Demographic aspects and regional income convergence in Brazil: a panel data approach

Rubiane Daniele Cardoso de Almeida, Benjamin M. Tabak and Tito Belchior Silva Moreira

Abstract

The framework of condition convergence forms the theoretical basis for different dynamic panel data approaches, but depending on the specifications, the method and the time period, results can vary significantly. This article presents empirical results of applying different panel data approaches to study the impact of demographic factors on regional growth of the 27 states of Brazil over 2000–2014. The results suggest that estimation using the generalized method of moments (GMM) is likely to be more consistent and efficient than the other methods studied. The results also point to a significant and negative relation between the demographic variables and regional economic growth. It is hoped that this study will contribute to the literature, by offering a comparative model framework and an analysis of demographics and economic growth.

Keywords

Economic growth, regional development, income, population aspects, population trends, macroeconomics, econometric models, Brazil

JEL classification

J11, O47, R11

Authors

Rubiane Daniele Cardoso de Almeida is a postdoctoral trainee of the Getulio Vargas Foundation School of Public Policy and Government, Brazil. Email: rubicardoso@yahoo.com.br.

Benjamin M. Tabak is a Professor and Senior Researcher at the Getulio Vargas Foundation School of Public Policy and Government in Brazil. Email: benjaminm.tabak@gmail.com.

Tito Belchior Silva Moreira is a professor and senior researcher in the Department of Economics of the Catholic University of Brasilia, Brazil. Email: titoeco@yahoo.com.br.

I. Introduction

Thoughts and discussions on income convergence between countries initially likely emerged over the course of the eighteenth and nineteenth centuries. Subsequently, in the twentieth century, Solow (1956) and Swan (1956) developed a complete theory on economic growth and convergence. Since then, different formulations have been developed, with multiple methodological and econometric applications.

This study addresses intraregional convergence, which is to say convergence of regions within a country. All 27 federative units of Brazil are analysed: 26 states and the Federal District, containing the capital Brasilia.[1] The econometric methodology used is based on dynamic panel data models, comparing four such models. In addition to these methods, indicators of inequality are used, such as the Gini coefficient and the Theil index, as well as the coefficient of variation. Econometric methods are used to find empirical evidence of β-convergence, and indicators of inequality to determine whether there is σ-convergence.

Anecdotal evidence suggests that inequality in income, wealth and opportunity are continuing to rise worldwide. Data from the *Human Development Report 2015* published by the United Nations Development Programme (UNDP, 2015), covering 2014, show that around 80% of the world's population holds just 6% of global wealth. One new development, however, is the decline between 1990 and 2014 in the number of people living in extreme poverty around the world, from 1.9 billion to 836 million. In addition, according to the same report, Brazil obtained an inequality-adjusted human development index (IHDI) value of 0.56 in 2014, below the average for Latin America and the Caribbean, which was 0.57. In this regard, in a recent study, King and Ramlogan-Dobson (2015) found evidence that growth in almost all Latin American economies is actually systematically related to that of the United States, but that some countries have converged to rather low relative income levels.

Economic activity in Brazil is highly concentrated geographically. The South-East region, the country's second smallest, holds more than 44% of the population and is responsible for more than half of national gross domestic product (GDP). However, a comparison of the years 2001 and 2012 shows a fall in the GDP share of the South-East region from 59.1% to 54.9% and rises for other regions such as the Midwest, whose share increased by more than 1 percentage point. This pattern may be indicative of a process of regional de-concentration.

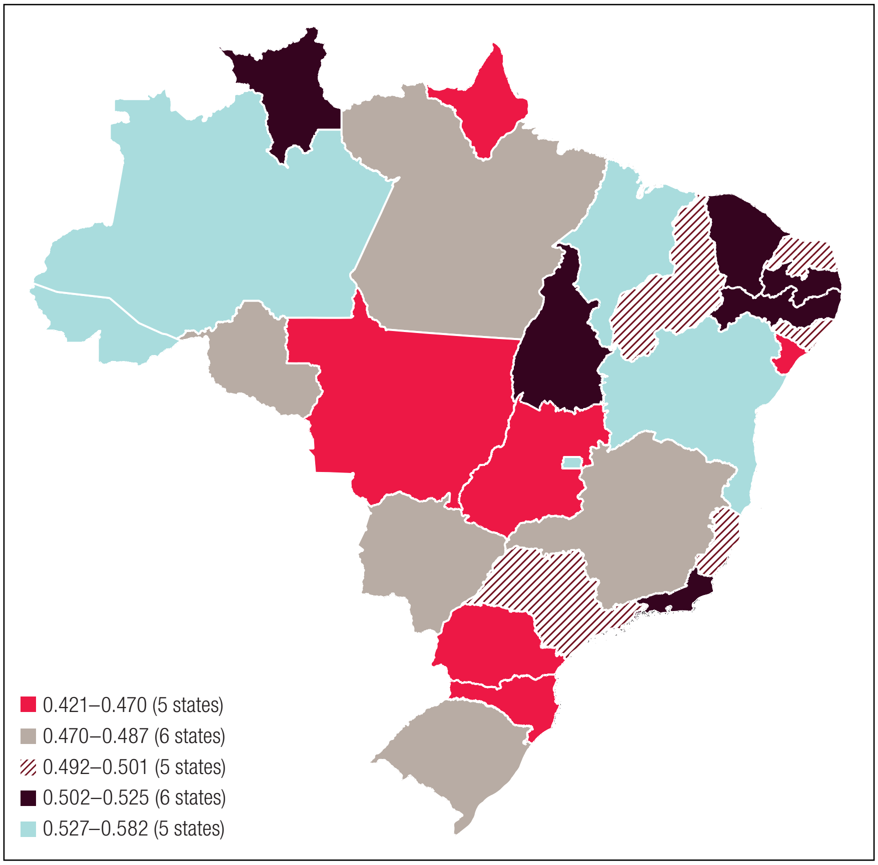
Examining labour force income in 2014, the Federal District reported the worst rates of income inequality in the country. In the Federal District, the Gini coefficient (which measures concentration of income in a given group) was 0.58, followed by the states of Acre, Amazonas, Maranhão and Bahia, with coefficients of around 0.53.[2] The state of Santa Catarina was the least unequal in Brazil, with a coefficient of 0.42. Overall, for Brazil, the coefficient was 0.60, indicating large disparities among regions. Map 1 shows the coefficients for each state in 2014.

The existence of a process of income convergence between regions is very important in the study of economic growth and inequality. According to Abramovitz (1986), technologically backward countries have the potential to grow faster than developed countries, provided they have enough capacity to take advantage of the technological level of leading countries. The question is whether poor countries have higher growth rates than rich countries.

[1] For the purpose of simplicity, all 27 federative units of Brazil will be referred to in this article collectively as states.

[2] The Gini coefficient varies from 0 to 1. The closer to 0, the lower the income inequality, and the closer to 1, the higher the income inequality. Data extracted from IPEA (2022).

Map 1
States of Brazil: Gini coefficients, 2014



Source: Prepared by the authors, on the basis of Institute of Applied Economic Research (IPEA), Ipeadata [online database] http://www.ipeadata.gov.br/Default.aspx.

In this context, the aim of this study is to analyse the process of convergence between Brazilian states from 2000 to 2014, considering the dynamic panel data approach. The article contributes to discussions in the literature by including demographic variables, seeking to analyse their impact on economic growth and convergence. Another contribution of this article is empirical discovery based on different panel data approaches and different estimation methods, focusing mainly on the generalized method of moments (GMM). The importance of this topic lies in the key information that can be provided to policymakers.

Our observations complement existing results on Brazil. Among the most recent studies, Cravo, Becker and Gourlay (2015) and Resende and others (2016) previously found conditional β-convergence but at relatively slow rates. The main methodological difference with respect to those studies is that we primarily control for the demographic characteristics of each state, through the fixed-effects model, while convergence in the two other papers is conditional on specific factors such as education, population density or infrastructure. The increase in convergence rates in fixed-effects (FE) estimates is in line with Islam (1995) who attributes the difference to omitted variable bias in regressions without FE.

Resende and others (2016) do not focus solely on the state-level but also compare their results to those obtained from other less aggregated spatial units. The differences, however, are minor, suggesting that the findings hold in a general context, irrespective of the level of analysis. The multilevel analysis in Dapena and others (2017) reveals internal divergence in the most developed states in the South-East region and in the states of the North-East. At the national level, their estimation confirms the previous positive relative convergence results. Thus, our work complements the existing literature for Brazil, specifically applying a demographic perspective to convergence and growth through GMM.

In short, the results show relevant changes in demographic characteristics among the Brazilian regions. Migration processes are less intense and are occurring in reverse, with people returning to their region of origin and fertility appears to be declining in all states. The empirical evidence shows a negative influence of population density, migration and fertility rates on per capita income; although these variables are in transition, they do appear to demonstrate the persistence of some socioeconomic and cultural characteristics of Brazil.

In addition to this introduction, the paper has four sections: section II outlines the basic concepts of convergence and empirical evidence on the subject; section III describes the data and empirical strategy used; section IV provides the results of the σ-convergence and β-convergence tests; lastly, section V summarizes the study and offers some final considerations.

II. Methodological aspects

1. Theory and estimation

In the recent debate about economic growth and income convergence, two theoretical approaches to empirical research stand out. The first relates to the model put forward by Solow (1956), which suggests that poorer regions tend to grow faster than richer regions owing to decreasing returns to scale on the capital stock. The second approach draws from the discussion of conditional convergence, which seeks to identify the factors that drive economic growth, according to Barro and Sala-i-Martin (1991). In order to test the convergence condition, these authors added, to the basic model of Solow (1956), a set of variables that refer to differences in the steady-state of the different economies. This approach is based on the supposition that there will only be convergence between countries or regions if they are similar to each other.

One critique of the model proposed by Barro and Sala-i-Martin (1991) is that it omits unobservable effects or considers them insignificant. Islam (1995) proposes a method that considers the different production functions of economies, which may overcome problems of this nature, comprising a dynamic panel data model that includes in its equation unobservable country effects. Islam (1995) compares his results with those of Mankiw, Romer and Weil (1992), emphasizing the changes that occurred when including specific effects.

For this study, we opt for a dynamic panel data model based on Islam (1995), in which the lagged dependent variable captures the short-run autoregressive behaviour of income. As shown by Islam (1995), the model can be written as:

$$Lny_{i,t} = \gamma Lny_{i,t-1} + \sum_{j=1}^{n} \beta_j X_{it} + \eta_t + \mu_i + v_{it} \quad (1)$$

where $\gamma = e^{-\lambda t}$, $\lambda = -\frac{ln\gamma}{\tau}$ being the speed of convergence and τ the variation in time, and the parameters μ_i and η_t are specific to each state and each year, respectively. The vector X_{it} represents the structural characteristics of each region. The convergence velocity can best be interpreted through the definition of half-life, i.e., half the time it takes an economy to reach half the distance to its steady-state.

Islam (1995) used the least squares dummy variable (LSDV) method in the dynamic version. As Roodman (2006) points out, the panel data approach is better than cross-sectional regression as it enables controlling for endogeneity and omitted variable bias. However, the LSDV method may not eliminate the existing bias in the case of dynamic panel estimation, since the lagged dependent variable is negatively correlated with the error term, underestimating the coefficient value.

In fact, Caselli, Esquivel and Lefort (1996) raised the endogeneity problem of explanatory variables not considered by Islam (1995). As a solution to the problem identified in Islam (1995), the authors propose the use of a GMM estimation developed by Arellano and Bond (1991). In this paper, in addition to the classical methods —ordinary least squares (OLS) and LSDV— we estimate the following models: two-stage least-squares or instrumental variables (IV) by Anderson and Hsiao (1981), difference GMM by Arellano and Bond (1991) and system GMM by Blundell and Bond (1998), which follows the analysis structure proposed by Roodman (2006). Thus, our intention is to provide a comparative table of estimates, to find the most consistent model for the sample, focusing primarily on demographic variables.

2. Data

A dataset was constructed for the 27 Brazilian states[3] and covering the period from 2000 to 2014. When selecting variables, the aim was to include ones traditionally used in the analysis of growth and convergence models, with some added demographic variables —ratio of urban to total population, population density, migration and fertility (see table 1). The main sources of data are the Brazilian Institute of Geography and Statistics (IBGE) and the Ipeadata database of the Institute of Applied Economic Research (IPEA). The economic variable considered is per capita GDP. This is the classical variable used in growth models. An additional proxy variable is included for human capital: the proportion of formal workers with a high school education or higher. The purpose of including this variable is to analyse how qualification of the workforce impacts on growth. The variable representing employment is the percentage of the total population that is in work. Capital expenditure as a proportion of each state's GDP is used as a proxy for public investment.

Table 1
Summary statistics

Description		Mean	Std. dev.	Min.	Max.	Source
Dependent variable						
GDP	Per capita GDP *(Reais)*	16 016	10 337	4 477	64 013	Institute of Applied Economic Research (IPEA), Ipeadata Brazilian Institute of Geography and Statistics (IBGE)
Independent variables						
Human capital	Proxy for human capital: share of formal workers with high school or higher education	0.243	0.108	0.070	0.654	IBGE
Investment	Capital expenditures to GDP *(Reais)*	31 269	22 686	7 993	130 681	Ipeadata
Employment	Ratio of employed persons to total population	0.410	0.068	0.254	0.526	Ipeadata
Urbanization	Ratio of urban population to total population	0.808	0.097	0.587	1.000	IBGE
Population density	Demographic density: inhabitants per square kilometre	66 634	101 080	1 546	483 258	IBGE
Migration	Migration rate: ratio of residents not belonging to the state to total population	0.194	0.136	0.038	0.537	IBGE
Fertility	Fertility rate: live births per woman aged 15–49 years	2.219	0.471	1.583	3.696	IBGE

Source: Prepared by the authors.
Notes: The table shows mean, standard deviation, minimum and maximum value for the main variables in 2000–2014. The number of observations is equal to 27 in each year.

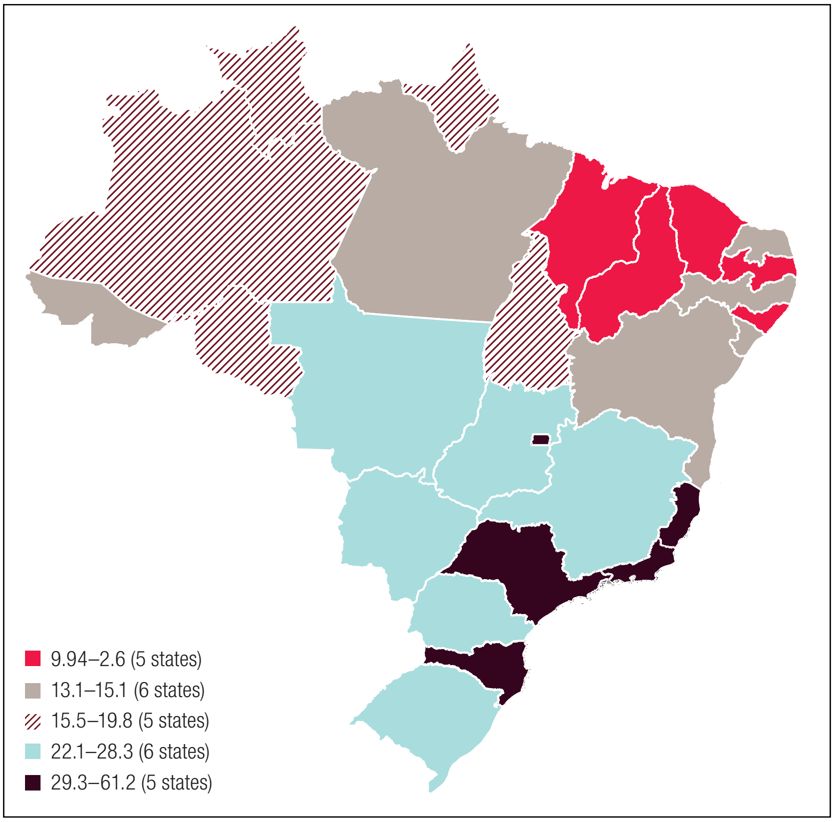
[3] Brazilian states: Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE) and Tocantins (TO).

The demographic variables are: the portion of the population that lives in cities; population density expressed as inhabitants per square kilometre; migration, expressed as the number of inhabitants not born in a state as a proportion of its total population; and the fertility rate, which is the number of children born to women of childbearing age (15-49 years). These variables were selected to verify the impact of the demographic characteristics of each region on growth and economic convergence.

In order to account for possible business cycle fluctuations, and in order not to capture short-term growth, triennial averages of the variables are used. As the series covers 15 years, this results in 5 time points, as in Islam (1995). The descriptive statistics are presented in table 1. The natural logarithm is used for all variables, in order to follow the classical methodology for convergence models. According to Wooldridge (2010), in some cases the logarithmic transformation can alleviate problems of heteroscedasticity. In addition, it can narrow the amplitude of values of the variables, making estimates less sensitive to outliers.

Map 2 shows the spatial distribution of the per capita GDP variable for 2014.

Map 2
States of Brazil: spatial distribution of per capita GDP, 2014
(Thousands of reais)



Source: Prepared by the authors.

In general, the states of the North and North-East regions have the lowest per capita GDPs, while the states of the South and South-East have the highest. Thus, it appears that the states with the highest product or income are surrounded by similar wealthy states and vice versa. This may be a result of socioeconomic interaction between neighbouring regions —technological overflows, labour mobility, economies of scale— or may also result from macroeconomic policies and sociocultural characteristics.

III. Results

An oft-repeated fact in the literature is that analysis of convergence uses two main measures —σ- and β-convergence— which are different, but also complementary. These two approaches gained prominence mainly through the seminal work of Barro and Sala-i-Martin (1992). Because neither kind of convergence necessarily entails the other, and because each has its own merits, we agree with Young, Higgins and Levy (2008) that it is worth studying and comparing the two forms.

1. σ-convergence

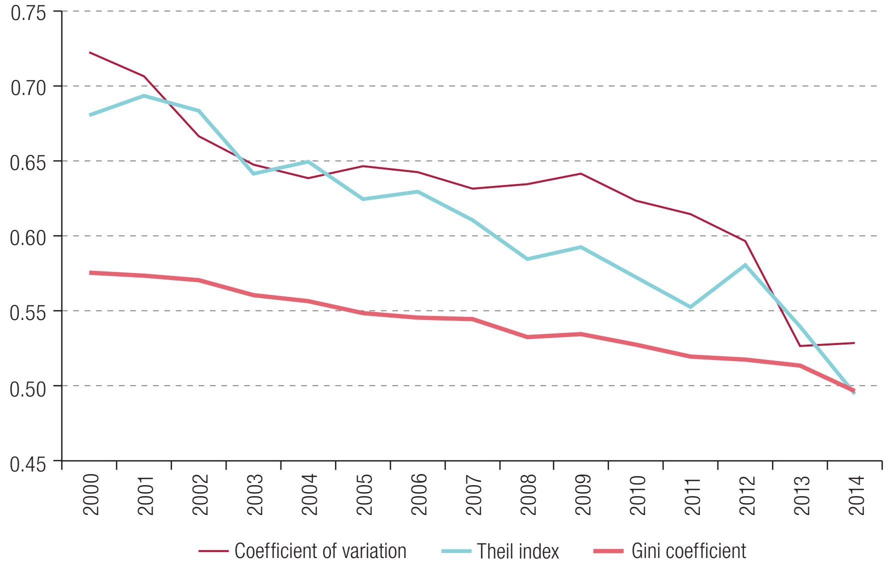
Even in the earliest studies of regional inequalities, such as those of Kuznets (1955) and Williamson (1965), convergence between territories was already attracting the attention of researchers in the area and has continued to do so. In these initial studies, an index similar to the standard deviation was used as a measure of income dispersion, known today as σ-convergence. This then became the simplest concept of convergence and, according to Barro and Sala-i-Martin (1991), can be understood as the continuous dynamics of reducing the differences between regional incomes, which leads to less dispersion and less inequality between economies. The literal concept of σ-convergence was introduced by Barro (1991).

The standard deviation, however, is not an adequate measure to compare variables that are measured at different scales or even when the scale of a single variable changes significantly over time. The coefficient of variation, which is defined as the standard deviation relative to the mean of a variable, is independent of the absolute magnitude of the variable and is therefore a more appropriate measure to test convergence.

Another means of obtaining evidence of this process would be to observe the behaviour of indicators of income inequality. These indicators have the capacity to demonstrate whether the differences between incomes in different states have become larger or smaller over a period. Figure 1 shows the main indicators of income inequality: the Gini coefficient, the Theil index and coefficient of variation. The Gini coefficient and Theil index are similar and vary between zero and one. The value zero corresponds to a complete equality between the incomes, whereas the value one corresponds to a complete inequality between the incomes.

As shown in Figure 3, the overall trend over time is a downward one. Although from 2004 to 2009, the coefficient of variation seemed to show some stability, it began falling again in 2009. The Theil index is the most unstable, with several peaks over the period but a general trend of declines, since at the start of the period the indicator stood at around 0.68 but had reached 0.50 by its end. The Gini coefficient also shows a decrease over the period, albeit a less pronounced one, but with a more linear trend than the other indicators. Overall, these indicators provide some evidence of σ-convergence. In a recent study for the provinces of China, Tian and others (2016) also use the coefficient of variation for σ-convergence analysis and find a reduction in the dispersion of income over time.

Figure 1
Indicators of income inequality, 2000–2014



Source: Prepared by the authors.
Note: The number of observations is 27 in each year.

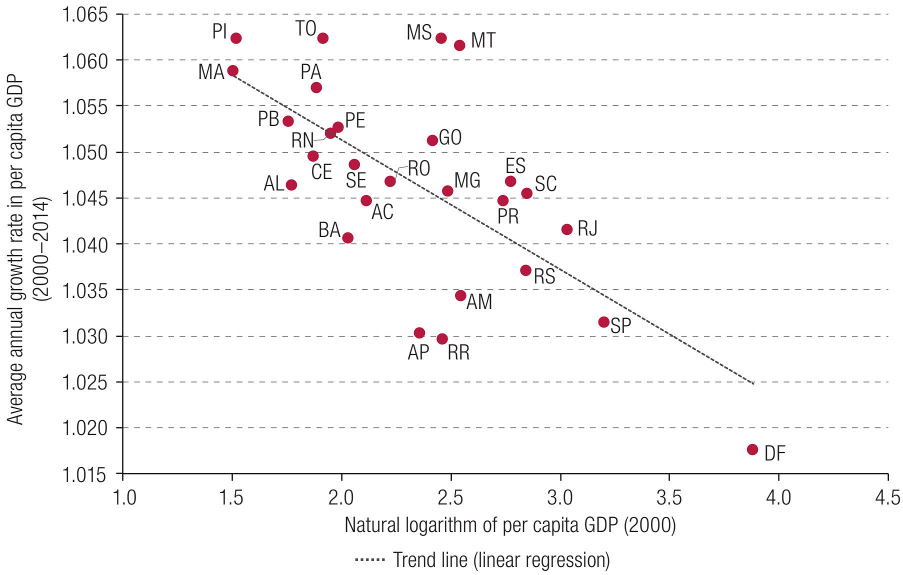
2. β-convergence

Before we analyse the empirical results obtained for convergence, figure 2 facilitates analysis of the convergence process according to Barro and Sala-i-Martin (1991), showing the relationship between the natural logarithm of per capita GDP in 2000 and its average annual growth rate for 2000–2014. To illustrate this relationship more clearly, a trend line is drawn, and a trend line has been drawn to show the linear regression. The decreasing relationship shows that those states with higher initial values experience lower growth rates in this variable. This behaviour is typical of a convergence process.

The states with the greatest economic lag at the beginning of the period and with the highest growth rates are: Piauí, Maranhão and Pernambuco, which are part of the North-East region. The characteristics that may have influenced the convergence process in that region include: (i) income redistribution policies, mainly through the Bolsa Família direct aid programme for the poorest families since 2003; (ii) expansion of the agricultural frontier in the Cerrado ecoregion of tropical savanna in eastern Brazil, particularly in terms of soybean, irrigated fruit growing and export diversification; (iii) expansion of credit with an emphasis on residential credit through the Minha Casa, Minha Vida programme (in 2006, loans were equivalent to 26% of the region's GDP, and in 2010 to 49%); and (iv) economic stability through implementation of the Plano Real which sought low inflation to increase the purchasing power of the low-income population and contribute to reducing social inequality.

We implement tests for the β-convergence conditional including characteristics of each state, according to tables 2 and 3. Adopting a comparative approach and based on the literature, we use the following models: OLS, FE, IV (Anderson and Hsiao, 1981), difference GMM and system GMM. The dependent variable is the natural logarithm of per capita GDP. Table 2 shows the results for the model without the demographic variables. Table 3 shows the results including the demographic variables.

Figure 2
States of Brazil: average per capita GDP growth rate, 2000–2014 and natural logarithm of per capita GDP in 2000 (simple relation)[a]
(Percentages and natural logarithms)



Source: Prepared by the authors.
Note: The graph includes the results of a simple linear regression between the two variables under consideration.
[a] Brazilian states: Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE) and Tocantins (TO).

Table 2
β-convergence analysis

	Dependent variable: $Ln(GDP)_{i,t}$				
	OLS	FE	IV	System GMM	Difference GMM
$Ln(GDP)_{i,t-1}$	0.910***	0.582***	0.852**	0.842***	0.615***
	(0.029)	(0.078)	(0.410)	(0.047)	(0.104)
Human capital	0.071*	0.282***	0.284***	0.139**	0.252***
	(0.038)	(0.054)	(0.091)	(0.051)	(0.075)
Investment	0.009	0.046	0.037	-0.003	0.027
	(0.011)	(0.028)	(0.034)	(0.016)	(0.027)
Employment	0.162***	0.204	0.299	0.229***	0.431
	(0.045)	(0.155)	(0.262)	(0.063)	(0.183)
Constant	0.583***	1.629***	-0.103	0.952	-
	(0.162)	(0.325)	(0.086)	(0.262)	
R^2	0.988	0.939	-	-	-
Implicit λ	0.031	0.180	0.053	0.057	0.162
Half-Life	22.04	3.84	12.98	12.09	4.22
Observations	108	108	108	108	108
Instruments	-	-	-	13	9
Hansen				10.78	10.70
				[0.21]	[0.05]
AB(2)				0.23	0.95
				[0.81]	[0.34]

Source: Prepared by the authors.
Note: Standard error in parentheses. P-value between square brackets. $VIF_{max} = 5.05$; Heteroscedasticity test: $\chi^2(1) = 0.15$, $p < 0.696$. The Wald test for the FE model: $\chi^2(27) = 777$, $p < 0.001$. GMM estimates (the last two columns) were performed in first-step robust, limited in 4 lags to prevent the proliferation of instruments. * significant at 10%, ** significant at 5% and *** significant at 1%.

Table 3

β-convergence analysis (including demographic variables)

	Dependent variable: Ln(GDP)$_{i,t}$				
	OLS	FE	IV	SYS	DIFF
Ln(GDP)$_{i,t-1}$	0.909***	0.433***	0.666**	0.779***	0.610***
	(0.030)	(0.085)	(0.310)	(0.088)	(0.141)
Human capital	0.047	0.147***	0.286***	0.152**	0.179***
	(0.034)	(0.061)	(0.095)	(0.077)	(0.056)
Investment	0.007	0.043*	0.037	-0.017	0.026
	(0.014)	(0.025)	(0.032)	(0.030)	(0.028)
Employment	0.090	0.280	0.413	0.216**	0.641***
	(0.067)	(0.182)	(0.273)	(0.103)	(0.140)
Density	-0.007	-0.551*	-0.825	-0.008	-0.878**
	(0.005)	(0.294)	(0.530)	(0.006)	(0.332)
Urbanization	-0.039	0.070	-0.078	0.105	-0.134
	(0.081)	(0.157)	(0.286)	(0.114)	(0.201)
Migration	0.014	-0.116	-0.081	0.029*	-0.266**
	(0.010)	(0.091)	(0.126)	(0.016)	(0.102)
Fertility	-0.144**	-0.990***	-0.110	-0.106	-0.674**
	(0.072)	(0.224)	(0.458)	(0.090)	(0.302)
Constant	0.639***	4.212***	-0.103	1.350**	-
	(0.177)	(1.158)	(0.086)	(0.515)	
R^2	0.989	0.953	-	-	-
Implicit λ	0.031	0.279	0.053	0.083	0.164
Half-Life	21.79	2.48	12.98	8.32	4.20
Observations	108	108	108	108	108
Instruments				17	13
Hansen test				12.03	9.01
				[0.21]	[0.10]
AB(2)				0.16	0.67
				[0.87]	[0.50]

Source: Prepared by the authors.

Note: Standard error in parentheses. P-value between square brackets. VIFmax = 9.34; Heteroskedasticity test: χ^2 (1) = 0.71, $p < 0.399$. The Wald test for the FE model: χ^2 (27) = 632, $p < 0.001$. The GMM estimates (last two columns) were performed in first-step robust, limited in 4 lags to prevent the proliferation of instruments. * significant at 10%, ** significant at 5% and *** significant at 1% level.

The estimates provide positive results for the presence of β-convergence, meaning that they will only converge to the same steady state if they are similar to each other. The coefficients obtained for the variable initial income are greater than 0 and less than 1 and statistically significant to less than 1%, except in the IV model. However, Islam (1995) emphasized in his study that the use of OLS to measure β-convergence may not be the most appropriate methodology. This estimator would be consistent if the individual effects of each state were captured by the error term and it was not correlated with the explanatory variables. As discussed in Islam (1995), one way of addressing individual effects would be through the FE estimator. However, in the case of a dynamic panel this method can generate unreliable estimators, since the lagged dependent variable is correlated with the mean of the errors.

Given the difficulty of finding good instruments for the variables that present problems, Roodman (2006) notes that researchers are often faced with the need to design instruments from the data set itself. This is the case for this research.

According to Ding, Haynes and Liu (2008), as it does not consider unobserved time and regional effects, the OLS estimate tends to tilt up the coefficient of the lagged dependent variable, given its correlation with the error. Roodman (2006) shows that although FE estimation is better than OLS, it does not eliminate the bias of the dynamic panel, since a negative correlation remains between the lagged dependent variable and the error term, tilting the value of the coefficient down. Therefore, the true value of the coefficient of the lagged dependent variable must be between the values found by the OLS and LSDV estimates, which serve as boundaries or intervals. Thus, an estimated coefficient for the lagged dependent variable that lies within this range offers some reliability.

As a starting point, the simplest way to incorporate any instrument into a regression is by using the two-stage least squares (2SLS) method, which refers to the Anderson and Hsiao (1981) instrumental (IV) estimators. This model was estimated in first difference using an instrument for the lagged dependent variable, such as $\Delta y_{i,t-2} = y_{i,t-2} - y_{i,t-3}$.

Our analysis focuses on the GMM estimators, because, according to the literature, they may present greater consistency. The results shown in the last two columns of table 2 indicate income convergence. Roodman (2006) recalls that the Sargan statistic is a special case of the Hansen statistic under the assumption of homoscedasticity. Thus, for robust GMM estimation, the Sargan test is inconsistent. Therefore, taking into account the Hansen statistic presented in table 2, the instruments used in the GMM models are valid. Considering the GMM difference and GMM system models, in table 2 the estimated convergence velocity is 5%–16%, indicating that the time required for the states to reach half the distance to their steady state is 4–12 years. As expected, the proxies for human capital and employment have a positive impact on per capita income.

Adding the demographic variables, in table 3, we observed that the speed of convergence increases, being between 8 and 16% for GMM models. The half-life passes to the interval between 4 and 8 years. It is important to highlight that these results are higher than the ones found by Barro and Sala-i-Martin (1990 and 1995), which were around 2% for European regions and North American states. This fact may contribute to the theory that poorer regions grow faster than richer regions. In recent studies, Mishra and Mishra (2018) and Tian and others (2016) also find evidence for convergence in regions of India and China, respectively.

In the case of the GMM difference model, the variables that represent human capital and labour continue to have a positive impact on income. With this model, there is no significant change in the speed of convergence when the variables of interest are added. Demographic variables —population density, migration and fertility— seem to have an impact on income, as expected. Population density has a negative impact, meaning that the more people per square kilometre, the lower the per capita income, suggesting that it is preferable for populations not to agglomerate in large urban areas. But this variable can also refer to population growth, which is naturally negatively correlated with income if income growth does not occur with the same intensity.

The fertility rate variable has a negative impact on income. The fertility rate is declining in Brazil, owing either to women's higher labour force participation or to the notion that successful families have few children. In 2000, the country's total fertility rate (TFR) was 2.36, but it had fallen to 1.79 in 2014, below the population replacement rate. There are two scenarios that complement each other with regard to this negative correlation between income and fertility. In the first scenario, the states with the lowest income are those with the highest fertility rates, even though fertility is declining for all states in the period analysed. In the second scenario, historically, the higher the income of a region, the lower the fertility rate tends to be; indeed, based on average values for the period, the Federal District is the state with the highest per capita income and also the lowest fertility rate (1.75).

According to IBGE data (2000 and 2010 censuses), education and income level are crucial factors for TFR. Among women with no schooling or up to three years of schooling, TFR fell from 3.8 to 3.0 between 2000 and 2010, and for those with four to eight years of schooling, from 2.8 to 2.6. In contrast, among women with a high school education, (9–11 years of schooling), TFR estimates show an increase from 1.6 to 1.8 over the same period and for those with 12 or more years of schooling, from 1.1 to 1.2. In terms of wage levels, there was a fall in TFR in all income strata, but among the lower-income strata the fall was sharper. In this regard, TFR was 1.3 among women with income of one to two minimum wages and 1.1 among those with incomes of more than two minimum wages in 2010. These TFR are called lowest-fertility levels and are comparable to European countries with very low fertility such as Italy and Portugal (see Kohler, Billari and Ortega, 2002; Morgan, 2003, Breton and Prioux, 2009).

Notably, the sample of women of childbearing age (15–49 years) increased between the two census years (2000 and 2010) from 46 million to 53 million. Schooling levels among women also underwent significant changes, as the number with 9–11 years of schooling rose from 13 million to 20 million and those with more than 12 years of schooling climbed from 4 million to 10 million. Conversely, the number of women with up to eight years of schooling fell.

Migration is an important means for people to improve their economic well-being and quality of life. Therefore, net population movements tend to be towards prosperous areas that offer better prospects in terms of real income. Although migration is considered to be a mechanism for reducing spatial income differentials (McCann, 2001), the impacts may depend on the characteristics of the migrants, such as entrepreneurship capacity and skills that can contribute to the economic growth of their destination (Poot, 2008). In the GMM difference model, this variable has a negative impact on income, perhaps meaning that characteristics of migrants are a critical factor. Indeed, in Brazil, many low-income families, mainly from the North and North-East regions, migrate in search of a better life.

However, according to data from IBGE, migration between regions of the country slowed between 2000 and 2010, and states from the North-East region, in addition to retaining population, began to record emigrants returning from the centre or south of the country. In 2009, the North-East states that reported the most significant return migration —more than 20% of total immigrants— were Pernambuco, Sergipe, Rio Grande do Norte and Paraíba. While São Paulo and Rio de Janeiro have begun to receive fewer immigrants in the last decade, states previously considered to have high outflows of migrants have begun to lose less population, such as Piauí and Alagoas. Bahia and Maranhão continued to post negative net migration, but also decreased flows.[4]

Since the 1990s, the literature on economic growth has included several studies of the role of internal migration in the convergence of per capita income. However, the current literature on the effects of migration is still inconclusive. Observed results may depend on several characteristics of the study, research methodologies, type of data and the spatial scale of measurement in which the research was performed.

Lastly, there are other important considerations. As this research consists of a test for the hypothesis of income convergence and of choosing the most appropriate estimation method, the results should not be seen as definitive, but rather as a stimulus for new studies on the subject. Although the GMM specification indicates a more robust analysis, Roodman (2006) emphasizes that this methodology should be used in the case of panels with a small T (period) and large N (individuals), in order to obtain better results. Nonetheless, a "large" number of individuals is not precisely defined. In any case, other studies use the same method with a small number of individuals, such as Ding, Haynes and Liu (2008) and Cabral and Varella Mollick (2012). This corroborates the relevance of the method and the possibility of obtaining more consistent results on economic growth and income convergence.

[4] Data from the National Household Survey (PNAD) 2009 and 2000 and 2010 censuses (IBGE, 2022a and 2022b).

IV. Conclusions

This article analyses the convergence process among Brazilian states, considering the impact of demographic variables on income. We examine the process of σ-convergence through inequality indicators, as well as the process of β-convergence using econometric models, establishing a comparative framework.

In general, the results show favourable economic convergence in the 2000–2014 period. In the analysis of σ-convergence, the indicators show a decreasing trajectory in their dispersion, indicating that the Brazilian states have become, on average, more equal in economic terms. Our analysis also finds β-convergence. The variables that refer to education and employment showed positive correlation with income, as expected.

Graphical analysis reveals that the states with the highest convergence rates are in the North-East region. The key factors behind the improvement in social conditions are the large income transfers from the government through social programmes, economic stability owing to the implementation of the Real Plan (Plano Real), the expansion of the agricultural sector and an increase in the supply of credit.

In addition, the results confirm the negative relation between economic growth and fertility. Given regional diversity and inequality, this may indicate that higher fertility rates are still occurring among the poorest populations in the states. Policymakers must therefore ensure they carefully incorporate the demographic transition and its social and economic effects into policy design.

As regards migration, the demographic scenario is changing. Migration processes are less intense and are occurring in reverse, with people returning to their region of origin. Even so, migration's impact on income is still negative, perhaps suggesting certain characteristics in this population group, particularly in terms of schooling, which may be more limited.

Our reading of the data is that, most probably there are several stationary states in the Brazilian regions, owing to their considerable economic and cultural diversity. Therefore, confirmation that there is a convergence process does not necessarily mean there will be a reduction in disparities. Without other regional development initiatives, regional economic inequalities are likely to persist.

The main contributions of this research are the provision of empirical evidence of income convergence over a recent period, the comparison of various methods of estimation and the focus on specific demographic factors. Among these factors, the migration rate and fertility rate are particularly important, since the Brazilian economy has a history of high levels of both, but more recently with significant changes. In addition, empirical results suggest that the high fertility rates of the poorest women affect GDP. Lastly, we propose that further studies on the subject look at other geographical levels.

Bibliography

Abramovitz, M. (1986), "Catching up, forging ahead, and falling behind", *The Journal of Economic History*, vol. 46, No. 2, June.
Anderson, T. W. and C. Hsiao (1981), "Estimation of dynamic models with error components", *Journal of the American Statistical Association*, vol. 76, No. 375, September.
Arellano, M. and S. Bond (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *The Review of Economic Studies*, vol. 58, No. 2, April.
Barro, R. J. (1991), "Economic growth in a cross section of countries", *The Quarterly Journal of Economics*, vol. 106, No. 2, May.
Barro, R. J. and X. Sala-i-Martin (1995), *Economic growth*, New York, McGraw-Hill.
___(1992), "Convergence", *Journal of Political Economy*, vol. 100, No. 2, April.
___(1991), "Convergence across states and regions", *Brookings Papers on Economic Activity*, vol. 1.

___(1990), "Economic growth and convergence across the United States", *Working Paper*, No. 3419, Cambridge, National Bureau of Economic Research (NBER).
Blundell, R. and S. Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, vol. 87, No. 1, November.
Breton, D. and F. Prioux (2009), "The one-child family: France in the European context", *Demographic Research*, vol. 20, No. 27, June.
Cabral, R. and A. Varella Mollick (2012), "Mexico's regional output convergence after NAFTA: a dynamic panel data analysis", *The Annals of Regional Science*, vol. 48, No. 3, June.
Caselli, F., G. Esquivel and F. Lefort (1996), "Reopening the convergence debate: a new look at cross-country growth empirics", *Journal of Economic Growth*, vol. 1, No. 3, September.
Cravo, T. A., B. Becker and A. Gourlay (2015), "Regional growth and SMEs in Brazil: a spatial panel approach", *Regional Studies*, vol. 49, No. 12.
Dapena, A. D. and others (2017), "Convergence in Brazil: new evidence using a multilevel approach", *Applied Economics*, vol. 49, No. 50.
Ding, L., K. E. Haynes and Y. Liu (2008), "Telecommunications infrastructure and regional income convergence in China: panel data approaches", *The Annals of Regional Science*, vol. 42, No. 1, January.
IBGE (Brazilian Institute of Geography and Statistics) (2022a), "PNAD - Pesquisa Nacional por Amostra de Domicílios" [online] https://www.ibge.gov.br/estatisticas/sociais/populacao/9127-pesquisa-nacional-por-amostra-de-domicilios.html?edicao=18338&t=microdados.
___(2022b), "Censo Demográfico" [online] https://www.ibge.gov.br/estatisticas/sociais/populacao/22827-censo-2020-censo4.html?=&t=destaques.
IPEA (Institute of Applied Economic Research) (2022), Ipeadata [online database] http://www.ipeadata.gov.br/Default.aspx.
Islam, N. (1995), "Growth empirics: a panel data approach", *The Quarterly Journal of Economics*, vol. 110, No. 4, November.
King, A. and C. Ramlogan-Dobson (2015), "International income convergence: Is Latin America actually different?", *Economic Modelling* vol. 49, September.
Kohler, H. P., F. C. Billari and J. A. Ortega (2002), "The emergence of lowest-low fertility in Europe during the 1990s", *Population and Development Review*, vol. 28, No. 4, December.
Kuznets, S. (1955), "Economic growth and income inequality", *The American Economic Review*, vol. 45, No. 1, March.
Mankiw, N. G., D. Romer and D. N. Weil (1992), "A contribution to the empirics of economic growth", *The Quarterly Journal of Economics*, vol. 107, No. 2, May.
McCann, P. (2001), *Urban and Regional Economics*, Oxford, Oxford University Press.
Mishra, A. and V. Mishra (2018), "Is there conditional convergence in the per capita incomes of BIMAROU states in India?", *Economic Modelling*, vol. 70, April.
Morgan, S. P. (2003), "Is low fertility a twenty-first-century demographic crisis?", *Demography*, vol. 40, No. 4, November.
Poot, J. (2008), "Demographic change and regional competitiveness: the effects of immigration and ageing", *International Journal of Foresight and Innovation Policy*, vol. 4, No. 1/2.
Resende, G. M. and others (2016), "Evaluating multiple spatial dimensions of economic growth in Brazil using spatial panel data models", *The Annals of Regional Science*, vol. 56, No. 1, January.
Roodman, D. (2006), "How to do xtabond2: an introduction to 'difference' and 'system' GMM in Sata", *Working Paper*, No. 103, Center for Global Development, December.
Solow, R. M. (1956), "A contribution to the theory of economic growth", *The Quarterly Journal of Economics*, vol. 70, No. 1, February.
Swan, T. W. (1956), "Economic growth and capital accumulation", *Economic Record*, vol. 32, No. 2, November.
Tian, X. and others (2016), "Regional income inequality in China revisited: a perspective from club convergence", *Economic Modelling*, vol. 56, August.
UNDP (United Nations Development Programme) (2015), *Human Development Report 2015: Work for Human Development*, New York.
Williamson, J. G. (1965), "Regional inequality and the process of national development: a description of the patterns", *Economic Development and Cultural Change*, vol. 13, No. 4, part 2, July.
Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*, 2nd edition, Cambridge, MIT Press.
Young, A. T., M. J. Higgins and D. Levy (2008), "Sigma convergence versus beta convergence: Evidence from U.S. county-level data", *Journal of Money, Credit and Banking*, vol. 40, No. 5, August.