The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1517-7580.htm

Poverty, population growth and agglomeration effects in all Brazil cities

André M. Marques
Department of Economics, UFPB, Joao Pessoa, Brazil

Received 13 November 2022
Revised 4 April 2023
Accepted 24 July 2023

Abstract

Purpose – This paper aims to test three hypotheses in city growth literature documenting the poverty reduction observed in Brazil and exploring a rich spatial dataset for 5,564 Brazilian cities observed between 1991 and 2010. The large sample and the author's improved econometric methods allows one to better understand and measure how important income growth is for poverty reduction, the patterns of agglomeration and population growth in all Brazilian cities.

Design/methodology/approach – The author identifies literature gaps and use a sizeable spatial dataset for 5,564 Brazilian cities observed in 1991, 2000 and 2010 applying instrumental variables methods. The bias-corrected accelerated bootstrap percentile interval supports the author's point estimates.

Findings – This manuscript finds that Brazilian data for cities does not support Gibrat's law, raising the scope for urban planning and associated policies. Second, economic growth on a sustainable basis is still a vital source of poverty reduction (The author estimates the poverty elasticity at four percentage points). Lastly, agglomeration effects positively affect the city's productivity, while negative externalities underlie the city's development patterns.

Originality/value – Data for cities in Brazil possess unique characteristics such as spatial autocorrelation and endogeneity. Applying proper methods to find more reliable answers to the above three questions is a desirable procedure that must be encouraged. As the author points out in the manuscript, dealing with endogenous regressors in regional economics is still a developing matter that regional scientists could more generally apply to many regional issues.

Keywords Growth, Cities, Poverty, Productivity, Agglomeration, Endogeneity

Paper type Research paper

1. Introduction

The world economy has made an impressive achievement in the past few decades. In 2015, only one-tenth of the world's population lived in extreme poverty, while more than a third of people lived in extreme poverty in 1990. Following world trends, Brazil has made immense social progress by reducing the poverty rate from 56.70% in 1991 to 23.20% in 2010, faster than countries with similar per capita income levels [1]. Compared to Brazil's past, we observe massive changes. The poverty reduction in Brazil between 1985 and 2004 was perfunctory. In those two decades, the poverty rate fell a mere four percentage points from 33% to 29% of the population (Ferreira, Leite, & Ravallion, 2007). The low growth rate in per capita gross domestic product (GDP) and the low elasticity of poverty can explain the low performance in poverty reduction (Ferreira *et al.*, 2007). Thus, even if the economy operates in a regime of low per capita income growth, it can alleviate poverty considerably if the elasticity is sufficiently high.

JEL Classification — O40, R11, R23

All remaining errors in this paper are the author's responsibility. The author thanks for helpful and constructive comments by an anonymous journal referee that substantially improved the paper.

Disclosure statement: No conflict of interest was reported by the author.

Nevertheless, social progress on a sustainable basis crucially depends on the proper policy measures and institutions (the size and the quality of government, for example) for recovery productivity growth in Brazil next years or decades. The slowdown in productivity growth observed in Brazil and other developing countries have occurred many years before the COVID-19 pandemic (World Bank, 2020, World Bank, 2020; Bacha & Bonelli, 2016).

Brazilian data show that labor productivity has stagnated in the last four decades of economic policy, preventing the Brazilian per capita income from getting closer to developed countries. In a recent contribution, Costa and Marcolino (2021) present empirical findings showing that output per worker in Brazil has grown at a rate of 4.04% over the period 1950–1980 and only 0.16% in 1980–2010. In the long run, the output per worker is the primary source of per capita income growth (Krugman, 1997). In turn, as this study documents, per capita income growth is still an important driver of poverty reduction in many circumstances and places (Krugman, 1997; World Bank, 2018, World Bank, 2020; ILO, 2003) [2].

Poverty reduction strategies bear on the speed by which economic growth reduces poverty. Thus, it is essential to measure the actual contribution of growth to alleviate poverty. Ravallion and Chen (1997) used a sample of developing countries. They estimated that the growth elasticity of poverty was around 3: 1 percentage point increase in income reduces the proportion of people living below the poverty line by 3%. Latter, World Bank (2000) has found a new estimate for countries, and the number is even lower: 2%. Specifically, Bourguignon (2003) reports that elasticity is around 2 for Brazil.

Most studies in this field use countries as the unit of analysis and do not seek to measure the actual contribution of growth in more disaggregated units. Moreover, these studies pay little attention to the characteristics of data when choosing appropriate methods for inference. Studying cities complement study in countries because cities are fully open economies that enjoy immense mobility of workers, firms and goods. National barriers often bar factor and goods mobility between countries and are less impeditive within cities. Besides, cities are more specialized than states and countries (Glaeser, Scheinkman, & Shleifer, 1995).

This paper aims to test three hypotheses in city growth literature documenting the poverty reduction observed in Brazil and exploring a rich spatial dataset for 5564 Brazilian cities observed between 1991 and 2010. The large sample and our improved econometric methods allows one to better understand and measure how important income growth is for poverty reduction, the patterns of agglomeration and population growth in all Brazilian cities.

The data set for cities used possess three essential characteristics. First, the data for all variables used in the analysis come from a single source, providing homogeneity. Second, these data are subject to endogeneity issues caused by simultaneity between the dependent and explanatory variables. Lastly, spatial autocorrelation is inherent in these data, in which poverty incidence, population and income growth tend to occur in clusters. By ignoring spatial dependency and endogeneity of regressors could lead to inefficient and biased estimates, invalid inferences and wrong conclusions (Anselin & Rey, 1991; Baltagi, Blien, & Wolf, 2012). Regional studies have only recently begun applying improved methods to correctly identify endogenous regressors' causal effects (Baum-Snow & Ferreira, 2015). This paper gives a small contribution to this line of research regarding the prior published studies focused on Gibrat's law (Resende, 2004; Eeckhout, 2004; Rose, 2006; Soo, 2014; Chauvin, Glaeser, Ma, & Tobio, 2017) and agglomeration effects (Chen & Partridge, 2013; Chauvin *et al.*, 2017).

Based on our improved statistical methods and the extensive sample data for all Brazilian cities, we test three widely known hypotheses on city growth literature. A positive correlation is evident between city size and productivity in cities of developed countries, as US cities (Chauvin *et al.*, 2017). However, this pattern of economic growth is still less documented using all Brazilian cities' data.

Further, the random growth hypothesis asserts that the city growth is independent of the number of residents (Eeckhout, 2004; Portnov, Reiser, & Schwartz, 2012). Lastly, the poverty–growth elasticity, initially suggested by Ravallion and Chen (1997) and Chambers and Dhondge (2011), can now be consistently estimated, accounting for endogeneity using a valid set of instruments derived from spatial interactions.

When population growth in a typical city is unrelated to the initial size, according to Portnov, Reiser, and Schwartz (2012), urban development policies aimed to enhance or restrict the size of the population in overpopulated cities are unfeasible objectives. Besides, when statistically significant, agglomeration economies can reduce the efficacy of policies designed to restrict population growth in overpopulated areas, even with negative externalities and congestion (Chauvin *et al.*, 2017).

The essential idea behind the agglomeration economies is that productivity increases with geographic proximity of economic activity, in which real wages are higher in larger and denser cities. For instance, Fujita and Thisse (2003) stress that the mobility of skilled workers and the R&D sector located in few core places in these economies appear to be a solid centripetal force and can generate strong externalities from core to peripheral regions based on additional growth boosted by the agglomeration effects. Chen and Partridge (2013) explain that cities' population density can influence firm productivity and wages because of improved labor market matching.

Further, LeSage and Fischer (2008) argue that higher population density represents urban agglomerations that contain massive human capital stocks as a repository of knowledge and could provide a boost to innovation and adoption of technological progress and economic growth. For instance, the metropolitan regions of Tokyo and New York are good examples of extreme forms of economic agglomerations within nations (Fujita & Thisse, 2009).

Being a country of migrants (Fiess & Verner, 2003), Brazil enjoys reasonable mobility of factors across different places and regions. This national characteristic contributes to reducing spatial inequality by improving economic opportunities for skilled and unskilled labor across locations (Fiess & Verner, 2003). While the goal of the present paper is empirical, a considerable body of general equilibrium models examines the relationship between exogenous changes in transport costs for goods and people (commuting costs) across places and its effects on population and income distribution across locations (Redding & Turner, 2015). Theoretically, improvements in transportation infrastructure, such as reducing commuting time across places, have uneven effects on wages, land prices and the size of cities (i.e. population). Workers' mobility across locations matters to ensure they have the same effect on welfare across populated locations (Redding & Turner, 2015). Summing up, empirically, the above theories can be formally tested via the following hypotheses, i.e.

H1. The response of poverty to economic growth is unimportant for cities.

H2. The city growth is independent of the number of residents.

H3. Absence of correlation between population density and productivity.

The layout of the paper is the following. Section two presents the econometric models and describes data. Section three presents and discusses the main findings, comparing them with related works, and Section four provides brief conclusions.

2. Econometric methodology

2.1 Econometric models

According to Ravallion and Chen (1997), and Chambers and Dhondge (2011), the relationship between poverty, economic growth and inequality, in its basic form, can be examined using the model given by

$$p_i = \alpha + \beta g_i + \phi \ln G_i + u_i, \tag{1}$$

$i = 1, \ldots, 5564$; where p_i denote the poverty rate for city i, β is the poverty elasticity measure, g_i denote the rate of economic growth for city i, $\ln G_i$ is the natural logarithm of Gini index for city i. Figueiredo and Laurini (2016) observe that there are two main difficulties in estimating poverty elasticity using the parametric approach we adopt here. First, as is common in a non-experimental environment, there is a likely simultaneity between poverty and income growth, leading to inconsistent estimates in least squares (LS) regression. Second, the degree of inequality can indirectly affect economic growth. The above specification does not account for this indirect effect.

We employ the instrumental variable (IV) estimator in a way that can solve both limitations of LS regression models simultaneously. We apply the Wu–Hausman test for endogeneity of economic growth, and we use as instruments for the endogenous economic growth in the city i the inequality level and economic growth in neighboring cities, $\ln WG_i$ and Wg_i. Our approach is similar to Baltagi, Blien and Wolf (2012) that use spatially lagged unemployment in a wage equation seeking to capture the labor market situation in neighboring regions. Following Baltagi *et al.* (2012), we employ the spatial lagged income growth and the spatial lagged inequality as valid instruments in the IV estimator. As a piece of complementary evidence, we adopt the limited information maximum likelihood (LIML) estimation method because we have a large sample of data. The LIML method delivers consistent estimates in the presence of endogenous regressors (Koutsoyiannis, 2013).

Further, we employ Staiger and Stock's (1997) approach in detecting weak instruments in the usual IV estimator. They consider instruments being relevant only when F-statistic is greater than 10. Lastly, we verify whether our instruments are valid using the Sargan test statistic for overidentification. The Wu–Hausman is a test for detecting endogeneity. When we observe a significant statistic value, it suggests that least squares regression method is inconsistent, and the IV estimator is consistent and shall be preferred.

The W matrix is an n by n non-stochastic inverse great circle distance, non-negative spatial weight matrix whose elements specify the strength of spatial dependence among the cities. If the city i is related to city j, then $w_{ij} > 0$. Otherwise, $w_{ij} = 0$, and the diagonal elements of W are set to zero as a normalization standard. Because row-sums to unity, Wg_i contains a linear combination of income growth from related cities. Wg_i captures the spatial dependence in g_i, and the variable $\ln WG_i$ captures the spatial dependence in the inequality index in neighboring cities.

According to Chauvin, Glaeser, Ma, and Tobio (2017), Eeckhout (2004), and Rose (2006), one can test Gibrat's law by regressing the city population growth on the initial level of the city size. The drawback of these works is that all they assume that cities are "float islands" in space (Fujita & Thisse, 2009), without testing this assumption. Further, all these works assume that the size of a city is exogenous concerning the population growth. Since we are studying a small fragment of an economy, this is also an unwarranted assumption. We include a control for spatial autocorrelation in addition to testing for endogenous population size. To study Gibrat's law, following Eeckhout (2004), in its basic form, the model is given by

$$n_i = \gamma + \pi \ln S_i + \mu_i, \tag{2}$$

$i = 1, \ldots, 5564$; where n_i denote the rate of population growth for city i, $\ln S_i$ is the natural logarithm of population size of a city i, π is the size elasticity, supposed to be zero if the Gibrat's law holds. However, the city growth rates in an environment where cities are not "float islands" in space (Fujita & Thisse, 2009) reads

$$n_i = \theta + \delta \ln S_i + \lambda \ln WS_i + e_i, \tag{3}$$

where the main parameters of interest are δ and λ, the variable $\ln WS_i$ captures the spatial dependence in $\ln S_i$ expressing the influence of size in neighboring places on city i, since labor mobility between nearby cities may provide better access to job opportunities, lower prices of housing, etc.

Assuming exogenous population density and independence across cities, we begin testing for agglomeration economies in all Brazilian cities data using the model specification suggested by Chauvin *et al.* (2017), given by

$$y_i = \tau + \chi \ln den_i + \varepsilon_i, \tag{4}$$

$i = 1, \ldots, 5564$; where y_i denote the natural logarithm of per capita income for city i, $\ln den_i$ is the natural logarithm of population density of a city i. If spatial autocorrelation is present in data and simultaneity exists between population density and real income level, it implies inconsistent LS parameter estimates. In this case, using an IV estimator, we may test the agglomeration effects employing the model given by

$$y_i = \eta + \omega \ln den_i + \kappa \ln Wden_i + v_i, \tag{5}$$

where the main parameters of interest are ω and κ, the variable $\ln Wden_i$ captures the spatial dependence in $\ln den_i$, that gives a measure of market dimension and economic opportunities in neighboring places. The population density can be seen as expressing a market potential for firms and workers (Fujita & Thisse, 2009). We follow Glaeser, Scheinkman, and Shleifer (1995) in using income growth to capture changes in productivity growth, even acknowledging that it also partially captures some changes in quality of life in cities.

2.2 Data description

The data used in this study come from two sources. First, the economic and social attributes of cities are freely available in the Atlas do Desenvolvimento Humano home page at http://www.atlasbrasil.org.br/acervo/biblioteca. The second data source is the Instituto de Pesquisa Econômica Aplicada that provides the shapefile for all Brazilian cities. The shapefile conveys all geographical information to generate choropleth maps of economic and social attributes and create the W matrix and spatial instruments used in the analysis. These geographical information are freely available at https://www.ipea.gov.br/ipeageo/malhas.html from Instituto de Pesquisa Econômica Aplicada. Thus, we have cross-sectional spatial data for 5564 Brazilian cities observed in 1991 and 2010 that vary within the cities and space.

Despite short-run crises and business cycle fluctuations, Brazil has been doing well in social progress when considering decades or more. Figure 1 exhibit a choropleth map, in which areas are shaded in proportion to the poverty rate, showing a decrease in poverty incidence after three decades of economic growth and distributional policies aiming to reduce poverty and inequality.

The noteworthy feature is that poverty is not random in space and tends to occur in clusters, as for the population growth and per capita income growth (not shown because space constraints). For brevity, we put in Appendix all summary statistics and the description of the variables (see Table A1) and the results of statistical tests for spatial autocorrelation (see Table A2) based on the global Moran's statistic (Bivand, Müller, & Reder, 2009; Anselin & Rey, 1991).

Figure 1 supports the conclusion about a substantial decline in poverty rates in all regions, mainly North and Northeast. The blue color spreads over from the Southeast to the North and Northeast, and the yellow areas decreased sharply. Summary statistics in Table A1 show that

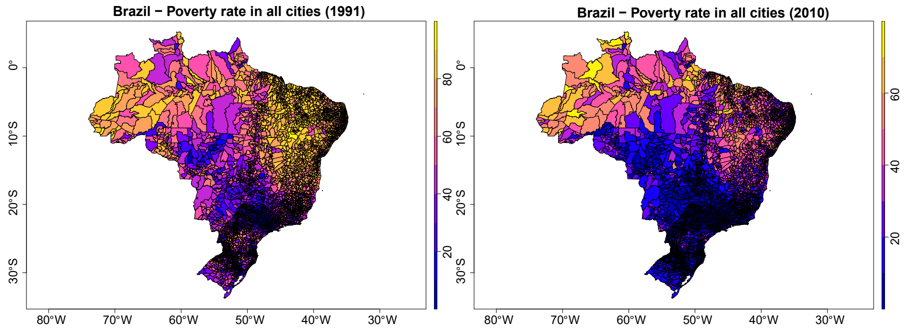

Figure 1.
Incidence of poverty in all cities of Brazil, 1991–2010

Source(s): Author's elaboration from data

Brazilian cities have grown around 30.1% in size, and the agglomeration index has increased by 30.2%, as shown by the greater population density. The poverty rate has decreased from 56.7% to 23.2%, showing signs of convergence across cities.

Following LeSage and Fischer (2008), we employ all the explanatory variables at the beginning of the sample period to minimize the simultaneity issues. As done in most growth regression literature, we also include the initial income level when estimating poverty elasticity to economic growth (LeSage & Fischer, 2008; Glaeser *et al.*, 1995). Even proceeding this way, the simulation study of Reed (2015) shows that this standard practice in lagging explanatory variables to avoid simultaneity in regression models does not prevent neglected endogeneity and implied inconsistent estimates. In this context, the Wu–Hausman test statistic is essential in testing the exogeneity assumption of a potential regressor.

3. Results and discussion

Table 1 displays the results for poverty elasticity under the two assumptions: exogenous economic growth and absence of spatial autocorrelation (LS estimates), beyond the tests for endogenous economic growth (Wu–Hausman statistic) and the estimates accounting simultaneously for all features present in data using IV estimator with the robust variance-covariance matrix estimator of Cribari-Neto (2004). The data reject the assumption of the absence of spatial autocorrelation in poverty rates (see Table A2), and the Wu–Hausman test statistic also rejects exogenous economic growth assumption at 1% level. Sargan test statistics show that instruments are valid and, according to Staiger and Stock (1997) criterion, the chosen instruments are relevant. The chosen instruments are relevant because the F-statistic is higher than 10. The Sargan p-value is higher than the 0.05 or even 0.10 significance level showing that the instruments are valid.

The poverty elasticity based on the LS estimator is biased, inconsistent and outside the bootstrap confidence interval ($\widehat{\beta}_{LS} = -11.98$). The LS-biased estimates likely reflect neglected endogeneity and uncontrolled spatial autocorrelation. Conversely, the consistent IV parameter estimates suggest that when the per capita income of a typical Brazilian city rise 1% point, the poverty rate declines 4.4%, on average. The LIML estimate supports the standard IV findings providing roughly the significant similar parameter estimate ($\widehat{\beta_{LIML}} = -4.35$). This result is in line with other works in the literature (see discussion below). The 95% bootstrap confidence interval based on the IV estimator and the set of valid instruments measures the uncertainty about the point parameter estimates and works as sensitivity analysis when it excludes zero. We employ the bias-corrected, accelerated

Dependent: $p_{i,\ 2010}$	LS	LIML	IV	Boot (BC_a) 95% CI
Constant	218.37***	186.68***	186.78***	[173.78; 196.26]
	(98.19)	(31.92)	(20.42)	–
*lnRDPC*1991	−32.78***	−28.24***	−28.26***	[−29.64; −26.42]
	(−98.39)	(−33.65)	(−21.15)	–
G	*−11.98****	*−4.35****	*−4.37***	[*−6.61*; *−1.35*]
	(−34.22)	(−3.15)	(−1.97)	–
ln *G*1991	9.66***	13.17***	13.16***	[11.36; 15.45]
	(13.44)	(13.10)	(8.46)	–
Weak instruments	–	*48.07****	*48.07****	–
Wu-Hausman	–	–	*46.24****	–
Sargan statistic (p value)	–	*0.60*	*0.60*	–
N	5,564	5,564	5,564	–
F-statistic	12,520***	–	6,997***	–

Note(s): (***) and (**) denote statistical significance at the 1% and 5% levels, respectively; z-statistics are in parentheses. We use as valid instruments for the endogenous per capita income growth from 1991 to 2010 (g_i) the natural logarithm of Gini index in neighboring cities in 1991 (ln *WG*1991), and the per capita income growth in neighboring cities over the same period, *Wg*1091. The Sargan test statistic has null the validity of instruments. The Wu-Hausman statistic tests the absence of correlation between the covariate and the error term (exogeneity). The weak instruments statistic tests the null of the absence of correlation between the instruments and the endogenous variable. In the LS and standard IV methods, we use robust standard errors of Cribari-Neto (2004) that present better performance, especially in the presence of influential observations. We calculate the 95% bootstrap confidence interval based on 10,000 resamples using the IV estimator and the same set of valid instruments employed for point estimates. The values in italics refer to the essential empirical findings
Source(s): Author's elaboration from data

Table 1. Poverty response to growth and inequality estimates (1991–2010) – Eq. (1)

bootstrap percentile interval (BC_a) because it conveniently combines precise point estimates and hypothesis testing in a single inferential statement (DiCiccio & Efron, 1996).

Annegues, Souza, Figueiredo, and Lima (2015) use data for Brazilian cities observed in the years 1991 and 2000 and have found that elasticity of poverty to growth is around 3%. Nonetheless, their estimates ignore endogeneity (simultaneity between growth and poverty), spatial autocorrelation and the indirect effect of inequality on growth. Bourguignon (2003) employs data on growth and poverty rates of 114 countries applying the LS estimator to show that poverty elasticity ranges from −2% to −6.3%, in an improved model. In the improved model, he interacts growth with the initial poverty rate and the initial Gini coefficient. Thus, our poverty elasticity estimate is in full accordance with the range of values found in other studies that apply parametric methods. However, when we consider the characteristics of city data and the suitable methods we adopt, our estimates can generate more reliable and sound conclusions than previous parametric studies.

Ferreira, Leite, and Ravallion (2007) conducted a comprehensive study seeking to reveal what factors help to explain Brazil's growth elasticity of poverty changes using data for 1985–2004. One of the study's main findings is that growth elasticity in the service sector was substantially more poverty reducing than growth in agriculture or industry, accounting for the change in policy regime around 1994. Further, the demise of hyperinflation (price stabilization), opening to international trade, social security and social assistance transfers significantly contribute to reducing poverty.

While Brazilian data show that the output per worker grew by 0.16% from 1980 to 2010, the observed decline in poverty is noticeable in city data: from 56.70% in 1991 to 23.20% in 2010. This considerable decline in the poverty rate can be explained almost exclusively by the magnitude of the poverty elasticity estimated at city-level data because we observe a stagnation of output per worker since 1980 (Costa & Marcolino, 2021). Further, as the main driver for poverty reduction (Ferreira *et al.*, 2007), the service sector has increased its labor share in output in the last decades. According to Costa and Marcolino (2021), by 2010, the labor share in agriculture has decreased to 17%, while the service sector had increased its labor share to 62.5%. The labor share in manufacturing was 20.5% in 2010. Among other factors, the persistent increase in the share of labor in services helps to explain the magnitude of the poverty elasticity in Brazil (Ferreira *et al.*, 2007).

Table 2 shows the results for Gibrat's law for cities under the two basic assumptions found in the literature using LS estimator: exogenous city size and absence of spatial autocorrelation. It also displays the test results for endogenous city size and the estimates accounting simultaneously for all data characteristics using the IV estimator. The data reject the assumption of absence of spatial interactions across population growth in Brazil's cities (see Table A2), and the Wu–Hausman test statistic also rejects exogenous city size assumption at 1% level. Sargan test statistics show that instruments are valid because the p-value is higher than 0.05 or even 0.10 significance level and, according to Staiger and Stock (1997) criterion, the chosen instruments are relevant because the F-statistic is higher than 10.

The response of city growth to initial size in the LS estimator is biased, inconsistent and outside the bootstrap confidence interval ($\widehat{\pi}_{LS} = -0.05$). Like poverty elasticity, the LS-biased estimates likely reflect neglected endogeneity and uncontrolled spatial autocorrelation. More accurate and improved methods using large city samples reject the apparent confirmation of the Gibrat law found in Chauvin *et al.* (2017) and Soo (2014). The IV parameter estimates displayed in Table 2 indicate that city growth suffers significant positive influence from own size ($\widehat{\delta}_{IV} = 0.32$) and negative and significant influence from neighboring cities ($\widehat{\lambda}_{IV} = -3.67$). Data reject Gibrat's law for Brazilian cities. The LIML estimation results support the standard IV findings providing roughly similar parameter estimates. It shows that some cities tend growth at the expense of others in Brazil. These

Dependent: n_i (population growth)	LS	LIML	IV	Boot (BC_a) 95% CI
Constant	0.70***	5.50***	5.50***	[3.02; 10.40]
	(1.96)	(3.44)	(3.17)	–
*lnPOP*1991	*−0.05*	*0.32****	*0.32****	[*0.26*; *0.40*]
	(−1.24)	(7.53)	(8.89)	–
*WlnPOP*1991	–	*−3.68****	*−3.67****	[*−6.06*; *−2.43*]
	–	(−4.65)	(−4.26)	–
Weak instruments	–	*681.39*	*681.39****	–
Wu–Hausman	–	–	*98.98****	–
Sargan statistic (p-value)	–	*0.28*	*0.29*	–
N	5,564		5,564	–
F-statistic	6.71***	–	28.99***	–

Note(s): (***) significant at the 1% level; z-statistics are in parentheses. We use as valid instruments for the endogenous size of population-level in city i the natural logarithm of per capita income level (ln *RDPC*91) in 1991, the natural logarithm of population-level in neighboring cities in 1991 (ln *WPOP*91), and the IDHM level in neighboring cities in 1991 (*WIDHM*91). The Sargan test statistic has null the validity of instruments. The Wu-Hausman statistic tests the null of the absence of correlation between the covariate and the error term (exogeneity). The weak instruments statistic tests the null of the absence of correlation between the instruments and the endogenous variable. In the LS and standard IV methods, we use robust standard errors of Cribari-Neto (2004) that present better performance, especially in the presence of influential observations. We calculate the 95% bootstrap confidence interval based on 10,000 resamples using the IV estimator and the same set of valid instruments employed for point estimates. The values in italics refer to the essential empirical findings
Source(s): Author's elaboration from data

Table 2.
Gibrat's law for population growth estimates (1991–2010) – Eq. (2)

findings suggest that conclusions based on the assumption of independence across cities and exogenous city size are unreliable and must be viewed cautiously.

Regarding Gibrat's law, Chauvin *et al.* (2017) aggregated 5564 Brazilian cities data into 144 microregions and employ the LS estimator to regress population growth on the initial level of the city population. Based on the untested assumptions of exogenous city size and the absence of spatial interactions across places, they conclude that city size has no significant effect on population growth in Brazil. However, data do not support both assumptions. In a closer and related study, Grüdtner and Marques (2020) apply a spatial Durbin model to analyze the relationship between the growth and size of 1,188 cities in the South Region of Brazil between 2000 and 2010. They find that the growth rate of cities is systematically dependent on the number of residents. Data for cities reject the independence assumption between growth and size at the 5% level. Further, after misspecification tests, the authors conclude that a poor model specification (i.e. assuming independence and absence of spillovers) will likely result in false confirmation of Gibrat's law.

Soo (2014) studied Gibrat's law using state-level data of Brazil. He concludes that lognormal distribution fits very well to Brazilian data and interprets it as evidence of Gibrat's law. The problem of this interpretation is that fitting a lognormal distribution is consistent with convergence in population growth, in which a negative and significant relationship between growth and size can generate this pattern (Kalecki, 1945; Portnov *et al.*, 2012), that invalidates Gibrat's law [3].

Soo (2014) also conducts a hypothesis testing procedure to verify whether the lagged population is the best predictor of the current population in Brazil based on a panel data model that accounts for endogeneity but not for spatial dependence, a feature present in state-level data in Brazil (Montenegro, Lopes, Ribeiro, Cruz, & Almeida, 2014). He assumes that Brazilian states do not interact spatially, being isolated entities in space. Neglected spatial dependence in panel models may lead to biased estimates, invalid inference and wrong conclusions (Baltagi & Pesaran, 2007). In this regard, our findings follow closer Portnov *et al.* (2012) conclusions since we also have found that the so-called Gibrat's law for cities can be considered a statistical artifact caused by the misspecified regression models. Black and Henderson (2003, pp. 351–354) have reached similar conclusions for US cities for 1900–1990 period and Grüdtner and Marques (2020) for city data of the South Region of Brazil.

Resende (2004) employs panel unit root tests proposed by Levin and Lin (1992, 1993) and Im, Pesaran, and Shin (2003) to verify whether a unit root is present in a panel of city size for Brazil over 1980–2000. He concludes that the city data panel has a unit root, confirming Gibrat's law for Brazilian cities. There are two main problems with the methods that he employs. First, according to Baltagi, Bresson, and Pirotte (2007), the major criticism of both Levin and Lin (1992, 1993) and Im *et al.* (2003) proposed tests is that they require cross-sectional independence. In Fujita and Thisse's words, these tests require that cities be "floating islands" in space to be valid. As can be seen from Table A2, Brazilian city-level data strongly reject this assumption at the 1% level for city growth. As one could expect from regional data, the growth of one city is significantly related to the growth of neighboring cities. Lastly, the simulation study of Baltagi *et al.* (2007) shows that there can be considerable size distortions in these tests in the presence of spatial autocorrelation. Employing such methods for spatial datasets may lead to unreliable and likely misleading conclusions.

Agglomeration economies can give rise to strong positive or negative externalities (Fujita & Thisse, 2009; Chauvin *et al.*, 2017). Hence, the employed model to measure it must simultaneously capture both direct and spillover effects. By omitting one of them, the estimates may suffer from omitted relevant variable problems (Baltagi *et al.*, 2012). Chauvin *et al.* (2017) employs the LS estimator by regressing natural logarithm of income to natural logarithm of population density for microregions of Brazil observed in 2010. The restricted model they employ does not allow spatial spillovers to measure externalities sign and significance. They report an estimate of 0.026, statistically significant. Exogenous density and the absence of spatial interactions across places are the assumptions of this estimate.

Table 3 shows the results for agglomeration effects under the two basic assumptions typically found in literature (LS): exogenous population density and absence of spatial autocorrelation. It also shows the test results for endogenous population density and the estimates accounting simultaneously for all data characteristics (IV). We conclude that the data reject the assumption of absence of spatial autocorrelation in productivity growth (see Table A2), and the Wu–Hausman test statistic also rejects exogenous density assumption at 1% level. The weak instruments test indicates the chosen instrument is relevant (Staiger & Stock, 1997).

Based on the IV parameters estimate exhibited in Table 3, our primary statistical findings suggest that productivity growth is subject to significant positive effects ($\widehat{\omega}_{IV} = 1.27$) from the agglomeration pattern in Brazilian cities caused by higher density in their own city. Further, as we can expect from theory, the spillover effect is negative and statistically significant ($\widehat{\kappa}_{IV} = -1.32$): it means that some cities grow at the expense of others. We can speculate that skilled workers and R&D firms, enjoying higher mobility, could be one driver for this city development pattern. Further, we use the F-statistic to test whether the impact of agglomeration is equal to one on the own city's productivity a decade later based on the IV method. The test delivered the F-statistic 16.203 with p-value equal to 0.0000. Thus, we may infer that in the Brazilian case, the agglomeration effect is more than proportional, i.e. higher than unity.

Interestingly, like the pattern in population growth, the way agglomeration effects operate also indicates that some cities tend growth at the expense of others in Brazil. When compared to Chauvin *et al.* (2017) estimate, based on exogenous density and independence assumption across places, we observe a substantial underestimation of positive agglomerations effects in Brazil in their findings. The likely reason for that is that data reject the assumptions they employ in estimating such effects. In summary, for a change in density in a neighboring city, the productivity response in a typical city is negative and significant. This finding suggests

Dependent: *lnypc*2010	LS	LIML	IV	Boot (BC_a) 95% CI
Constant	5.91***	6.29***	6.29***	[6.12; 6.46]
	(191.38)	(75.76)	(63.94)	–
*lnDen*1991	*0.06****	*1.27****	*1.27****	*[1.15; 1.43]*
	(6.98)	(18.86)	(15.23)	–
*WlnDen*1991	–	*−1.32****	*−1.32****	*[−1.50; −1.18]*
	–	(−17.29)	(−14.09)	–
Weak instruments	–	*370.67****	*370.70****	–
Wu–Hausman	–	–	*19489.20****	–
N	5,564	5,564	5,564	–
F-statistic	208.90***	–	179.10***	–

Note(s): (***) significant at the 1% level; z-statistics are in parentheses. We use as valid instrument for the endogenous population density in 1991 the natural logarithm of per capita income level (ln *RDPC*91) in the same year. The Wu-Hausman statistic tests the null of the absence of correlation between the covariate and the error term (exogeneity). The weak instruments statistic tests the null of the absence of correlation between the instruments and the endogenous variable. In the LS and standard IV methods, we use robust standard errors of Cribari-Neto (2004) that present better performance, especially in the presence of influential observations. We calculate the 95% bootstrap confidence interval based on 10,000 resamples using the IV estimator and the same set of valid instruments employed for point estimates. The values in italics refer to the essential empirical findings
Source(s): Author's elaboration from data

Table 3.
Agglomeration effects estimates (2010–1991) – Eq. (4)

that when a market potential expands in surrounding places, it attracts people and firms from neighboring locations.

We found that density, our agglomeration measure, significantly raises the productivity in a given city and decreases productivity in neighboring locations. Following the interpretation of LeSage and Fischer (2008, p. 295) as the *average total impact on an observation*, an autonomous decrease in the density of a set of neighbors raises the productivity in a typical city in Brazil significantly. According to Redding and Turner (2015), the cost of transporting people is a significant source of changes in density. Lowering commuting costs induces people to leave over-populated areas, likely reducing adverse agglomeration effects in a set of neighbors and boosting output per worker in a typical city. As the output per worker is the primary determinant of the per capita income in the long run (Krugman, 1997), the data suggest likely a new indirect source of alleviating poverty.

Regarding the movement of people within and between cities in Brazil, city's data for 1995–2008 (Carvalho & Pereira, 2009, p. 88) show an undesirable trend in which household per capita income lags behind the public transport fare (bus and subway) in urban places, harming the job opportunities and welfare distribution across cities [4]. From the above-reported empirical evidence and considering some theoretical predictions from the general equilibrium model deployed for urban development (Redding & Turner, 2015), we conclude that a reasonable way to boost job opportunities and contribute to alleviating poverty in cities would be to invert this unfortunate trend observed in public transport fare. Lowering commuting costs increases real income, worker mobility across urban locations and labor supply in each city (Redding & Turner, 2015). Further, there is evidence that adverse agglomeration effects (e.g. congestion costs) are sensitive to changes in commuting costs. Specifically, reducing commuting costs may induce the population to migrate from the central, heavily populated areas to lower-density neighboring areas (Redding & Turner, 2015).

In general, these findings shed some light on the other studies that ignore simultaneity bias and patterns of spatial dependence across cities and regions. Brazilian city-level data reject the assumptions of exogenous poverty rates, exogenous population growth and exogenous population density regarding productivity at a 1% level. The studies that ignore both characteristics of data deserve to be reviewed and read with caution.

When we account simultaneously for spatial autocorrelation, endogeneity caused by simultaneity between dependent and explanatory variables, and heteroskedasticity of unknown form, Brazilian city-level data reject all three hypotheses of our study significantly. Hence, we can extract three lessons from our study.

First, productivity and income growth, even accounting for the indirect effect of inequality on economic growth, remain an essential source of poverty reduction in sustainable basis for Brazil. Second, data show that the random growth hypothesis is not suited to describe population growth in Brazil city-level data. Hence, it does not follow that population growth is difficult to predict or anticipate for urban planning in Brazilian cities overall.

Lastly, agglomeration effects play a statistically and positive role in own city productivity, but negative externalities are underlying the city development patterns. We can interpret this finding by saying that most Brazilian cities are growing at the expense of the others when they work as a centripetal force for skilled labor and R&D firms, for instance.

4. Conclusions

This paper explored spatial cross-section data of 5564 cities of Brazil to test three widely known hypotheses on city growth literature. We employed tests for endogeneity, spatial autocorrelation, and our inference method utilizes a variance-covariance matrix robust to autocorrelation and heteroskedasticity of unknown form. In addition, we have used the 95%

bootstrap confidence interval based on the IV estimator to measure the uncertainty about the point parameter estimates for robustness.

Brazilian city-level data reject all three hypotheses of our study significantly. Hence, we can extract three lessons from our study when we account for all essential data characteristics. First, economic growth remains an essential source of poverty reduction in sustainable basis for Brazil, accounting for the indirect effect of inequality on growth. Second, data show that the random growth hypothesis, also called Gibrat's law, is not suited to describe population growth in Brazil city-level data.

Lastly, agglomeration effects play a significant and positive role in own city productivity, while negative externalities are underlying the city development patterns. Thus, most Brazilian cities are growing at the expense of the others when they work as a centripetal force for skilled labor and R&D firms, for instance .

Notes

1. See Gupta & Gupta (2020) for a comparison of Brazil to China, India, Indonesia, Mexico, and Vietnam.
2. (IMF, 2021) summarizes and discusses much of the recently published works on productivity growth.
3. Some authors argue that when city size distribution fits a lognormal distribution, it proves Gibrat's law for cities (see Giesen, Zimmermann, & Suedekum, 2010). This conclusion does not follow because lognormal distribution can be generated by a convergence pattern in cities, in which there exists a statistically significant inverse correlation between growth and city size (Kalecki, 1945; Portnov *et al.*, 2012). Hence, fitting a lognormal distribution to city size does not provide empirical evidence of Gibrat's law. As observed by Portnov *et al.* (2012), "Lognormality can, however, hold even where proportionate growth does not. Thus log-normality does not provide sufficient evidence for Gibrat's law."
4. The data for 1995–2008 show that even the inflation rate lags behind the public transport fare (bus and subway) in Brazilian metropolitan cities (Carvalho & Pereira, 2009, p. 86).

References

Annegues, A. C., Souza, W. P. S. F., Figueiredo, E., & Lima, F. S. (2015). Elasticidade de pobreza: apliacação de uma nova abordagem empírica para o Brazil. *Planejamento e Políticas Públicas*, *44*, 145–166.

Anselin, L., & Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. *Geographical Analysis*, *23*, 112–131.

Bacha, E.; Bonelli, R. (2016). Coincident growth collapses: Brazil and Mexico since the early 1980s, *Novos Estudos CEBRAP, São Paulo*, *35*(2), 151–181.

Baltagi, B. H., & Pesaran, M. H. (2007). Heterogeneity and cross section dependence in panel data models: Theory and applications. *Journal of Applied Econometrics*, *22*, 229–232.

Baltagi, B. H., Bresson, G., & Pirotte, A. (2007). Panel unit root tests and spatial dependence. *Journal of Applied Econometrics*, *22*, 339–360.

Baltagi, B. H.; Blien, U.; Wolf, K. (2012). A dynamic spatial panel data approach to the German wage curve. *Economic Modelling*, *29*, 12–21.

Baum-Snow, N., & Ferreira, F. (2015). Causal inference in urban and regional economics. In Duranton, G., & Henderson, V. (Eds.), *Handbook of Regional and Urban Economics* (Vol. 5A, pp. 3–68). Oxford: Elsevier.

Bivand, R., Müller, W. G., & Reder, M. (2009). Power calculations for global and local Moran's I. *Computational Statistics and Data Analysis*, *53*, 2859–2872.

Black, D., & Henderson, V. (2003). Urban evolution in the USA. *Journal of Economic Geography*, *3*, 343–372.

Bourguignon, F. (2003). The growth elasticity of poverty reduction: Explaining heterogeneity across countries and time periods. In Eicher, T. S., & Turnovsky, S. J. (Eds.), *Inequality and Growth* (pp. 3–26). Cambridge: MIT Press.

Carvalho, C. H. R., & Pereira, R. H. M. (2009). Efeitos da variação da tarifa e da renda da população sobre a demanda de transporte público coletivo urbano no Brasil. *Regional, urbano e ambiental*, (3), 85–92.

Chambers, D., & Dhondge, S. (2011). A non-parametric measure of poverty elasticity. *Review of Income and Wealth*, *57*(4), 683–703.

Chauvin, J. P., Glaeser, E., Ma, Y., & Tobio, K. (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics*, *98*, 17–49.

Chen, A., & Partridge, M. D. (2013). When are cities engine of growth in China? Spread and backwash effects across the urban hierarchy. *Regional Studies*, *47*(8), 1313–1331.

Costa, D., & Marcolino, M. (2021). Structural transformation and labor productivity in Brazil. *Revista Brasileira de Economia*, *75*(4), 464–495.

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics & Data Analysis*, *45*(4), 215–233.

DiCiccio, T. J., & Efron, B. (1996). Bootstrap confidence intervals. *Statistical Science*, *11*(3), 189–212.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review*, *94*(5), 1429–1451.

Ferreira, F. H. G., Leite, P. G., & Ravallion, M. (2007). Poverty reduction without economic growth? Explaining Brazil's poverty dynamics, 1985–2004. Policy Research Working Paper 4431, Development Research Group, World Bank.

Fiess, N. M., & Verner, D. (2003). Migration and human capital in Brazil during the 1990s. Policy Research Working Paper 3093, World Bank.

Figueiredo, E. A., & Laurini, M. P. (2016). Poverty elasticity: a note on a new empirical approach. *Review of Income and Wealth*, *62*(2), 394–401.

Fujita, M., & Thisse, J.-F. (2003). Does geographical agglomeration foster economic growth? And who gains and loses from it?. *Japanese Economic Review*, *54*(2), 121–145.

Fujita, M., & Thisse, J.-F. (2009). New economic geography: An appraisal on the occasion of Paul Krugman's 2008 Nobel prize in economic sciences. *Regional Science and Urban Economics*, *39*, 109–119.

Giesen, K., Zimmermann, A., & Suedekum, J. (2010). The size distribution across all cities – Doble Pareto lognormal strikes. *Journal of Urban Economics*, *68*(1), 129–137.

Glaeser, E. L., Scheinkman, J. A., & Shleifer, A. (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics*, *36*, 117–143.

Grüdtner, V., & Marques, A. M. (2020). Is Gibrat law robust when cities interact each other?. *Papers in Regional Science*, *99*, 1087–1111.

Gupta, K., & Gupta, A. S. (2020). Enhancing productivity for poverty reduction in India. *Asian Development Briefs BRF200117-2*, (129). doi:10.22617/BRF200117-2.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, *115*, 53–74.

International Labour Organization (2003). Productivity growth and poverty reduction in developing countries. 2004 World Employment Report, Geneva.

International Monetary Fund (2021). *Boosting productivity in the aftermath of COVID-19*. Group of Twenty. Available from https://www.imf.org/external/np/g20/pdf/2021/061021.pdf

Kalecki, M. (1945). On the Gibrat distribution. *Econometrica*, *13*(2), 161–170.

Koutsoyiannis, A. (2013). *Theory of econometrics*. New York: Palgrave.

Krugman, P. (1997). *The age of diminished expectations: U.S. Economic policy in the 1990s*. Cambridge: MIT Press.

LeSage, J. P., & Fischer, M. M. (2008). Spatial growth regressions: Model specification, estimation, and interpretation. *Spatial Economic Analysis*, *3*(3), 275–304.

Levin, A., & Lin, C. (1992). Unit root tests in panel data: asymptotic and finite-sample properties. Working Paper 92-93, University of California, San Diego, CA.

Levin, A., & Lin, C. (1993). Unit root tests in panel data: New results. Working Paper 93-56, University of California, San Diego, CA.

Montenegro, R. L. G., Lopes, T. H. C. R., Ribeiro, L. C. S., Cruz, I. S., & Almeida, C. P. C. (2014). Efeitos do crescimento econômicos sobre os estados brasileiros. *Economia Aplicada*, *18*(2), 215–241.

Portnov, B. A., Reiser, B., & Schwartz, M. (2012). Does Gibrat's law for cities hold when location counts. *Annals of Regional Science*, *48*, 151–178.

Ravallion, M., & Chen, S. (1997). What can new survey data tell us about recent changes in distribution and poverty?. *World Bank Economic Review*, *11*(2), 357–382.

Redding, S. J., & Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. In Duranton, G., Henderson, J. V., & Strange, W. (Eds.), *Handbook of Regional and Urban Economics* (Vol. 5B, pp. 1339–1398). Oxford: Elsevier.

Reed, W. R. (2015). On the practice of lagging variables to avoid simultaneity. *Oxford Bulletin of Economics and Statistics*, *77*(6), 897–905.

Resende, M. (2004). Gibrat's law and the growth of cities in Brazil: a panel data investigation. *Urban Studies*, *41*(8), 1537–1549.

Rose, A. K. (2006). Cities and countries. *Journal of Money, Credit, and Banking*, *38*(8), 2225–2245.

Soo, K. T. (2014). Zipf, Gibrat and geography: Evidence from China, India and Brazil. *Papers in Regional Science*, *93*(1), 159–181.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586.

World Bank (2000). *Attacking poverty: World development report 2000/2001*. Washington, DC: World Bank.

World Bank (2018). *Piecing together the poverty puzzle*. Washington, DC: World Bank.

World Bank (2020a). *Global productivity: Trends, drivers, and policies*. Washington, DC: World Bank.

World Bank (2020b). *A decade after the global recession*. Washington, DC: World Bank.

Appendix

Summary statistics and Moran statistic results

The economic and social attributes of cities were collected in the Atlas do Desenvolvimento Humano home page at http://www.atlasbrasil.org.br/acervo/biblioteca. We used the following variables:

p_i is the headcount poverty rate, defined as the proportion of households whose real per capita income is equal or less than R$ 140.00 monthly (based on consumer price index at August 2010).

y_i is the natural logarithm of real per capita income of a given city.

G_i: Gini index, it expresses the inequality degree across individuals based on the real per capita income.

g_i is the rate of economic growth, calculated as $(Y_{i,2010} - Y_{i,1991})/Y_{i,1991}$, where Y is the per capita income level.

n_i is the rate of population growth, calculated as $(POP_{i,\,2010}\text{-}POP_{i,\,1991})/POP_{i,\,1991}$.

S_i is the number of inhabitants of a given city.

den_i is the ratio of the number of inhabitants to km^2.

$lnypc2010$ is the natural logarithm of real per capita income in 2010.

$lnDen1991$ is the natural logarithm of population density in 1991.

$lnPOP1991$ is the natural logarithm of the population in 1991.

$IDHM$ is the human development index in a given city.

Variable	Statistic	1991	2010
Poverty rate (%)	Mean	56.7	23.2
	SD	23.6	17.9
City size (population)	Mean	26025	34050.9
	SD	164650.2	201594.8
Per capita income level (R$/population)	Mean	234.9	493.7
	SD	143.6	243.3
Gini index	Mean	0.53	0.49
	SD	0.07	0.07
Density (population/km^2)	Mean	101.2	131.8
	SD	703.9	880.8
Population growth (1991–2010) (%)	Mean	–	27.9
	SD	–	141
Per capita income growth (1991–2010) (%)	Mean	–	131.5
	SD	–	69.5

Source(s): Author's elaboration from data

Table A1. Summary statistics

Null hypothesis	Statistic	Model
Absence of spatial autocorrelation	0.1204*** [0.0000]	Eq. (1)
Absence of spatial autocorrelation	0.0138*** [0.0000]	Eq. (2)
Absence of spatial autocorrelation	0.5687*** [0.0000]	Eq. (4)

Note(s): (***) significant at the 1% level. The p values are into square brackets
Source(s): Author's elaboration from data

Table A2. Results for Moran test statistic

Corresponding author
André M. Marques can be contacted at: andremmarques@yahoo.com.br