China. *Nature Energy*, 2(10):795–803.