

Article

Economic Complexity and Inequality: Does Regional Productive Structure Affect Income Inequality in Brazilian States?

Margarida Bandeira Morais ¹, Julia Swart ² and Jacob Arie Jordaan ^{2,*}

¹ Department of Geography and Environment, London School of Economics, Houghton Street, London WC2A 2AE, UK; m.bandeira-morais@lse.ac.uk

² Utrecht University School of Economics, Adam Smith Hall, International Campus Utrecht, Kriekenpitplein 21-22, 3584 EC Utrecht, The Netherlands; j.swart@uu.nl

* Correspondence: j.a.jordaan@uu.nl

Abstract: Recent research on the effects of the productive structure of an economy has turned to examining whether economic complexity is associated with lower income inequality. In contrast to the commonly adopted approach that estimates the impact of economic complexity in a cross-country setting, we use panel data for Brazilian states to identify the relationship between economic complexity and income inequality at the sub-national level. Our findings show that the relationship between economic complexity and income inequality has an inverted U-shape, indicating that growing levels of complexity first worsen and then improve the income distribution in Brazilian states. Our findings also show that this relationship is particularly prominent in those states that have relatively high levels of urbanization and overall development. Furthermore, we identify separate effects on income inequality from the degree to which regional productive structures are characterised by diversity in terms of industries and occupations. These effects are particularly pronounced in less developed states with a more rural character. In combination, these findings confirm the important role that the productive structure plays in processes that drive improvements in income distributions and suggest that more research on this impact is warranted at the regional level.

Keywords: income inequality; productive structure; economic complexity; Brazil; regional



Citation: Bandeira Morais, M.; Swart, J.; Jordaan, J.A. Economic Complexity and Inequality: Does Regional Productive Structure Affect Income Inequality in Brazilian States? *Sustainability* **2021**, *13*, 1006. <https://doi.org/10.3390/su13021006>

Received: 18 December 2020

Accepted: 16 January 2021

Published: 19 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The notion that the productive structure and the structural transformation of an economy play important roles in processes of economic growth and development can be traced back to original contributions by economists including Rosenstein-Rodan [1], Prebisch [2] and Singer [3]. However, conceptual and data limitations have substantially constrained empirical research in their attempts to accurately identify the association between productive structure and economic growth. Predominantly, research has been characterised either by the use of broad indicators of structural transformation, such as changing shares of manufacturing in economic activity [4,5], or aggregate indicators of degrees to which economic or industrial activity are diversified [6–8].

With the introduction of the concept of economic complexity [9,10], this research strand has received an important impetus. Economic complexity, relating to the degree of sophistication of a country's productive structure, is affected by a multitude of factors that are related to economic growth, including factor endowments, geography, institutions, social capital and a country's historical trajectory [11]. In line with this interpretation, it is likely that countries with a higher level of economic complexity are characterised by higher growth and that increases in economic complexity exercise a positive effect on economic growth rates. This is confirmed by a large number of cross-country studies that identify positive effects of a country's level of sophistication of export baskets or its degree of economic complexity on economic growth [9,10,12,13].

More recently, research on the effects of economic complexity has been broadened by examining whether a country's productive structure also matters for the distribution of income. It is increasingly understood that, similarly to processes of economic growth, the income distribution in a country is affected by a number of underlying factors, many of which are related to, or directly captured by, the concept of economic complexity. For instance, the degree of sophistication of a country's product mix reflects the level of knowledge that is embodied in the country's population, as well as the availability of ample job opportunities [11,14]. Furthermore, the quality and inclusiveness of institutions is likely to co-evolve with the level of economic complexity of an economy. This suggests that countries with a high degree of economic complexity will be characterised by lower levels of income inequality, as evidenced for instance by the cross-country studies by Hartmann et al. [15] and Lee and Vu [16].

However, it may also be the case that economic complexity, especially in countries that are operating in earlier stages of economic development, is associated with higher income inequality. Processes of structural transformation that involve moving from low to higher value-added activities are often characterised by increasing levels of capital intensity and the introduction of and increasing reliance on new technologies. This will favour relative returns for high-skilled workers in an economy, resulting in an increase in income inequality [17]. Given that the relationship between economic complexity and income distribution is difficult to predict a priori as it is open to multiple interpretations, more research is needed in order to improve our understanding of the exact nature of this relationship.

In this paper we extend upon this emerging research strand by conducting a novel empirical analysis to obtain further evidence on the potential significance of the relationship between economic complexity and the income distribution as well as on the nature of this relationship. The contribution of our paper is three-fold. First, instead of using a cross-country framework, we focus on the impact of economic complexity on income inequality within an individual country. It is increasingly understood that drivers of economic growth, such as institutions and knowledge spillovers caused by the agglomeration of firms, workers and innovation, as well as inter-firm linkages between multinational corporations and domestic firms, can differ substantially across regions within countries [18–21]. This not only suggests that economic complexity is likely to be subject to substantial variation at the regional level within a country, but also that an analysis of the relationship between economic complexity and income inequality at the regional level will be particularly informative.

Second, inspired by the original contribution of Kuznets [22] on the nature of the relationship between the income level of a country and the distribution of income, we examine whether the relationship between economic complexity and income inequality is characterised by non-linearity. Kuznets [22] proposed that the relationship between the income of a country and its income inequality follows an inverted U-shape, whereby, following an initial phase where increasing income is accompanied with growing inequality, further increases in income generate a more equal income distribution. The relationship between economic complexity and income inequality may have a similar shape. A country with a low level of economic complexity may experience an increase in income inequality when the production of more sophisticated products disproportionately benefits high-skilled workers. At some level of economic complexity, other forces such as inclusive institutions, rising job opportunities and stronger worker representation may then become more important, resulting in the effect of economic complexity on income inequality becoming negative.

Third, we conduct our analysis on Brazil, a country that is characterised by high levels of poverty and inequality. Brazil can be seen to be representative of a substantial number of countries in the world economy who, whilst according to their average income can be classified as lower-middle or upper-middle income countries, are still facing relatively high levels of poverty and income inequality. Furthermore, these countries are often

characterised by marked regional differences in terms of income and inequality, further supporting our focus on analysing the relationship between economic complexity and income inequality at the regional level. In this way, our findings for Brazil may also be relevant for this wider group of countries that find themselves in a similar situation.

The paper is constructed as follows. Section two surveys the small body of empirical evidence on the relationship between economic complexity and income inequality. Section three discusses income inequality in Brazil. Section four discusses the dataset and the specification of the econometric model. In section five we present our main empirical findings, which show that the relationship between economic complexity and income inequality follows an inverted U-shape, indicating that growing levels of complexity first worsen and then improve the income distribution in Brazilian states. This relationship is particularly prominent in states with a relatively high level of urbanisation and overall development. We also find that industry and occupational diversity exercise additional effects on income inequality, further strengthening the notion that the regional productive structure constitutes an important factor in processes impacting on the distribution of income in Brazilian states. Finally, section six summarises and concludes.

2. Literature Review

Following an initial focus on the impact of the productive structure of a country on economic growth, the introduction of the economic complexity index (ECI) by Hidalgo and Hausman [9] has most recently fostered research on the effect of ECI on income inequality [14–17,23]. Economic complexity can be seen as a “high-resolution expression” ([15], p. 85) of a number of underlying factors, including institutions, human capital, the availability of job opportunities and worker representation. This suggests that countries with a higher level of economic complexity will be characterised by a lower level of income inequality, resulting from better institutions, more job opportunities, among other improvements. However, the findings of the small number of available studies are mixed, with some studies finding a negative association and other studies reporting a positive effect.

Hartmann et al. [15] were the first to estimate the impact of the ECI on income distribution. Using data for over 150 countries for the period 1963–2008, they test for a linear relationship and find that economic complexity is a negative predictor of income inequality. Their findings also show that, when controlling for economic complexity, the rising part of the Kuznets curve-like relationship between GDP and income inequality becomes more pronounced. Furthermore, they examine the effect of the Product Gini Index (PGI), linking product categories to different levels of income inequality. Their findings show that products associated with the highest levels of income inequality (i.e., a high PGI score) consist mainly of commodities (e.g., cocoa beans and animal hair), which have a low level of economic complexity. In contrast, low PGI products include more sophisticated forms of machinery and manufacturing products (e.g., textiles, machinery and road rollers), involving a high level of economic complexity. Using the same framework, Hartmann et al. [11] look at the structural constraints of income inequality in Latin America, by comparing the productive sophistication and structural constraints on income inequality of Latin American and Caribbean countries with that of China and other high-performing Asian economies. Their results show that Latin American and Caribbean countries continue to export products associated with high levels of inequality and low levels of economic complexity, and their productive structure strongly constrains their ability to generate and distribute income. The intuition behind these findings is that complex products require a larger network of skilled workers, related industries, and inclusive institutions for economic competitiveness. Such characteristics are conducive to more equal societies. In contrast, the competitiveness of simple industrial products and resource-exploiting activities is mainly based on resource richness, low labour costs, routinised activities and economies of scale, characteristics that foster more unequal economies.

In contrast, other studies present evidence that indicates that economic complexity fosters income inequality. It is likely that economic complexity co-evolves with higher

levels of value added, productivity, use of modern technologies, etc. In such cases, it may be that rising levels of economic complexity generate larger wage differentials between low and high-skilled workers, resulting in increased income inequality. An example of a study that presents evidence of such a positive effect is Lee and Vu [16]. They estimate the effect of the productive structure using a system-GMM (Generalised Method of Moments) dynamic panel with 113 countries with 5-year averages for the period 1965–2014 and find a positive relationship between ECI and income inequality. Similarly, Chu and Hoang [17] carry out a system-GMM estimation covering 88 countries for the period 2002–2007 and find that economic complexity is associated with higher income inequality. Lee and Wang [23] also report a positive impact of economic complexity on income inequality, based on the use of a finite mixture model for 43 countries for the period 1991–2016.

In extension of these findings, some studies provide evidence that suggests that the relationship between the ECI and income inequality may be characterised by non-linearity. Kuznets [22] argued that it is likely that the relationship between the level of income and income inequality has an inverted U-shape. At early levels of economic development, the beneficial effects from economic growth are likely to materialise amongst sub-groups of a country's population, generating increased income inequality. With ongoing increases in economic development more and more people will start to participate in and benefit from the growing economy, resulting in a negative effect on income inequality. A similar process may apply to the effect of economic complexity on income distribution. Starting at low levels of economic complexity, increases in complexity may benefit capital owners and high-skilled workers in particular, resulting in a worsening of income distribution. With ongoing increases in economic complexity, other components of economic complexity (e.g., institutions, worker representation, job opportunities) may become relatively more important, which would change the effect of economic complexity on income inequality from positive into negative at some stage.

Hausmann et al. [10] present findings that suggest that such a non-linear relationship may be important. They look at opportunity value, or the rewards of knowledge accumulation, and how it relates to economic complexity. Their data reveal that countries with a low ECI have low rewards for knowledge; this is because countries with a low ECI are unable to effectively put knowledge to productive use. However, countries with high levels of productive knowledge also have low rewards for knowledge. In such countries, productive knowledge already occupies a large fraction of the product space, limiting the returns from further knowledge accumulation. Countries with an intermediate level of complexity vary much more widely in their opportunity value. If we associate opportunity value with wage levels, their findings imply that the relationship between the ECI and wage differentials is characterised by non-linearity. Le et al. [24] also provide evidence for this. They use data for 90 countries for the period 2002–2014 to estimate the effect of export diversification on income inequality and find evidence of an inverted U-shape relationship. Although economic complexity is a more comprehensive indicator of a country's productive structure than the level of export diversification, their findings can be taken as support for the notion that the effect of economic complexity on income inequality is non-linear.

Summing up, the small body of evidence indicates that economic complexity plays an important role as driver of income inequality. However, the evidence is mixed, with different studies reporting negative or positive effects, indicating that more research is urgently required. Considering the multitude of factors that are encapsulated by the concept of economic complexity, the possibility that the relationship between economic complexity and the income distribution has a non-linear character is particularly interesting to examine further empirically. More precisely, whereas early increases in economic complexity can foster an increase in income inequality, ongoing increases in economic complexity may start to generate a negative impact on income inequality.

3. Income Inequality in Brazil

Figure 1 shows the evolution of the Gini coefficient for Brazil from 1976 to 2014. The Gini coefficient is a frequently used measure capturing the degree of inequality in income between households. The value of the Gini coefficient ranges between 1 (perfect inequality where total income is concentrated in 1 household) and 0 (perfect equality where all households have a similar share in total income). Despite some fluctuations in the first half of this period, Brazil has experienced a steady decrease in inequality from 1993 onwards, with the Gini coefficient decreasing from 0.604 in 1993 to 0.518 in 2014. Nevertheless, income inequality is still very high, and there are marked disparities in the Gini coefficient across Brazilian states, thus attracting a lot of attention from researchers.

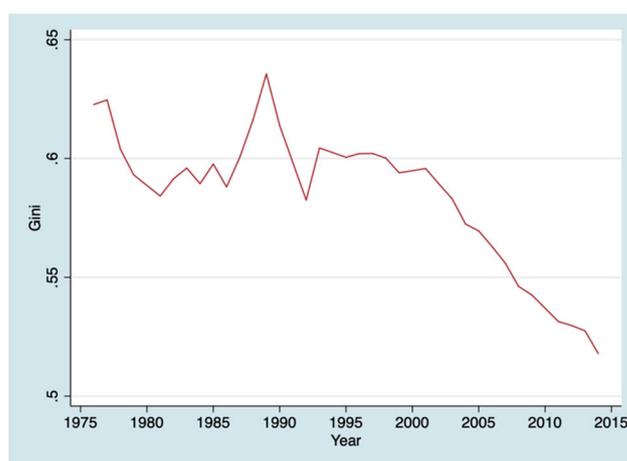


Figure 1. Income inequality in Brazil, from 1976 to 2014, measured by the Gini coefficient. Authors' elaboration using data from the Institute for Applied Economic Research (IPEA).

Most of the studies on income inequality in Brazil distinguish between two main periods, 1981 to early 1990s and 1993 onwards [25]. A few authors focus solely on one of these periods (e.g., [26–28]), while others focus on both periods and explicitly try to differentiate the determinants of inequality in each of the periods (e.g., [29]). While there is agreement that inequality in Brazil was not driven by the same factors throughout each of the two time periods, there is no overall consensus with respect to the relative importance of different driving forces. There are three broad groups of variables commonly distinguished as important drivers of inequality in Brazil. The first group is related to education and socioeconomic aspects, the second one relates to macroeconomic aspects (in particular unemployment and inflation) and the third one concerns international trade.

First, when focusing on the period from the 1990s onwards, a number of studies identify several factors that have contributed to the decline in income inequality, which are also linked to stark differences in income inequality between Brazilian states. In particular, education and average returns to schooling [26,29–31], government transfers and social assistance programmes [29,30], job formality [31], spatial and sectoral labour market integration [30] and changes in racial inequality [26,29,31] have been identified as significant predictors of income inequality in Brazil.

Second, some authors argue that, next to education and other socioeconomic aspects, inflation and unemployment were more important in explaining inequality in Brazil, in particular during the 1980s [27,29]. Both inflation and unemployment can generate higher income inequality—the former by pushing middle-income groups into poverty (inflation reduces the real income of all but more strongly affects the group in the middle), and the latter by decreasing the incomes of those who are unemployed.

The third group of variables relates to the impact of international trade on income inequality. Brazil went through a period of significant trade liberalisation between 1989 and 1995, with tariff levels remaining relatively stable in subsequent periods, particularly

from the early 2000s onwards. The effects of trade liberalisation have received significant attention in Brazil, because it impacted the country differently than it did other Latin American countries. In countries such as Colombia and Mexico, trade liberalisation fostered a pronounced increase in inequality, counter to theoretical predictions [32]. In Brazil, however, trade liberalisation impacted wage inequality in the opposite direction [25,33]. This was likely driven by a reduction in the wage premium of skilled workers and a movement of workers away from previously protected industries [25]. Nevertheless, some ambiguity in the empirical evidence remains, with some papers finding no evidence of any effect from trade liberalisation on the Brazilian wage distribution [34], and other studies presenting evidence of an initial positive and subsequently negative impact of trade liberalisation on wage inequality [35].

Finally, an important contribution is made by Castilho et al. [32], who look at trade liberalisation and its impact on inequality and poverty across Brazilian states from 1987 to 2005. They find that trade liberalisation significantly impacted inequality levels in Brazilian states. However, the direction of the impact differed between rural and urban areas—while trade liberalisation led to an increase in both inequality and poverty in urban areas, it led to a decrease in inequality in rural areas. As a possible explanation for this, Castilho et al. [32] point out that trade liberalisation in Brazil had a particularly pronounced impact on manufacturing sectors, which are typically set up in urban areas.

Summarising, despite a process where the income distribution has been improving over the last few decades, the issue of high income inequality remains an important feature of the Brazilian society. Furthermore, there are also important differences in levels of income inequality between Brazilian states. Various types of factors have been examined as possible drivers of income inequality but there is no clear consensus on their relative importance, and a substantial part of the evolution of income inequality remains unexplained. Against this background, and in line with existing evidence on other countries, we hypothesise that economic complexity may play an important role as a driver of income inequality in Brazil and that an empirical study on the relative importance of productive structures will generate important new insights into the process underlying the evolution of the income distribution in this country.

4. Data and Regression Model

In contrast to previous studies that have examined the relationship between economic complexity and income inequality in a cross-country framework, our empirical analysis is focused on estimating the impact of economic complexity on income inequality at the regional level within Brazil. Our motivation for doing so is that many factors that are incorporated into or closely linked with economic complexity are likely to be subject to substantial heterogeneity across regions within a country. For instance, there is growing evidence that industries are often spatially concentrated in agglomerations and that institutions and their impacts are subject to regional heterogeneity [18,19]. This implies that the relationship between economic complexity and income inequality will have a regional dimension, which is masked when using national level data.

The regression model specification that we use in the present analysis is based primarily on Hartmann et al. [15] and Castilho et al. [32]. We follow Hartmann et al. [15] in linking economic complexity to inequality. Castilho et al. [32] provide insight into the main factors that we need to consider as drivers of income inequality across Brazilian states. This leads to the following specification of the regression model:

$$y_{it} = \alpha_i + \beta_1 ECI_{it} + \beta_2 ECI_{it}^2 + \beta_3 GDPcap_{it} + \beta_4 GDPcap_{it}^2 + \gamma X'_{it} + \varepsilon_{it} \quad (1)$$

This model expressed in Equation (1) posits inequality y in state i and period t as a function of economic complexity (ECI) and its squared term (ECI^2), GDP per capita ($GDPcap$) and its squared term ($GDPcap^2$), a number of additional control variables captured in X , an idiosyncratic error term and state-specific effects α_i . The dataset that we composed contains 27 federative units (26 states and one federal district) with annual data for the

years 2002–2014. The time scope of the dataset is determined by the availability of the inequality indices. Table 1 presents all the variables that we use in this study with their data sources.

Table 1. Variables and data sources.

Variable Name	Description	Source
Gini	Gini coefficient	Institute of Applied Economic Research (IPEA)
Theil	Theil coefficient	Institute of Applied Economic Research (IPEA)
ECI	Economic Complexity Index (ECI)	Secretariat of Foreign Trade (SECEX) (*)
ECI ²	ECI squared	
GDPcap	Gross Domestic Product per capita (**)	Institute of Geography and Statistics (IBGE)
(GDPcap) ²	GDPcap squared	
Schooling	Average years of schooling of people aged 25 or older	Institute of Applied Economic Research (IPEA)
Population	Ln (state population)	Institute of Geography and Statistics (IBGE)
Skilled	Share of population aged 25 or older with 11 or more years of schooling	National Household Sample Survey (PNAD)
Semi-skilled	Share of population aged 25 or older with 4–10 years of schooling	National Household Sample Survey (PNAD)
White	% of workforce that is self-declared white	National Household Sample Survey (PNAD)
Informal	% of workforce without work contract or self-employed	National Household Sample Survey (PNAD)
Share_agri	% of workforce employed in Agriculture and Animal Farming	National Household Sample Survey (PNAD)
% Urban	% of economic active population in urban area	National Household Sample Survey (PNAD)
CNAE diversity	Number of unique industries	Annual Social Information Report (RAIS) (*)
CBO diversity	Number of unique occupations	Annual Social Information Report (RAIS) (*)

(*) Obtained from Dataviva at <http://legacy.dataviva.info/en/>. (**) Measured in 2010 Brazilian Reais.

As a dependent variable, we use two alternative indices of income inequality in the form of the Gini coefficient and the Theil coefficient. The Gini coefficient is the most widely used inequality measure, but has some shortcomings such as being sensitive to transfers at all income levels [36]. The Theil coefficient, on the other hand, can be more sensitive to variations in the lower and higher tail of the income distribution, but is less intuitive than the Gini coefficient. As such, it is common to consider both inequality measures for robustness. ECI and GDP per capita are our main variables of interest. We include both *ECI* and *ECI*² in the model to assess whether the relationship between economic complexity and inequality is characterised by non-linearity. We also include GDP per capita and its squared term, to test for the presence of a Kuznets curve relationship.

The ECI is calculated by taking the average complexity of the products that a region exports with international comparative advantage, weighed by the share of overall exports for that location. Product complexity is based on the concepts of diversity (the number of products that a region exports with comparative advantage) and ubiquity (the number of regions that export a given product with comparative advantage). The underlying idea is that more complex products are produced and exported by a more limited number of regions, requiring more productive knowledge. A region with a high ECI therefore produces a higher number of more complex products, which are produced by a limited number of regions (see [9,10]). The ECI data originates from the Secretariat of Foreign Trade (SECEX) and can be downloaded through DataViva, which is a large platform providing official social and economic data for Brazil at several regional levels. For each of the years, we z-transformed the ECI scores of the states and use the transformed variables in our empirical analysis.

Table 2 presents sample averages for the 27 federative units for the variables Gini, ECI and GDP per capita. São Paulo ranks first regarding economic complexity. Compared to the other states, the productive structure is characterised by a markedly higher level of complexity. Other states with a relatively high level of complexity include Rio de Janeiro, Amazonas, the Federal District and Acre. Among these states, the Federal District, São

Paulo and Rio de Janeiro also have the highest average levels of GDP per capita. The Federal District, Acre, Alagoas, Piauí and Paraíba are the states with the most unequal income distributions.

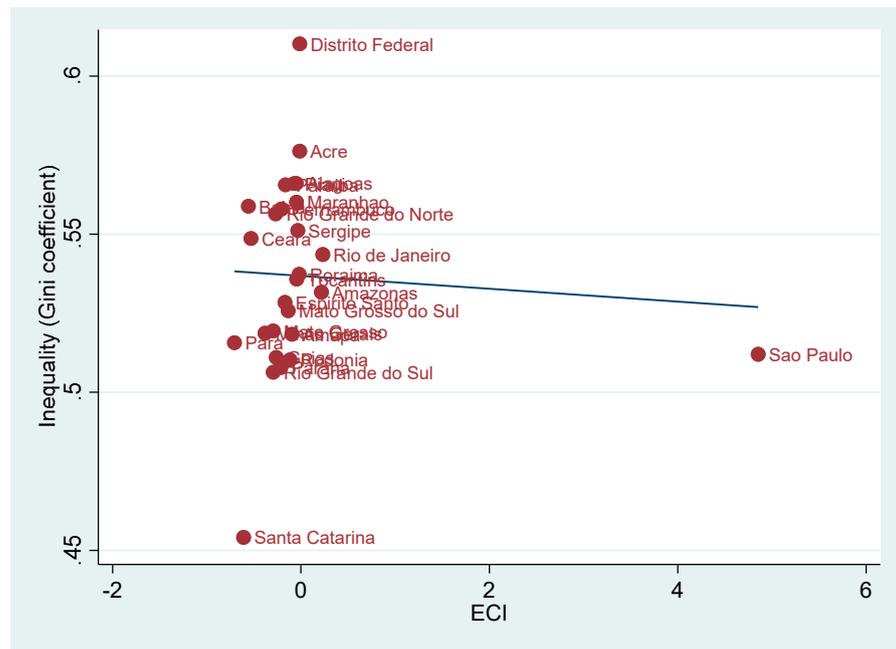
Table 2. Key variables of Brazilian states.

ECI Rank	State	Gini	ECI	GDP per Capita
1	São Paulo	0.512	4.8543	26.393
2	Rio de Janeiro	0.544	0.2321	23.436
3	Amazonas	0.532	0.2172	15.443
4	Distrito Federal	0.610	−0.0133	53.166
5	Acre	0.576	−0.0133	10.120
6	Roraima	0.537	−0.0198	12.782
7	Sergipe	0.551	−0.0340	10.329
8	Tocantins	0.536	−0.0456	10.431
9	Maranhão	0.560	−0.0468	6.147
10	Alagoas	0.566	−0.0499	7.027
11	Piauí	0.566	−0.0700	6.068
12	Amapá	0.518	−0.0955	11.388
13	Rondônia	0.510	−0.1170	12.522
14	Mato Grosso do Sul	0.526	−0.1359	15.766
15	Paraíba	0.566	−0.1650	7.361
16	Espírito Santo	0.528	−0.1696	20.280
17	Pernambuco	0.558	−0.2066	9.434
18	Paraná	0.508	−0.2147	18.308
19	Goiás	0.511	−0.2624	14.389
20	Rio Grande do Norte	0.556	−0.2692	9.412
21	Rio Grande do Sul	0.506	−0.2939	20.961
22	Mato Grosso	0.519	−0.2951	17.734
23	Minas Gerais	0.519	−0.3816	15.943
24	Ceará	0.549	−0.5306	8.041
25	Bahia	0.559	−0.5581	9.535
26	Santa Catarina	0.454	−0.6096	22.910
27	Pará	0.516	−0.7059	9.606

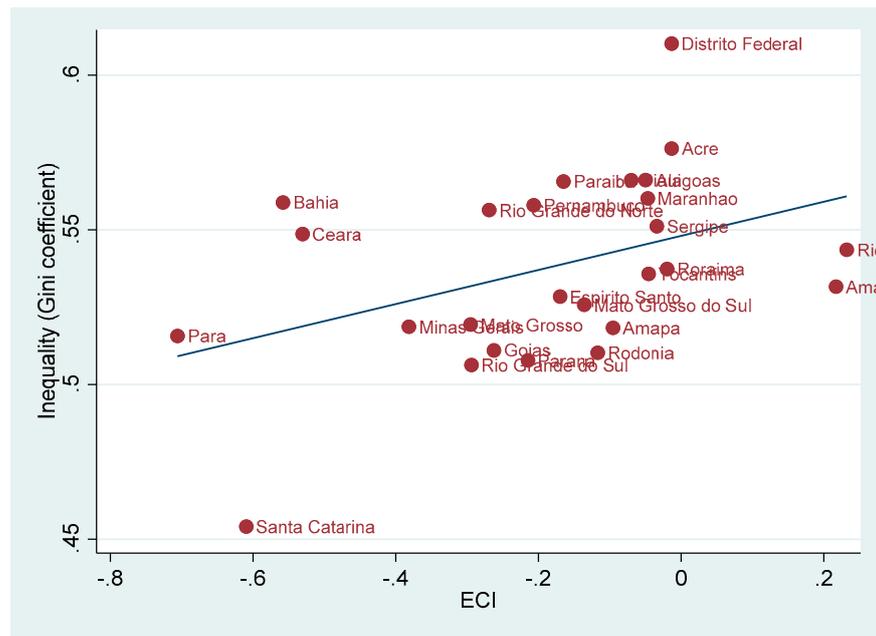
Gini, economic complexity index (ECI) and GDP per capita are state averages.

Figure 2 contains scatterplots between the average values of economic complexity and inequality measured by the Gini coefficient. Panel (a) shows the scatter plot for all states; the outlier state São Paulo can be clearly seen. Panel (b) shows the scatter plot where we omit São Paulo. According to the latter scatterplot, there is a positive association between the level of economic complexity of a state's economy and the level of inequality in the state.

Next, Figure 3 shows scatterplots between the average values of GDP per capita and inequality. The figure on panel (a) does not reveal any meaningful association, caused by the outlier state of the Federal District. Omitting the Federal District produces the scatterplot that is shown on panel (b). This scatterplot suggests that states with a higher level of GDP per capita are characterised by a lower level of inequality.

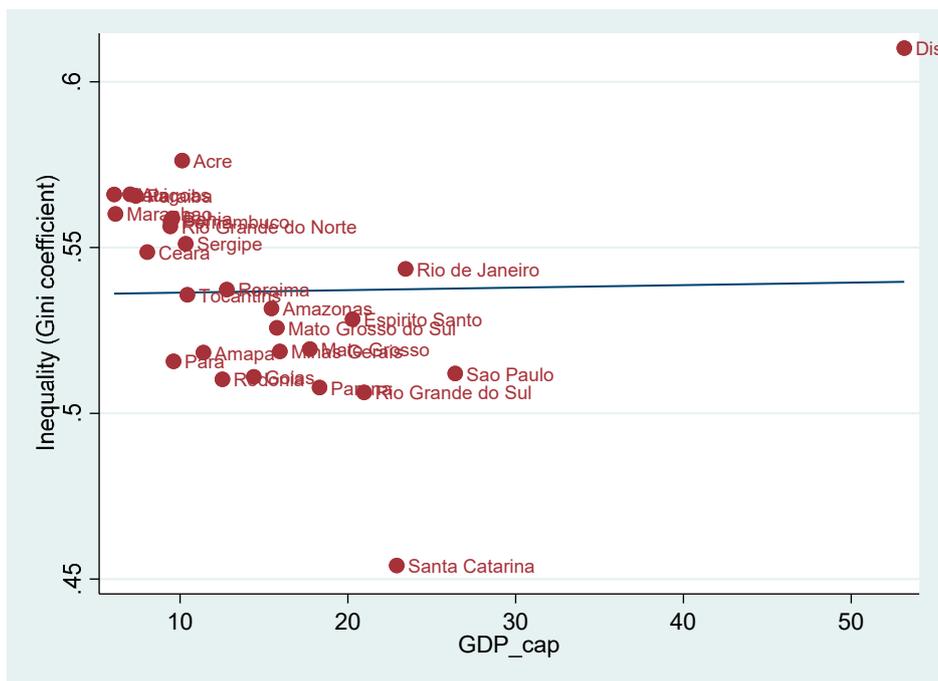


(a)

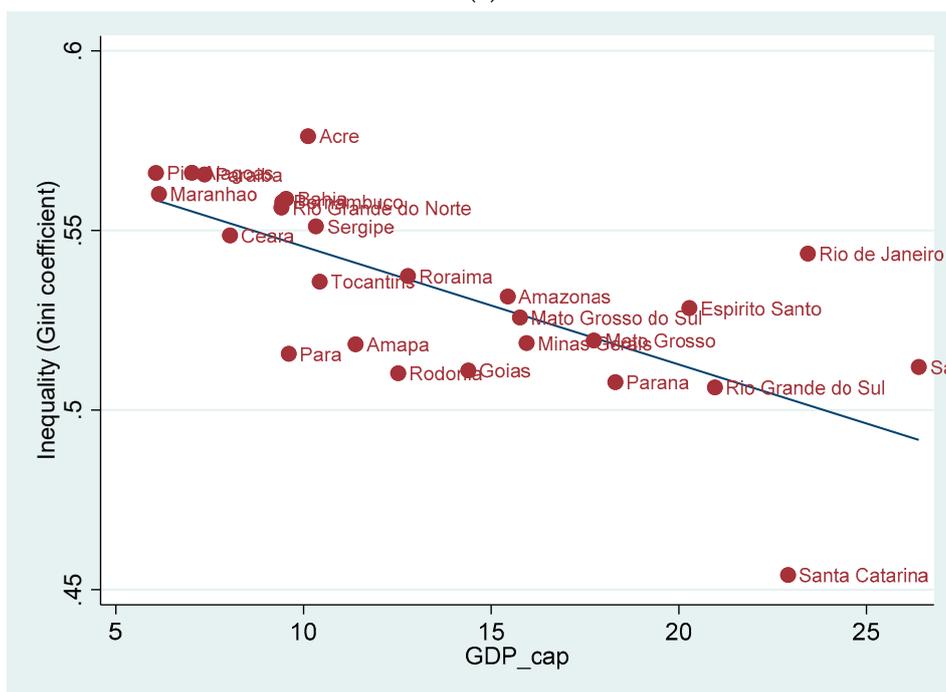


(b)

Figure 2. Scatterplots of economic complexity and inequality: (a) all states; (b) excluding São Paulo.



(a)



(b)

Figure 3. Scatterplots of GDP per capita and inequality; (a) all states; (b) excluding the Federal District

Other Control Variables

We control for the level of human capital to test whether states with a higher level of human capital are characterised by lower inequality. We use two different variables to capture human capital. One indicator is the overall level of education of a state’s population older than 24 years. The other indicator captures the share of highly skilled and medium skilled workers in a state’s labour force. We also include the size of the population across

states, given findings presented by Hartmann et al. [15] that population size is associated with inequality.

We include several variables following Castilho et al. [32]. One variable labelled “white” captures the percentage of a state’s workforce that declare themselves to be white; the expectation is that a higher percentage is associated with a lower degree of inequality. Another variable is the percentage of workers employed in the informal sector, which is expected to be positively associated with income inequality. We also include the relative importance of the agriculture sector in a state’s economic structure, given findings that it exercises a positive effect on poverty levels (see [32]). Finally, we add a variable capturing the degree of urbanisation across states to control for the difference in income inequality between states that have a more urban or rural character.

5. Empirical Findings

5.1. Baseline Model

We start our discussion of the empirical findings with the results that we obtained from estimating the baseline model, controlling for regional economic complexity, level of income per capita, schooling and population. Table 3 presents the findings for the baseline model.

Table 3. Economic complexity and inequality: findings from the baseline model.

Estimator	POLS	POLS	POLS	POLS	POLS	POLS	RE	RE	RE	RE
Sample	Full	Full	Full	Full	No SP	No SP	Full	Full	No SP	No SP
Dep. var.	GINI	GINI	GINI	THEIL	GINI	THEIL	GINI	THEIL	GINI	THEIL
<i>ECI</i>	0.05 *** (0.01)	0.04 *** (0.007)	0.045 *** (0.008)	0.12 *** (0.02)	0.026 * (0.015)	0.066 * (0.036)	0.02 ** (0.009)	0.08 *** (0.03)	0.02 (0.02)	0.06 (0.05)
<i>ECI</i> ²	−0.01 *** (0.002)	−0.01 *** (0.001)	−0.01 *** (0.001)	−0.03 *** (0.004)	−0.04 * (0.022)	−0.11 ** (0.05)	−0.005 ** (0.002)	−0.02 *** (0.006)	−0.014 (0.02)	−0.04 (0.04)
<i>GDP_CAP</i>		−0.32 *** (0.03)	−0.26 *** (0.03)	−0.73 *** (0.08)	−0.27 *** (0.03)	−0.75 *** (0.08)	−0.20 ** (0.08)	−0.73 *** (0.15)	−0.23 *** (0.07)	−0.78 *** (0.13)
<i>GDP_CAP</i> ²		0.06 *** (0.005)	0.05 *** (0.005)	0.14 *** (0.01)	0.05 *** (0.005)	0.14 *** (0.01)	0.04 ** (0.02)	0.14 *** (0.03)	0.044 ** (0.016)	0.14 *** (0.02)
Schooling			−0.02 *** (0.004)	−0.05 *** (0.01)	−0.02 *** (0.004)	−0.05 *** (0.01)	−0.02 *** (0.005)	−0.05 *** (0.01)	−0.02 *** (0.005)	−0.05 *** (0.01)
Population			−0.002 (0.002)	0.008 (0.005)	−0.002 (0.002)	0.009 * (0.005)	−0.006 (0.005)	0.003 (0.009)	−0.006 (0.005)	0.004 (0.009)
Constant	0.56 *** (0.003)	0.99 *** (0.04)	1.03 (0.04)	1.79 *** (0.11)	1.02 *** (0.04)	1.78 *** (0.11)	1.03 *** (0.14)	1.89 *** (0.21)	1.06 *** (0.13)	1.92 *** (0.19)
n	243	243	243	243	234	234	243	243	234	234
Adj. R2	0.13	0.48	0.53	0.56	0.53	0.56	0.42	0.55	0.44	0.54

POLS = pooled OLS; RE = random effects. No SP = Sao Paulo omitted. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The first set of estimated effects concerns findings from the regression model for the full set of states with pooled OLS (Ordinary Least Squares) and Gini or Theil as dependent variable. The relationship between inequality and economic complexity has an inverted U-shape, with an estimated positive effect of *ECI* and an estimated negative effect of its square term. The estimated effect of GDP per capita and its square term reveal a U-shaped relationship with inequality. The estimated effect of schooling is negative and significant, indicating that regions with higher levels of human capital are characterised by lower levels of inequality. The estimated effect of population is not statistically significant.

Next, we re-estimate the model on a restricted sample that omits the state São Paulo, an outlier due to its high level of economic complexity. The nature and significance of the estimated effects of income per capita, schooling and regional population are the same as with the full sample. The main difference is seen in the estimated effects of the *ECI* and its

square term. The nature of the estimated effects of these two variables remains the same, with a positive effect of *ECI* and a negative effect of the square term. The significance of the estimated effect of these two variables is lower, however, especially when using the Gini coefficient as dependent variable.

The main drawback of the pooled OLS estimations is that state effects are assumed to be part of the error term. The variables *ECI* and *GDPcap* vary over time to only a very limited degree, as shown in Figures A1 and A2 in Appendix A. This prevents us from estimating the regression model with a standard fixed effects estimation. With such an estimation, the variation across the states is wiped out and the limited variation of the variables over the years does not allow for an identification of the effect of these two variables. In order to capture the effects of time invariant characteristics in the estimations, we resort to estimating the regression model with a random effects specification, whereby we also cluster the standard errors at the state level.

The findings from the random effects estimations are largely similar to the results that we obtain with the pooled OLS estimations, especially when looking at the estimations with the full set of states. The main difference concerns the results from the estimations with the restricted sample. When São Paulo is omitted from the sample, the nature of the estimated effect of *ECI* and its squared term is similar to previous results but the significance of the effect of the two variables is affected.

5.2. Extended Model

The main findings from estimating the extended regression model are presented in Table 4. Looking at the full set of results, the inclusion of the additional control variables does not impact on the nature and significance of the estimated effects of economic complexity and GDP per capita. Replacing schooling with the indicators of the regional variation of skilled and semi-skilled workers produces results that indicate that inequality is lower in states with a relatively large presence of these types of workers. The other variable that carries a significant coefficient in several estimations is the share of informal workers in the regional workforce. States with a relatively high presence of informal workers are characterised by a higher level of inequality.

Turning to the estimated effects of economic complexity, a similar picture emerges as with the estimations of the baseline model. Overall, the findings indicate that the relationship between inequality and economic complexity follows an inverted U-shape. Again, the main difference between the pooled OLS and the random effects estimation is that the estimated significance of this relationship weakens when we control for time-invariant characteristics. However, the importance of examining the non-linear nature of the relationship is further indicated by the findings in the last two columns that contain the results when we omit the squared term of *ECI* from the model. When the model only includes *ECI*, the results suggest that economic complexity generates a significant positive effect on inequality. The findings from the various specifications of the model indicate that it is likely that the effect of economic complexity turns negative at some stage, suggesting that ongoing increases and improvements of the productive structure of states will generate improvements in the distribution of income. This important aspect of the relationship between economic complexity and inequality is missed when the model only captures the effect of the level of *ECI*.

Table 4. Economic complexity and inequality: findings from the extended model.

Estimator	POLS	POLS	POLS	POLS	RE	RE	RE	RE	RE	RE
Sample	Full	No SP	Full	No SP	Full	No SP	Full	No SP	No SP	No SP
Dep. var.	GINI	GINI	THEIL	THEIL	GINI	GINI	THEIL	THEIL	GINI	THEIL
<i>ECI</i>	0.03 *** (0.01)	−0.02 (0.02)	0.07 ** (0.03)	−0.035 (0.05)	0.03 *** (0.008)	0.015 (0.02)	0.08 *** (0.03)	0.02 (0.05)	0.03 *** (0.01)	0.09 *** (0.03)
<i>ECI</i> ²	−0.007 ** (0.003)	−0.08 *** (0.02)	−0.016 ** (0.007)	−0.19 *** (0.06)	−0.006 *** (0.002)	−0.025 (0.02)	−0.017 *** (0.005)	−0.10 ** (0.05)		
<i>GDPcap</i>	−0.16 ** (0.06)	−0.14 *** (0.05)	−0.56 *** (0.14)	−0.51 *** (0.14)	−0.19 ** (0.08)	−0.21 *** (0.07)	−0.70 *** (0.16)	−0.69 *** (0.15)	−0.21 ** (0.08)	−0.74 *** (0.16)
<i>GDPcap</i> ²	0.04 *** (0.01)	0.04 *** (0.008)	0.11 *** (0.02)	0.11 *** (0.02)	0.04 ** (0.02)	0.04 *** (0.01)	0.13 *** (0.03)	0.13 *** (0.02)	0.04 ** (0.015)	0.14 *** (0.03)
Skilled	−0.005 *** (0.001)	−0.005 *** (0.001)	−0.013 *** (0.003)	−0.013 *** (0.002)	−0.002 *** (0.0007)	−0.003 *** (0.0007)	−0.007 *** (0.002)	−0.008 *** (0.002)	−0.002 *** (0.0007)	−0.007 ** (0.002)
Semi-skilled	−0.003 *** (0.0006)	−0.003 *** (0.0006)	−0.008 *** (0.002)	−0.009 *** (0.002)	−0.0002 (0.0008)	−0.0005 (0.0007)	−0.014 (0.002)	−0.003 (0.002)	−0.0002 (0.0008)	−0.0015 (0.002)
White	−0.003 (0.003)	−0.001 (0.002)	0.0004 (0.0007)	0.0002 (0.0007)	0.0005 (0.0004)	0.0003 (0.0004)	0.0007 (0.0008)	0.0005 (0.0008)	0.0004 (0.0004)	0.0007 (0.0008)
Informal	0.0009 * (0.0005)	0.0009 (0.0006)	0.001 (0.001)	0.0014 (0.0013)	0.002 *** (0.0007)	0.0016 ** (0.007)	0.003 ** (0.0015)	0.0025 * (0.0015)	0.0017 ** (0.0007)	0.003 * (0.0016)
Share_agri	0.06 (0.09)	−0.01 (0.09)	0.31 (0.20)	0.15 (0.22)	0.03 (0.09)	0.028 (0.09)	0.24 (0.22)	0.19 (0.22)	0.05 (0.09)	0.28 (0.22)
% Urban	0.001 ** (0.0005)	0.0009 ** (0.0004)	0.0015 (0.01)	0.001 (0.001)	0.001 (0.0001)	0.00009 (0.0004)	−0.0003 (0.001)	−0.0001 (0.001)	0.001 (0.004)	−0.0003 (0.001)
Constant	0.79 *** (0.08)	0.77 *** (0.08)	1.58 *** (0.20)	1.54 *** (0.21)	0.70 *** (0.13)	0.74 *** (0.12)	1.53 *** (0.28)	1.56 (0.26)	0.73 *** (0.13)	1.58 *** (0.28)
n	243	234	243	234	243	234	243	234	234	234
Adj. R2	0.61	0.63	0.60	0.60	0.48	0.53	0.56	0.57	0.48	0.55

POLS = pooled OLS; RE = random effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

5.3. Rural–Urban and Level of Development

In order to get a better understanding of the conditions under which the relationship between economic complexity and inequality is most prominent, we examine the impact of the levels of urbanisation and overall development of Brazilian states. To do so, we estimate the regression model on sub-samples of states, whereby we separate the states according to whether their levels of urbanisation and GDP per capita are below or above the sample median values. In the extended model underlying Table 4, we control for the effects of the level of urbanisation and the share of agriculture in the regional economies on inequality. By estimating the model separately for urban and rural states, we allow for the coefficients of all the variables to differ between these two types of state. The reasoning for separating states according to their level of GDP per capita is similar. The findings from estimating the model on the sub-samples of states are shown in Tables 5 and 6.

Table 5. Economic complexity and inequality: rural–urban.

Estimator	RURAL				URBAN			
	POLS	RE	POLS	RE	POLS	RE	POLS	RE
Dep. var.	GINI	GINI	THEIL	THEIL	GINI	GINI	THEIL	THEIL
<i>ECI</i>	−0.01 (0.03)	−0.003 (0.02)	0.01 (0.06)	0.02 (0.06)	0.09 *** (0.01)	0.05 *** (0.01)	0.19 *** (0.03)	0.12 *** (0.03)
<i>ECI</i> ²	−0.08 * (0.046)	−0.02 (0.02)	−0.16 (0.09)	−0.08 (0.07)	−0.02 *** (0.002)	−0.01 *** (0.002)	−0.04 *** (0.006)	−0.03 *** (0.007)
<i>GDPcap</i>	0.11 (0.13)	−0.14 (0.13)	0.10 (0.34)	−0.37 (0.32)	−0.44 *** (0.07)	−0.37 *** (0.11)	−0.98 *** (0.18)	−0.87 *** (0.25)
<i>GDPcap</i> ²	−0.03 (0.03)	0.02 (0.02)	−0.04 (0.07)	0.05 (0.06)	0.08 *** (0.01)	0.07 *** (0.02)	0.18 *** (0.03)	0.16 *** (0.04)
Schooling	−0.016 *** (0.005)	−0.01 ** (0.005)	−0.05 ** (0.017)	−0.04 ** (0.015)	−0.03 *** (0.005)	−0.04 *** (0.008)	−0.07 *** (0.01)	−0.08 *** (0.02)
Population	−0.002 (0.006)	−0.008 (0.006)	0.01 (0.01)	0.0008 (0.01)	0.002 (0.002)	−0.002 (0.003)	0.015 *** (0.004)	0.007 (0.007)
Constant	0.58 *** (0.16)	0.97 *** (0.16)	0.80 * (0.43)	1.46 *** (0.43)	1.30 *** (0.12)	1.29 *** (0.19)	2.17 *** (0.29)	2.15 *** (0.41)
n	121	121	121	121	122	122	122	122
Adj. R2	0.45	0.41	0.48	0.46	0.71	0.62	0.59	0.55

POLS = pooled OLS; RE = random effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 6. Economic complexity and inequality: level of development.

Estimator	Low GDP				High GDP			
	POLS	RE	POLS	RE	POLS	RE	POLS	RE
Dep. var.	GINI	GINI	THEIL	THEIL	GINI	GINI	THEIL	THEIL
<i>ECI</i>	−0.08 (0.06)	−0.035 (0.07)	−0.19 (0.17)	−0.13 (0.18)	0.08 *** (0.02)	0.025 ** (0.01)	0.17 *** (0.03)	0.08 *** (0.03)
<i>ECI</i> ²	−0.14 ** (0.07)	−0.06 (0.06)	−0.38 * (0.20)	−0.27 (0.20)	−0.017 *** (0.003)	−0.0048 * (0.0025)	−0.04 *** (0.007)	−0.018 *** (0.007)
<i>GDPcap</i>	0.18 (0.26)	−0.11 (0.22)	0.04 (0.66)	−0.48 (0.65)	−0.65 *** (0.11)	−0.69 *** (0.14)	−1.57 *** (0.22)	−1.67 *** (0.28)
<i>GDPcap</i> ²	−0.05 (0.06)	0.013 (0.05)	−0.025 (0.16)	0.09 (0.15)	0.12 *** (0.02)	0.12 *** (0.02)	0.27 *** (0.03)	0.28 *** (0.04)
Schooling	−0.024 *** (0.008)	−0.017 ** (0.009)	−0.07 *** (0.02)	−0.06 ** (0.02)	−0.03 *** (0.008)	−0.03 *** (0.008)	−0.07 *** (0.01)	−0.07 *** (0.016)
Population	−0.007 (0.006)	−0.008 (0.006)	−0.001 (0.02)	0.0002 (0.02)	0.005 (0.003)	0.002 (0.007)	0.02 ** (0.007)	0.018 (0.013)
Constant	0.61 ** (0.27)	0.94 *** (0.24)	1.08 ** (0.69)	1.59 *** (0.64)	1.58 *** (0.19)	1.73 *** (0.22)	2.99 *** (0.41)	3.21 *** (0.46)
n	124	124	124	124	119	119	119	119
Adj. R2	0.41	0.36	0.44	0.43	0.71	0.49	0.62	0.55

POLS = pooled OLS; RE = random effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The separate results for states with a more rural or urban character show clear differences. The main variable lowering inequality in rural states is the level of human capital, as captured by education. It is the urban states where the non-linear relationship between economic complexity and inequality materialises. Similar to rural states, the estimated

effect of human capital on inequality is significant and negative in the urban states, which are states where the GDP per capita variables also carry significant coefficients.

The findings that we obtain from estimating the regression model separately for states with a relatively low or high level of development are very similar. The effects of economic complexity are prominent in states with a relatively high level of GDP per capita, as are the effects of GDP per capita. In states with a relatively low level of income, there is some evidence that economic complexity does lower inequality, as evidenced by the estimated significant negative effect of the squared ECI variable in the pooled OLS estimations. The estimated effect of schooling on inequality is significant and negative in both low- and high-income states.

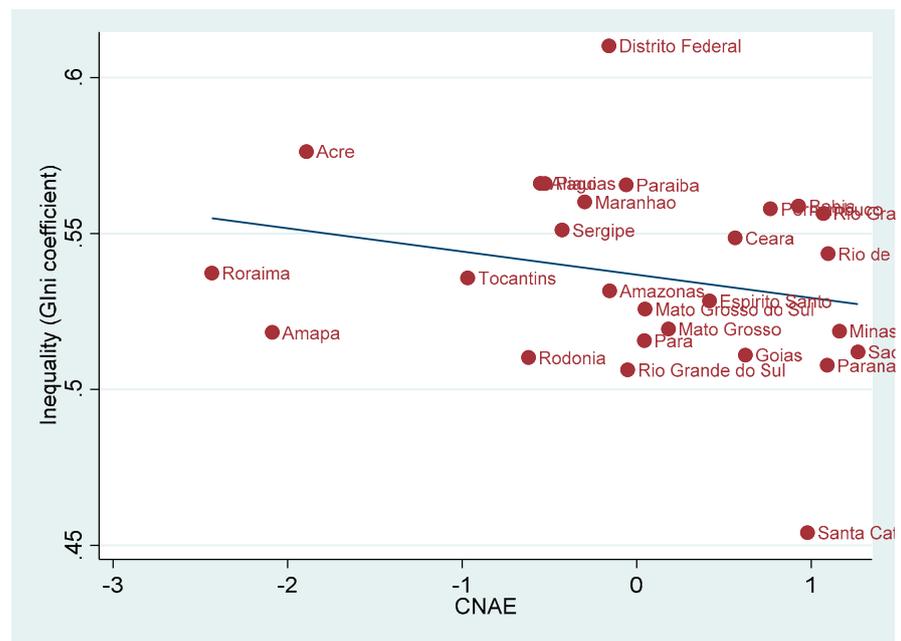
5.4. Industry and Occupation Diversity

So far, the findings clearly confirm that economic complexity is an important driver of inequality in Brazilian states. However, it may be that the *ECI* captures only part of the effect of the productive structure of the regions. The *ECI* is measured with international export data, taking into account international comparative advantages. This means that industries that produce for the domestic market are not captured by economic complexity. While for national-level data this seems like a reasonable assumption, when using data for a set of regions within a country this may be less so.

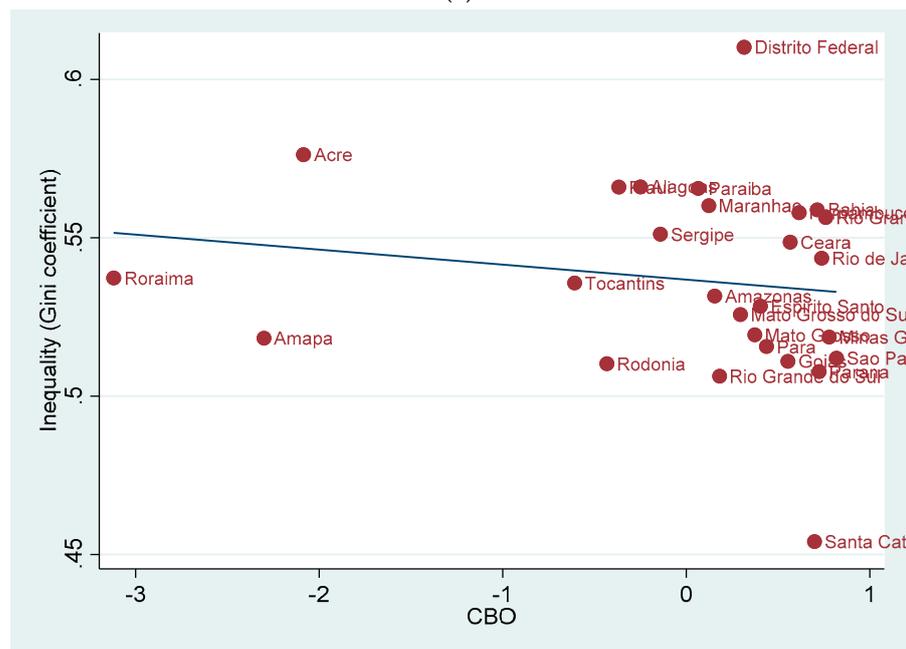
One reason why the complexity of the productive structure may not be fully captured by the *ECI* when using regional data is that a region may be involved in the production of complex products for the domestic market, whilst producing less complex products for international markets. Second, it is well known that economic activity tends to spatially concentrate within countries to benefit from agglomeration economies [37,38]. This is also the case in Brazil, where exporting activities are concentrated in a limited number of regions [32]. Other states may incorporate complex intermediate products that they supply to the exporting states. As the *ECI* is based on international export data, it will assign complexity only to the states that export products, without taking into account that part of the complexity of these states is linked to complexity of economic activity in other states that provide the intermediate products.

In an attempt to widen the measurement of the productive structure of the economic activities of the Brazilian states, we use two additional variables. One variable, labelled *CNAE* diversity (as it follows the Brazilian National Classification of Economic Activities), captures the level of industry diversity of a state, measured by the number of industries in the states. The second variable, labelled *CBO* diversity (named in line with the Brazilian Occupational Classification), is measured as the number of occupations in a state. In order to facilitate their incorporation into the regression model, we z-transform both these variables.

To obtain an impression of the nature of the relationship between the diversity variables and income inequality, we create scatterplots between state averaged Gini and *CNAE* Diversity and *CBO* Diversity, as shown in Figure 4. Both scatterplots show a negative association between industry diversity and state level inequality, suggesting that this dimension of the regional productive structure may also generate an impact on inequality.



(a)



(b)

Figure 4. Scatterplots of economic diversity measures and inequality: (a) industry diversity (CNAE); (b) occupation diversity (CBO)

To identify the effect of diversity in a multivariate setting and to assess whether the effect of economic complexity is robust to the inclusion of this dimension of productive structure, we augment the extended regression model with the indicators of industry and occupation diversity. The results from estimating several specifications are shown in Tables 7 and 8. Table 7 contains the results from the pooled OLS estimator, using the two alternative inequality indicators as dependent variable. In all four columns, the estimated effects of both industry diversity and occupation diversity are significant. In contrast to the scatterplot in Figure 4, the estimated effect of industry diversity is positive, suggesting that regions with a high presence of different industries are characterised by a higher level of

inequality. In contrast, the estimated effect of occupation diversity indicates that regions with a high number of occupations have a lower level of inequality. Importantly, the inclusion of the two diversity variables does not affect the estimated effects of the economic complexity variables, which are the same as in the previous tables. This indicates that, in addition to economic complexity, the diversity of economic activity across states is also an important driver of income inequality in Brazilian states.

Table 7. Economic complexity, industry and occupation diversity: pooled OLS estimations.

Estimator	POLS	POLS	POLS	POLS
Sample	Full	No SP	Full	No SP
Dep. var.	GINI	GINI	THEIL	THEIL
<i>ECI</i>	0.04 *** (0.007)	−0.006 (0.01)	0.10 *** (0.02)	0.01 (0.03)
<i>ECI</i> ²	−0.008 *** (0.002)	−0.07 *** (0.02)	−0.03 *** (0.004)	−0.16 *** (0.05)
<i>GDPcap</i>	−0.22 *** (0.04)	−0.20 *** (0.04)	−0.71 *** (0.11)	−0.65 *** (0.12)
<i>GDPcap</i> ²	0.05 *** (0.006)	0.05 *** (0.006)	0.14 *** (0.02)	0.13 *** (0.02)
Skilled	−0.005 *** (0.0007)	−0.005 *** (0.0007)	−0.012 *** (0.002)	−0.012 *** (0.002)
Semi_skilled	−0.0025 *** (0.0007)	−0.003 *** (0.0007)	−0.005 ** (0.002)	−0.006 *** (0.02)
White	−0.0002 (0.0016)	−0.0002 (0.0016)	−0.0002 (0.0004)	−0.0003 (0.0004)
Informal	0.0007 (0.00046)	0.0008 * (0.00046)	0.001 (0.001)	0.0014 (0.0012)
Share_agri	0.09 * (0.05)	0.03 (0.06)	0.33 ** (0.14)	0.19 (0.16)
% Urban	0.0008 ** (0.0004)	0.0007 * (0.00039)	0.0009 (0.001)	0.0007 (0.001)
CNAE Diversity	0.02 *** (0.005)	0.02 *** (0.005)	0.06 *** (0.015)	0.06 *** (0.015)
CBO Diversity	−0.02 *** (0.005)	−0.02 * (0.005)	−0.05 *** (0.016)	−0.044 *** (0.015)
Constant	0.88 *** (0.07)	0.85 *** (0.07)	1.76 *** (0.20)	1.70 *** (0.21)
n	243	234	243	234
Adj. R2	0.64	0.65	0.62	0.62

POLS = pooled OLS; RE = random effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The next set of findings is obtained with the random effects estimator, presented in Table 8. Overall, the findings are in line with the pooled OLS results. Industry diversity generates a positive impact on inequality and occupation diversity creates a negative effect on inequality. Again, the estimated effect of economic complexity is similar to the previous tables.

Table 8. Economic complexity, industry and occupation diversity: Random effects estimations.

Estimator	RE	RE	RE	RE	RE	RE	RE	RE
Sample	Full	No SP	Full	No SP	Rural	Urban	Low GDP	High GDP
Dep. var.	GINI	GINI	THEIL	THEIL	GINI	GINI	GINI	GINI
<i>ECI</i>	0.03 *** (0.008)	0.018 (0.018)	0.10 *** (0.02)	0.06 (0.045)	0.08 ** (0.04)	0.035 *** (0.012)	0.07 (0.009)	0.02 ** (0.009)
<i>ECI</i> ²	−0.006 *** (0.0015)	−0.02 (0.018)	−0.02 *** (0.005)	−0.08 * (0.048)	0.06 (0.05)	−0.007 *** (0.002)	0.06 (0.11)	−0.005 ** (0.0018)
<i>GDPcap</i>	−0.21 *** (0.08)	−0.23 *** (0.07)	−0.80 *** (0.16)	−0.80 *** (0.16)	−0.07 (0.10)	−0.33 ** (0.15)	−0.05 (0.26)	−0.37 *** (0.12)
<i>GDPcap</i> ²	0.04 *** (0.016)	0.05 *** (0.013)	0.15 *** (0.03)	0.15 *** (0.03)	0.0008 (0.02)	0.06 *** (0.02)	0.0006 (0.06)	0.07 *** (0.02)
Skilled	−0.0024 *** (0.0007)	−0.0025 *** (0.0007)	−0.007 *** (0.002)	−0.007 *** (0.002)	−0.004 *** (0.0009)	−0.001 (0.001)	−0.003 *** (0.001)	−0.005 *** (0.001)
Semi_skilled	−0.0001 (0.0008)	−0.0002 (0.0007)	−0.004 (0.02)	−0.0009 (0.002)	−0.003 *** (0.0008)	0.001 * (0.007)	−0.002 ** (0.0008)	−0.003 *** (0.001)
White	−0.0004 (0.0004)	0.0002 (0.0004)	0.00016 (0.0007)	−0.0005 (0.0007)	−0.0009 (0.004)	0.0007 (0.0006)	0.0009 (0.0006)	0.0005 (0.0003)
Informal	0.0017 ** (0.0007)	0.0016 ** (0.0007)	0.003 * (0.0017)	0.0026 (0.0016)	−0.0009 *** (0.0003)	0.003 *** (0.001)	−0.0002 (0.0008)	0.003 *** (0.0008)
Share_agri	0.06 (0.09)	0.06 (0.09)	0.26 (0.17)	0.23 (0.19)	0.25 *** (0.09)	0.01 (0.12)	0.36 ** (0.18)	−0.06 (0.11)
% Urban	−0.00006 (0.0004)	−0.0002 (0.0003)	−0.0006 (0.01)	−0.0005 (0.001)	0.003 *** (0.0006)	−0.0008 (0.001)	0.0003 (0.0006)	0.002 *** (0.0006)
CNAE Diversity	0.015 * (0.08)	0.017 ** (0.008)	0.06 *** (0.02)	0.06 *** (0.018)	0.05 *** (0.009)	−0.003 (0.02)	0.04 *** (0.01)	−0.001 (0.01)
CBO Diversity	−0.016 ** (0.008)	−0.017 ** (0.008)	−0.05 *** (0.018)	−0.05 *** (0.018)	−0.045 *** (0.01)	0.005 (0.01)	−0.04 *** (0.01)	0.006 (0.01)
Constant	0.74 *** (0.14)	0.78 *** (0.13)	1.65 *** (0.296)	1.68 *** (0.29)	0.76 *** (0.13)	0.81 *** (0.32)	0.73 ** (0.30)	0.91 *** (0.22)
n	243	234	243	234	121	122	124	119
Adj. R2	0.50	0.54	0.60	0.60	0.65	0.50	0.50	0.77

POLS = pooled OLS; RE = random effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

To further examine the conditions under which diversity impacts inequality, we also estimate the regression model on subsamples of states, separating states according to their level of urbanisation and their overall level of development. The last four columns in Table 8 present the results for the subsamples. The findings show a structural difference between the effects of diversity and economic complexity on inequality. The two diversity variables are significantly associated with inequality in states with a below median level of urbanisation or level of development. Their estimated effect is statistically insignificant in states with a relatively high level of urbanisation or development. In contrast, the effect of economic complexity is prominent only in states that have levels of urbanisation and development above the median values. This suggests that diversity is a particularly important factor influencing the distribution of income in states at relatively early stages of development, whereas economic complexity takes over this role in states that have surpassed a certain level of development.

6. Summary and Conclusions

With the introduction of the concept of economic complexity, the research strand investigating the impact of productive structures on structural transformation and eco-

economic growth has received an important impetus. In particular, there is growing evidence that countries that are involved in the production and export of complex products are characterised by higher average growth rates and that increases in economic complexity lead to faster economic growth.

Most recently, this research strand has started to widen its scope by investigating the impact of economic complexity on the income distribution within countries. Economic complexity is linked to a multitude of factors including factor endowments, knowledge, the availability of job opportunities and the quality of institutions. This suggests that more complex productive structures will exercise positive effects on the distribution of income, generating the hypothesis that economic complexity is negatively associated with income inequality. Given the small number of studies that look at this relationship and the heterogeneity of the available evidence, more research into this relationship is urgently required.

The purpose of this paper is to contribute to and extend upon this emerging research strand, adding to the literature in the following three ways. First, in contrast to existing studies that look at the relationship between economic complexity and income inequality in cross-country frameworks, our study investigates this at the regional level in Brazil. Our focus on the regional dimension is linked with the notion that many of the factors that are directly connected to economic complexity can be subject to substantial heterogeneity across regions, warranting empirical research on the impact of regional productive structures on the income distribution of regions.

Second, in line with the concept of the Kuznets curve, we examine whether the relationship between economic complexity and income inequality is characterised by non-linearity. Initial increases of economic complexity may disproportionately benefit high-skilled workers, generating increased inequality. With ongoing increases of economic complexity, other components of productive structures that lower inequality may become more important, resulting in a negative impact on income inequality at higher levels of complexity.

Third, Brazil is a country that—whilst having a level of economic development that classifies it as a lower- or upper-middle income—is facing high levels of poverty and inequality. As such, it is representative of a substantial group of countries in the world economy that find themselves in a similar situation and whose governments are searching for ways to implement policies that promote economic development whilst also impacting levels of poverty and inequality.

Using panel data for 27 federative units for the period 2002–2014, we obtain evidence in support of the notion that the relationship between economic complexity and income inequality has an inverted U-shape. Controlling for a number of drivers of income inequality, our pooled OLS estimations show a significant negative effect of economic complexity and a significant positive effect of its squared term on two different measures of income inequality. In extension, given the low variability of the main variables over the years, we rely on random effects estimations to capture time-invariant effects. With this estimator, although the statistical significance of the estimated effects of economic complexity is affected in some specifications, the nature of the estimated effect of the two economic complexity variables remains the same, in further support of the inverted U-shaped nature of the relationship. Other factors that play an important role as drivers of income inequality in Brazilian states include GDP per capita, human capital and the regional importance of the informal sector.

Next, our findings show that the effect of economic complexity on income distribution is affected by the level of development of the regions. When estimating the regression models separately for regions with low or high levels of urbanisation or GDP per capita, the non-linear relationship between economic complexity and income inequality materialises in those regions characterised by relatively high levels of these two indicators. An explanation for this finding is that a certain level of development needs to be reached before the regional productive structure starts to exercise meaningful effects on income inequality.

The indicator of economic complexity is based on international exports by regions based on revealed comparative advantage. This can pose a problem when using regional data, as the indicator does not capture inter-regional trade within a country. It may be that regions supply (complex) intermediate inputs to other regions that in turn export products to international markets. To test whether other characteristics of regional productive structures are important, we augment the regression models with two variables that capture the level of regional industry and occupation diversity. Our findings indicate that these features of regional productive structures are important, as they are—independently from economic complexity—significantly impacting on income inequality. Separate estimations for regions with a relatively low or high level of development show that, whereas economic complexity is an important driver of income inequality in the more developed states, industry and occupation diversity are important in states with a lower level of development.

In conclusion, our findings for Brazil confirm that economic complexity has an important impact the income distribution. The inverted U-shape of the relationship between economic complexity and income inequality indicates that governments can use policymaking to promote economic development and economic complexity, as increasingly complex productive structures will lead to beneficial effects on the income distribution. Of course, more research is needed to obtain further evidence on this relationship and on the conditions that affect the impact of economic complexity on income distributions.

In extension, we believe that further research will benefit from incorporating the regional dimension of the impact of economic complexity, given that productive structures and their determinants can be subject to substantial variation across regions within countries. Such research will also benefit from considering a wider interpretation of the concept of productive structures beyond levels of economic complexity. Indicators of economic complexity are based on levels of international trade according to revealed comparative advantages. When conducting a regional analysis, such indicators do not capture the economic complexity of regions that are less involved in international trade but supply intermediate products to other regions that are more trade intensive. More work is therefore needed to design indicators of economic complexity that incorporate the relative importance of indirect exports by such regions to further improve the identification of the impact of regional productive structures on income inequality at the subnational level.

Author Contributions: Conceptualisation: M.B.M., J.S.; methodology: M.B.M., J.S., J.A.J., formal analysis: M.B.M., J.S., J.A.J.; writing—original draft preparation: M.B.M., J.S., J.A.J.; project administration M.B.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analysed in this study. The data can be found in the following links: <http://legacy.dataviva.info/en/> and <http://www.ipeadata.gov.br/Default.aspx>.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

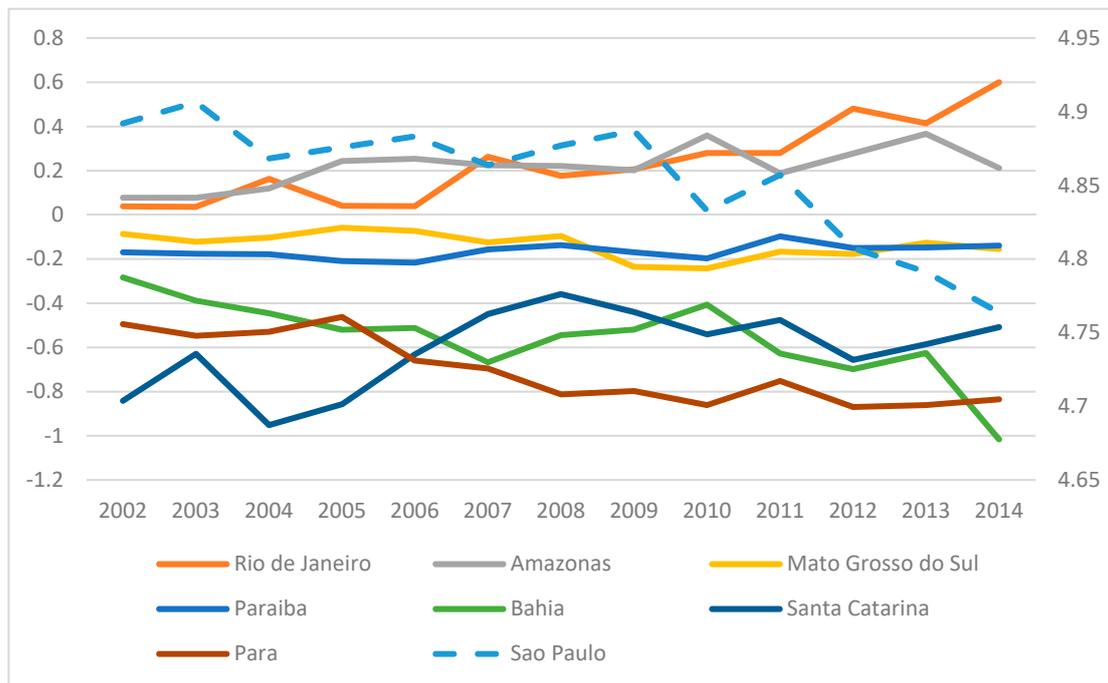


Figure A1. Economic Complexity Index 2002–2014: selected states. Note on Y-axis: right-hand side São Paulo; left-hand side all other states.

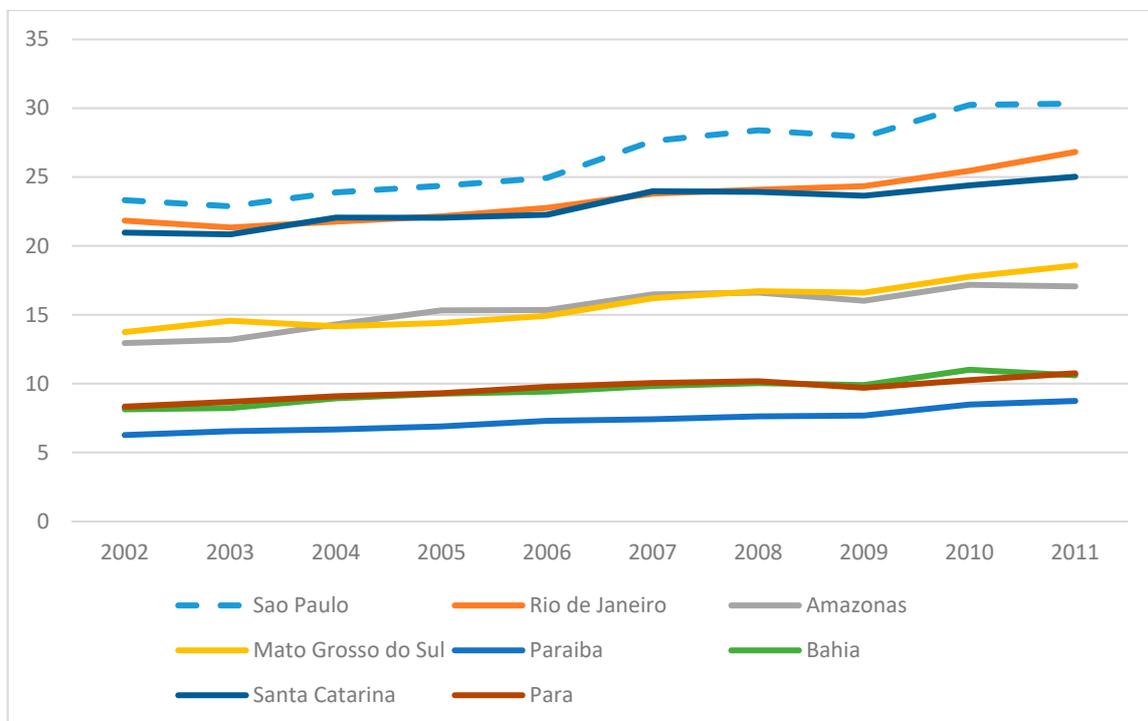


Figure A2. GDP per capita 2002–2011: selected states.

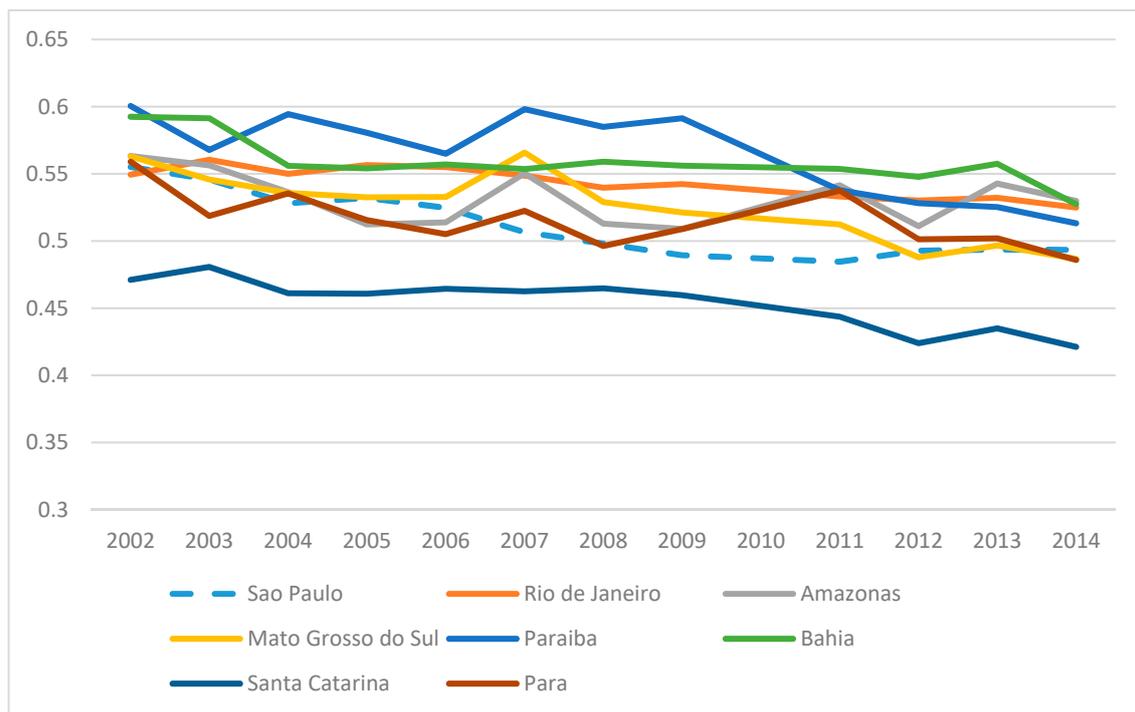


Figure A3. Income inequality (Gini coefficient) 2002–2014: selected states.

References

- Rosenstein-Rodan, P.N. Problems of industrialisation of eastern and south-eastern Europe. *Econ. J.* **1943**, *53*, 202–211. [\[CrossRef\]](#)
- Prebisch, R. The economic development of Latin America and its principal problems. *Econ. Bull. Lat. Am.* **1962**, *7*, 1–22.
- Singer, H.W. The distribution of gains between borrowing and investing countries. *Am. Econ. Rev.* **1950**, *40*, 473–485.
- Naudé, W.; Szirmai, A. *The Importance of Manufacturing in Economic Development: Past, Present and Future Perspectives*; UNU-MERIT Working Papers: Maastricht, The Netherlands, 2012.
- Dervis, K.; Page, J.M., Jr. Industrial policy in developing countries. *J. Comp. Econ.* **1984**, *8*, 436–451. [\[CrossRef\]](#)
- Lederman, D.; Klinger, B. *Diversification, Innovation, and Imitation Inside the Global Technological Frontier*; Policy Research Working Paper Series, No. 3872; The World Bank: Washington, DC, USA, 2006.
- Lectard, P.; Rougier, E. Can developing countries gain from defying comparative advantage? Distance to comparative advantage, export diversification and sophistication, and the dynamics of specialization. *World Dev.* **2018**, *102*, 90–110. [\[CrossRef\]](#)
- Frenken, K.; Van Oort, F.; Verburg, T. Related variety, unrelated variety and regional economic growth. *Reg. Stud.* **2007**, *41*, 685–697. [\[CrossRef\]](#)
- Hidalgo, C.A.; Hausmann, R. The building blocks of economic complexity. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10570–10575. [\[CrossRef\]](#)
- Hausmann, R.; Hidalgo, C.A.; Bustos, S.; Coscia, M.; Simoes, A.; Yildirim, M.A. *The Atlas of Economic Complexity: Mapping Paths to Prosperity*; MIT Press: Cambridge, MA, USA, 2014; ISBN 978-0-262-52542-8.
- Hartmann, D.; Jara-Figueroa, C.; Guevara, M.R.; Simoes, A.; Hidalgo, C.A. The structural constraints of income inequality in Latin America. *Integr. Trade J.* **2016**, *40*, 70–85.
- Hausmann, R.; Hwang, J.; Rodrik, D. What you export matters. *J. Econ. Growth* **2007**, *12*, 1–25. [\[CrossRef\]](#)
- Pugliese, E.; Chiarotti, G.L.; Zaccaria, A.; Pietronero, L. Complex economies have a lateral escape from the poverty trap. *PLoS ONE* **2017**, *12*, e0168540. [\[CrossRef\]](#)
- Hartmann, D.; Bezerra, M.; Lodolo, B.; Pinheiro, F.L. International trade, development traps, and the core-periphery structure of income inequality. *Economia* **2020**, *21*, 255–278. [\[CrossRef\]](#)
- Hartmann, D.; Guevara, M.R.; Jara-Figueroa, C.; Aristarán, M.; Hidalgo, C.A. Linking economic complexity, institutions and income inequality. *World Dev.* **2017**, *93*, 75–93. [\[CrossRef\]](#)
- Lee, K.K.; Vu, T.V. Economic complexity, human capital and income inequality: A cross-country analysis. *Jpn. Econ. Rev.* **2020**, *71*, 695–718. [\[CrossRef\]](#)
- Chu, L.K.; Hoang, D.P. How does economic complexity influence income inequality? New evidence from international data. *Econ. Anal. Policy* **2020**, *68*, 44–57. [\[CrossRef\]](#)
- Ketterer, T.D.; Rodríguez-Pose, A. Institutions vs. ‘first-nature’ geography: What drives economic growth in Europe’s regions? *Pap. Reg. Sci.* **2018**, *97*, S25–S62. [\[CrossRef\]](#)

19. Charron, N.; Lapuente, V.; Annoni, P. Measuring quality of government in EU regions across space and time. *Pap. Reg. Sci.* **2019**, *98*, 1925–1953. [[CrossRef](#)]
20. Jordaan, J.A.; Douw, W.; Qiang, C.Z. *Multinational Corporation Affiliates, Backward Linkages, and Productivity Spillovers in Developing and Emerging Economies: Evidence and Policy Making*; Policy Research Working Paper Series. No. 9364; The World Bank: Washington, DC, USA, 2020.
21. Fischer, M.M.; Scherngell, T.; Reismann, M. Knowledge spillovers and total factor productivity: Evidence using a spatial panel data model. *Geogr. Anal.* **2009**, *41*, 204–220. [[CrossRef](#)]
22. Kuznets, S. Economic growth and income inequality. *Am. Econ. Rev.* **1955**, *45*, 1–28.
23. Lee, C.C.; Wang, E.Z. Economic Complexity and Income Inequality: Does Country Risk Matter? *Soc. Indic. Res.* **2020**, 1–26. [[CrossRef](#)]
24. Le, T.H.; Nguyen, C.P.; Su, T.D.; Tran-Nam, B. The Kuznets curve for export diversification and income inequality: Evidence from a global sample. *Econ. Anal. Policy* **2020**, *65*, 21–39. [[CrossRef](#)]
25. Ferreira, F.H.G.; Leite, P.G.; Wai-Poi, M. *Trade Liberalization, Employment Flows, and Wage Inequality in Brazil*; The United Nations University World Institute for Development Economics Research (UNU-WIDER): Helsinki, Finland, 2007.
26. Azzoni, C.R.; Servo, L.M.S. Education, cost of living and regional wage inequality in Brazil. *Pap. Reg. Sci.* **2002**, *81*, 157–175. [[CrossRef](#)]
27. Cardoso, E.; Urani, A. Inflation and unemployment as determinants of inequality in Brazil: The 1980s. In *Reform, Recovery, and Growth: Latin America and the Middle East*; University of Chicago Press: Chicago, IL, USA, 1995; pp. 151–176.
28. Borraz, F.; Ferrés, D.; Rossi, M. Assessment of the distributive impact of national trade reforms in Brazil. *J. Econ. Inequal.* **2013**, *11*, 215–235. [[CrossRef](#)]
29. Ferreira, F.H.G.; Litchfield, J.A.; Leite, P.G. *The Rise and Fall of Brazilian Inequality, 1981–2004*; World Bank Policy Research Working Paper Series, No. 3867; The World Bank: Washington, DC, USA, 2006.
30. Barros, R.; De Carvalho, M.; Franco, S.; Mendonça, R. *Markets, the State, and the Dynamics of Inequality in Brazil*; United Nations Development Programme Discussion Paper: New York, NY, USA, 2010; Volume 169.
31. Signor, D.; Kim, J.; Tebaldi, E. Persistence and determinants of income inequality: The Brazilian case. *Rev. Dev. Econ.* **2019**, *23*, 1748–1767. [[CrossRef](#)]
32. Castilho, M.; Menéndez, M.; Sztulman, A. Trade liberalization, inequality, and poverty in Brazilian states. *World Dev.* **2012**, *40*, 821–835. [[CrossRef](#)]
33. Gonzaga, G.; Filho, N.M.; Terra, C. Trade liberalization and the evolution of skill earnings differentials in Brazil. *J. Int. Econ.* **2006**, *68*, 345–367. [[CrossRef](#)]
34. Pavcnik, N.; Blom, A.; Goldberg, P.; Schady, N. Trade liberalization and industry wage structure: Evidence from Brazil. *World Bank Econ. Rev.* **2004**, *18*, 319–344. [[CrossRef](#)]
35. Helpman, E.; Itskhoki, O.; Muendler, M.-A.; Redding, S.J. Trade and inequality: From theory to estimation. *Rev. Econ. Stud.* **2017**, *84*, 357–405. [[CrossRef](#)]
36. Swart, J. Gini Index: Conceiving inequality in one single number. In *No Poverty. Encyclopedia of the UN Sustainable Development Goals*; Filho, W.L., Azul, A.M., Brandli, L., Salvia, A.L., Özuyar, P.G., Wall, T., Eds.; Springer: Cham, Switzerland, 2020.
37. Henderson, J.V. Marshall's scale economies. *J. Urban Econ.* **2003**, *53*, 1–28. [[CrossRef](#)]
38. Duranton, G.; Puga, D. Micro-foundations of urban agglomeration economies. In *Handbook of Regional and Urban Economics*; Elsevier: Amsterdam, The Netherlands, 2004; Volume 4, pp. 2063–2117, ISBN 1574-0080.