Economic Effects of Demographic Dividend in Brazilian Regions

Diogo Baerlocher [*] Stephen L. Parente [†] Eduardo Rios-Neto [‡]

First Draft: October, 2017
This Version: February 5, 2019

Abstract

Exploiting heterogeneity across Brazilian micro-regions over the 1970-2000 period, this paper examines whether the demographic dividend extends beyond a pure accounting effect. Using a Sys-GMM approach, it finds evidence that changes in age structure have only pure accounting effects after controlling for human capital. Therefore, in the case of Brazilian micro-regions, there is a second demographic dividend, which is associated with education. This second dividend is the far more important of the two dividends in terms of economic growth. In a counterfactual exercise, we show that the accounting effect is responsible for less that 10% of the income gap between the poorest and richest regions in Brazil.

Keywords: Demographic Dividend, Education Dividend, Convergence
JEL J11, I25, O47

[*]Corresponding author. Department of Economics, University of Illinois at Urbana-Champaign, David Kinley Hall, 1407 Gregory Dr, Illinois, USA. (email: dcarval2@illinois.edu)

[†]Department of Economics, University of Illinois at Urbana-Champaign, David Kinley Hall, 1407 Gregory Dr, Illinois, USA. (email: parente@illinois.edu)

[‡]Department of Demography and Cedeplar, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, 3o andar, Belo Horizonte, Minas Gerais, Brazil. (email: eduardo@cedeplar.ufmg.br)

1 Introduction

A well-documented feature of the development process is the demographic transition whereby an economy's population growth rate first increases and then decreases due to a lag between the fall in its mortality rate and the fall in its fertility rate. Accompanying this transition is an important change in the age structure of the population whereby the percentage of working age citizens increases. This increase in the share of the working age population is referred to the "demographic bonus" or "demographic dividend" as it has the potential to raise per capita income simply through an accounting effect associated with having more workers to non-workers in the population, and not because of any direct effect on the productivity of factors.

Several researchers argue, however, that the demographic dividend may extend beyond this pure accounting or translation effect, a so-called second demographic dividend. It is easy to think of several channels whereby demographic variables may have growth effects beyond the translation effect.[1] For example, life cycle considerations suggest that an increase in the share of the working age population will lead to an increase in an economy's savings rates associated with physical capital and human capital accumulation.

Early empirical work in this area examined the importance of demographic variables for the growth of nations without attempting to determine if the demographic dividend extends beyond the accounting effect.[2] For example, Bloom and Williamson (1998) looking at the experiences of the East Asian tigers, found that nearly one third of growth in per capita income is attributed to increases in the share of the working age population. For another example, Kelley and Schmidt (2005) reached a similar conclusion in extending the analysis to almost all of Asia.

A second generation of research, however, does try to determine if the demographic dividend extends beyond a pure accounting effect. For example, Cuaresma et al. (2014), using a panel of countries conclude that there is a secondary dividend associated with education. They arrive at this conclusion in three steps. First, they derive a growth rate regression equation from economic theory. Next, they show that the size of the estimated coefficients on the demographic variables in the regression that fails to include lagged education as a control imply an effect larger than the pure translation one. Lastly, they

[1] Lee and Mason (2006) call a *second demographic dividend* the growth induced by the accumulation of factors of production resulting from the changing in age structure. It is usually related to supply-side effects. Kuhn and Prettner (2018) investigate demand-side explanations for the second demographic dividend.

[2] See Williamson (2013) for a recent literature review.

show that this extra effect vanishes once lagged education is added as a control variable. Rentería et al. (2016) arrive at the same conclusion for Mexico and Spain but use a different method, the National Transfer Account (NTA) method.[3]

In this paper, we reexamine the question of whether the demographic dividend extends beyond a pure accounting or translation effect by studying the development experiences of Brazilian micro-regions between 1970 and 2000. Following Cuaresma et al. (2014), we estimate a growth rate equation using the Generalized Method of Moments (GMM), but extend the analysis in two important ways. First, we break down population growth into its natural rate of increase (NRI) component and its net migration component. Second, we decompose the natural rate of increase into a crude birth rate component and a crude death rate component.

These decompositions may provide new insights on how demographic variables affect development. Although both NRI and net migration have the same effect on population growth, it is easy to think that they could have very different impacts on a region's growth of per capita income for a number of reasons. For example, the effects of migration and NRI could be different if migrants are more likely to be young and have higher fertility than the native population. These differences in characteristics would affect the demographics of both origin and destination economies.[4] Moreover, migration may be driven by low-skilled individuals looking for better opportunities or by the dire need to flee adverse natural conditions. Alternatively, it may be driven by demand for high-skilled workers by firms in richer areas (brain drain). Depending on the balance of these forces, migration may have a significant effect on the quality of labor force. Although the literature on the demographic dividend seems to neglect these effects – possibly on account that migration is inconsequential at the national level –, we believe they may be important in the context of regional economies where migration accounts for a higher share of population growth.

Similarly, by breaking down the natural rate into a crude birth rate component and a crude death rate component, new insights can be potentially gained. Given the quality versus quantity trade off, it is natural to think that the crude birth rate may have a greater effect beyond the pure translation effect. Moreover, although the mechanical effect of the

[3]The National Transfer Account method calculates the economic support ratio that weights the age structure by the labor earning and consumption age profiles. The decomposition of the economic support ratio by education segments presents an alternative view on the demographic dividend and its relationship with education.

[4]See Zaiceva and Zimmermann (2016) for a review of the effects of migration on demographic structure in an international context.

crude death rate should be opposite in sign to that of the crude birth rate, improvements in health have the potential to increase economic growth.[5] Therefore, each of these components of the NRI can possibly induce growth effects beyond the translation effect of different magnitudes.

Our analysis both confirms the findings of others and offers some new insights into how demographic variables affect development. Like Cuaresma et al. (2014), we find evidence of both a first demographic dividend and a second demographic dividend, with the latter being associated with education. Quantitatively, the effect of the first dividend is small; in a counterfactual experiment based on our estimation results, we show that the translation effect accounts for no more than 10 percent of the income gap between the richest and poorest regions in Brazil. The second dividend, namely, the changes in age structure that are correlated with the accumulation of human capital, is crucial for explaining the disparities in regional development in Brazil. Furthermore, we show that the negative effect of population growth – and more importantly, fertility – disappears after controlling for education, suggesting important behavioral changes accompanying the demographic transition. In particular, it suggests that the long-run Beckerian trade-off between child quantity (fertility) and child quality (education) is at the heart of some portion of the second demographic dividend in its fertility decline component.[6]

There are a number of virtues to reexamining the question of whether the demographic dividend extends beyond a pure accounting effect using regional data rather than international data. First, as institutions and exogenous shocks, which are clearly important for economic growth, are more homogeneous across regions within a country than across countries, biases that arise from the omission of variables are less likely. Second, as there is far more consistency in the collection of data and definition of variables inside a nation, measurement error is less of a concern.

Despite these advantages, there are very few papers that examine the importance of demographic variables for economic growth using subnational data and that take an econometric approach. Indeed, a contribution of this paper is to fill a gap in the regional development literature that has for the most part neglected the potential importance of the demographic transition for economic growth. One exception is Wei and Hao (2010) who

[5] Bloom et al. (2000), for example, suggest that increases in life expectancy were conducive for growth in East Asian countries.

[6] The other component of the second demographic dividend is associated with increasing savings due to a decline in adult mortality. This does not seem to be relevant for the paper here.

show that fertility has a positive effect on economic growth in Chinese provinces.[7] Our paper is also innovative in that it is the first to take an econometric approach to test for the presence of a demographic dividend in Latin American countries. All the other tests on the demographic dividend in Latin American countries are based on the NTA method.

Of course, at the subnational level, there is the issue of whether there is sufficient heterogeneity across regional units to validate the analysis. Figs. 1 and 2 speak to this point. Fig. 1 shows the national statistics for a number of relevant demographic variables whereas Fig. 2 shows a subset of these statistics for Brazil's five macro regions. Starting with the national statistics shown in Fig. 1, Brazil experienced a demographic transition in the post-World War II period typical of middle-income developing countries, and is currently at the latter stages of this transition. As measured on the right-hand axis, population growth rates have declined and are now converging to replacement rates. The demographic dividend, measured by the share of the working-age population in the total population and measured on the left-hand vertical axis, has risen steadily since 1965 and is expected to peak at around 70% sometime between 2020 and 2025. Turning to Brazil's five macro regions shown in Fig. 2, we observe large differences in the timing of the demographic transition. Although all regions are characterized by declining population growth rates (Panel a), these rates were far lower in the Northeast, Southeast and South than in the North and Central-West in all years of the period. Turning to Panel b, which shows the working age share of the population, the Southeast, South and Central-West display a steady increase since the 1970s. In contrast, the North and Northeast, which start with lower shares, experienced a far less steady increase with an acceleration beginning around 1980. As we shall subsequently document, these differences are more striking at the micro-region level, making Brazil an excellent laboratory for testing the demographic dividend hypothesis.

This paper is divided into five sections. Following this introduction, Section 2 briefly describes the theoretical model that underlies the regression analysis as well as the derivation of the key regression equations. Section 3 describes the data sources and definitions of micro-regions, and then presents summary statistics of the variables used in the regression analysis. Section 4 describes the empirical strategy and reports the results of the regression analysis. Section 5 concludes the paper.

[7] Potter et al. (2002) use Brazilian microregions to show the impacts of economic development on fertility trends. We test this relationship in the opposite direction. Barros et al. (2015) find that change in age structure in Brazil alleviated poverty.

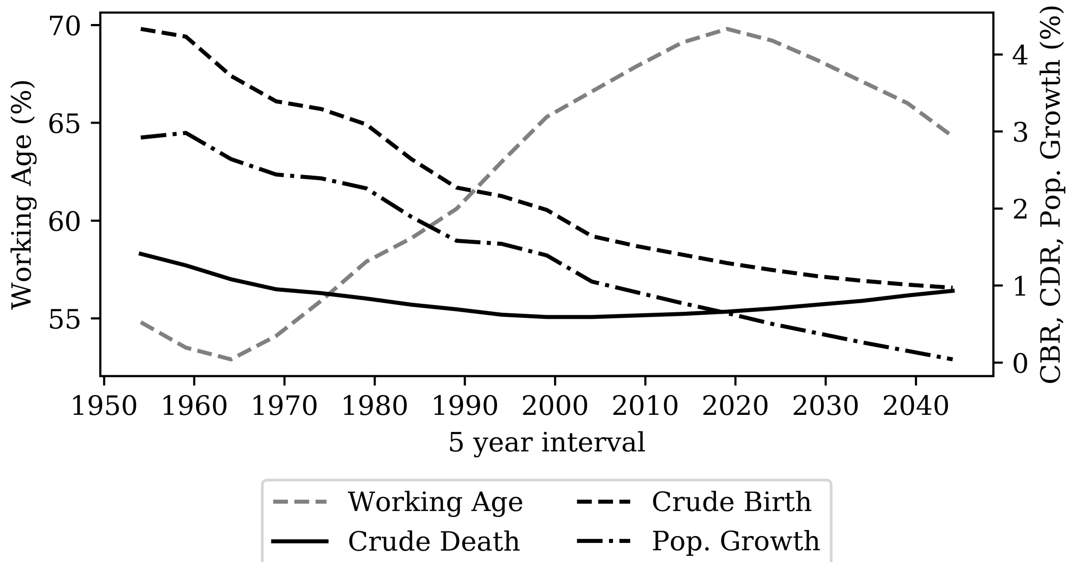

Figure 1: Brazilian Demographic Transition

Note: Data from United Nations – Population division.

2 A Model of Convergence

The theoretical structure that underlies the empirical analysis is the same one utilized by most of the literature on the economics effects of demographic change. Namely, it effectively uses the model of Hall and Jones (1999) to derive a regression equation that gives a region's growth rate of per capita GDP as a function of demographic and economic variables. The model is not truly endogenous or exogenous. [8] For the sake of exposition, we briefly describe the derivation of this main regression equation.

Most growth models, whether they be endogenous or exogenous in nature, make no distinction between the labor force and the population. Hence, some modifications are needed to allow for the age structure of the population to affect income per capita. To understand this, begin with an aggregate production function for country i, $Y_{it} = A_{it} K_{it}^{\alpha} H_{it}^{1-\alpha}$, where Y_{it} is GDP, A_{it} is Total Factor Productivity (TFP), K_{it} is the aggregate physical capital stock, and H_{it} is efficiency units of labor, which is equal to the labor force, L_{it}, multiplied

[8] Mankiw et al. (1992) being an extension of the Solow-Swan model is an exogenous growth model. It differs importantly from the Hall and Jones (1999) specification in that the coefficients on physical capital and human capital in the production function sum to a number less than one.

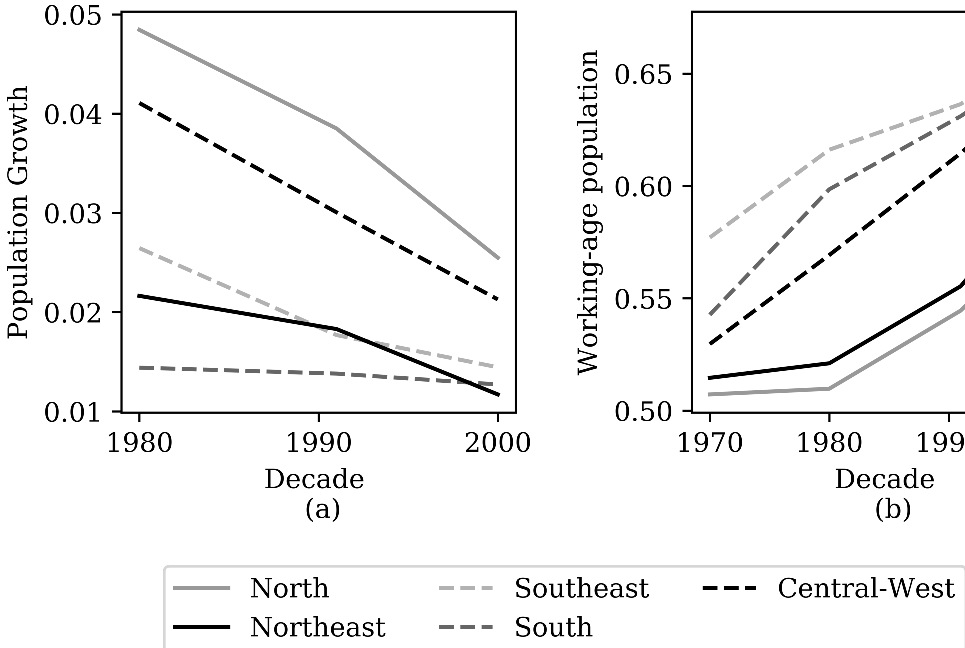

Figure 2: Population Growth Rates Working Age Population and for Brazilian Macro Regions

Note: Data from Brazilian Censuses.

by the human capital of a labor force participant, h_{it}. Unemployment issues are ignored so the labor force is assumed to be the appropriate measure of the raw labor input. Human capital is assumed to be given by the following equation $h_{it} = e^{\theta s_{it}}$, where θ is the return to schooling and s_{it}, is the years of schooling.

The first step in deriving the regression equation is to divide total output by the labor force, so as to obtain the per worker expression, namely, $\bar{y}_{it} = A_{it}\bar{k}_{it}^{\alpha}h_{it}^{1-\alpha}$. To derive the per capita variables, one next multiplies per worker output by the ratio of the population to the labor force, N_{it}/L_{it}. Letting y_{it} denote GDP per capita, one arrives at, $y_{it} = \bar{y}_{it}\frac{L_{it}}{N_{it}} = A_{it}\bar{k}_{it}^{\alpha}h_{it}^{1-\alpha}\frac{L_{it}}{N_{it}}$. From here, one takes the logs of both sides and differentiates to derive the growth rate of per capita income,

$$\frac{\dot{y}_{it}}{y_{it}} = \frac{\dot{A}_{it}}{A_{it}} + (1-\alpha)\Delta s_{it} + \alpha\frac{\dot{\bar{k}}_{it}}{\bar{k}_{it}} - \frac{\dot{N}_{it}}{N_{it}} + \frac{\dot{L}_{it}}{L_{it}}, \tag{1}$$

where $\dot{x}/x \equiv \Delta \ln x$ denotes the growth rate of a variable x. This is the equation used by Cuaresma et al. (2014) in their study on the demographic dividend.[9]

As a matter of principle, a truly endogenous growth model would not consider continuous changes in TFP. However, as a practical matter, it is considered so as to allow for the possibility of catch-up along the lines of Barro (1991). Catch-up depends on how far a region is from the technological frontier with the idea being that regions closer to the frontier find it more difficult to increase their TFPs. The standard approach is to proxy for a region's distance from the technological frontier by its lagged per worker GDP in the TFP growth rate equation. However, for the purpose of considering the effects of demographic variables, per worker GDP is converted to a per capita equivalent in the TFP growth rate equation. Additionally, following Nelson and Phelps (1966) and Benhabib and Spiegel (1994) who argue that better educated societies are more likely to innovate, lagged years of schooling is also added to the TFP growth equation. These considerations give rise to the following equation for the growth rate of a region's TFP,

$$\frac{\dot{A}_{it}}{A_{it}} = \delta + \rho s_{it-1} + \mu \ln y_{it-1} - \mu \ln wa_{it-1} - \mu \ln p_{it-1}, \tag{2}$$

where p_{it} is the participation rate and wa_{it} is the age structure (i.e, working age populuation). Let W_{it} denote the population between 15 and 64 years of age, then $p_{it} \equiv L_{it}/W_{it}$ and $wa_{it} \equiv W_{it}/N_{it}$. To reiterate, the inclusion of the working age share and the participation rate is a consequence of converting per worker GDP to per capita GDP. If the demographic dividend is limited to a pure translation effect, then the demographic variables will have no effect on the growth rate of TFP.

Substituting the growth rate of TFP into the growth rate of GDP per capita equation yields

$$\frac{\dot{y}_{it}}{y_{it}} = \delta + \rho s_{it-1} + \mu \ln y_{it-1} - \mu \ln wa_{it-1} - \mu \ln p_{it-1} + (1-\alpha)\Delta s_{it} + \alpha \frac{\dot{\bar{k}}_{it}}{\bar{k}_{it}} - \frac{\dot{N}_{it}}{N_{it}} + \frac{\dot{L}_{it}}{L_{it}}. \tag{3}$$

Eq. (3) is the basis for three regression models used in this paper to test the effects of demographic variables on economic growth. All three allow for region and time fixed effects. The first regression model is

[9] Other translations are possible. See Kelley and Schmidt (2005) for a discussion.

$$\ln y_{it} = \beta_0 + \beta_1 \ln y_{it-1} + \beta_2 s_{it-1} + \beta_3 \ln p_{it-1} + \beta_4 \ln wa_{it-1} + \beta_5 \Delta s_{it} + \beta_6 \frac{\dot{\bar{k}}_{it}}{\bar{k}_{it}} + \beta_7 \frac{\dot{N}_{it}}{N_{it}} + \beta_8 \frac{\dot{L}_{it}}{L_{it}} + \nu_i + \varphi_t + \epsilon_{it}, \tag{4}$$

where ν_i is a fixed effect of region i, φ_t a time fixed effect, and ϵ_{it} is structural random shock. As is standard in this literature, the dependent variable is transformed into levels to emphasize the autoregressive structure of the model.[10] This transformation implies $\beta_1 = 1 + \mu$. If demographic variables have only an accounting effect on growth, then $\beta_3 = \beta_4 = 1 - \beta_1$ and $-\beta_7 = \beta_8 = 1$. This model is the regression equation used by Cuaresma et al. (2014).

Models 2 and 3 extend Model 3 so as to consider decompositions of some of the demographic variables. The second model decomposes population growth as follows: $\dot{N}_{it}/N_{it} = nr_{it} + nm_{it}$, where nr_{it} is the natural rate of increase (NRI) and nm_{it} is the net migration per capita.[11] The regression equation is then

$$\ln y_{it} = \beta_0 + \beta_1 \ln y_{it-1} + \beta_2 s_{it-1} + \beta_3 \ln p_{it-1} + \beta_4 \ln wa_{it-1} + \beta_5 \Delta s_{it} + \beta_6 \frac{\dot{\bar{k}}_{it}}{\bar{k}_{it}} + \beta_7 \frac{\dot{L}_{it}}{L_{it}} + \beta_8 nr_{it} + \beta_9 nm_{it} + \nu_i + \varphi_t + \epsilon_{it}. \tag{5}$$

For Model 2, if demographic variables have only an accounting effect, then $\beta_3 = \beta_4 = 1 - \beta_1$, $\beta_7 = 1$ and $\beta_8 = \beta_9 = -1$. The third regression model decomposes the NRI so that it is the difference between the crude birth rate (CBR) and the crude death rate (CDR). Namely, $nr_{it} = cbr_{it} - cdr_{it}$, where cbr_{it} is the CBR and cdr_{it} is the CDR. Therefore, we have our third regression equation:

[10] See, for example, Caselli et al. (1996) who started a tradition of GMM estimation of growth models.

[11] The natural rate of increase is calculated as crude birth rates minus crude death rates, whereas net migration per capita is the difference between population growth and natural rate of increase. Details are provided in the next section.

$$\ln y_{it} = \beta_0 + \beta_1 \ln y_{it-1} + \beta_2 s_{it-1} + \beta_3 \ln p_{it-1} + \beta_4 \ln wa_{it-1} + \beta_5 \Delta s_{it} + \beta_6 \frac{\dot{\bar{k}}_{it}}{\bar{k}_{it}} + \beta_7 \frac{\dot{L}_{it}}{L_{it}} + \beta_8 cbr_{it} + \beta_9 cdr_{it} + \beta_{10} nm_{it} + \nu_i + \varphi_t + \epsilon_{it}. \quad (6)$$

Here demographic variables only have a pure accounting effect if $\beta_3 = \beta_4 = 1 - \beta_1$, $\beta_7 = \beta_9 = 1$ and $\beta_8 = \beta_{10} = -1$.

With Eq. (4) we can determine if the conclusions of of Cuaresma et al. (2014) hold for subnational economies with significant heterogeneity. With Eq. (5) we can study the effects of net migration and with Eq. (6) we can study the individual effects of fertility and mortality.

3 Data

Brazilian municipalities are the building blocks for the data set we construct. By definition, municipalities are the smallest independent jurisdictions in Brazil with elected mayors as chief executives and city council legislators as elected representatives. There are over 5,500 municipalities in Brazil, unevenly distributed across states. For example, Minas Gerais has the most with 853 municipalities whereas Roraima has the least with 15. Municipalities are diverse localities with some being rich, fully urbanized areas and others being poor, rural areas.

Starting with these municipalities, we then aggregate the units into micro-regions. Micro-regions are identified based on the definition of homogeneous spaces established by the Brazilian Institute of Geography and Statistics (IBGE) that takes into account natural, social and economic characteristics. There are 558 micro-regions in total in Brazil according to this definition.[12]

The data relating to the demographic variables for these micro-regions are taken from the Brazilian Census of 1970, 1980, 1991 and 2000 whereas the data relating to economic variables are taken from the Institute of Applied Economic Research (IPEA) with the exception of GDP, which is taken from the regional accounts. Economic variables like demographic ones are collected at the municipality level. GDP is the estimate of total

[12] This level of regional disaggregation was used, for instance, by Dix-Carneiro (2014) and Lima and Silveira Neto (2016).

output at factor costs measured in 2000 prices. The labor force is the economically active population and the working-age population is the number of citizens between ages 15 and 64. Years of schooling is the average years of schooling of the population above age 25.

The construction of crude birth rates relies on Census data using Brass P/F method (United Nations, 1983). This method uses single year data on lifetime fertility (alive children ever born) adjusted by recent fertility (alive children born last year) to produce estimates of crude births.[13] Crude death rates are computed using life expectancy data for municipalities. With the life expectancy data in hand, the Coale-Demeny West life table (United Nations, 2012) is applied to calculate mortality rates by age group, which, in turn, are multiplied by the population by age group to compute CDRs by municipality. The data for life expectancy is taken from the Brazilian Atlas of Human Development sourced from IPEA and the population by age is sourced from the Brazilian Censuses.

Given the panel structure of the data and given that many variables in the regression model are expressed in growth rates, we work with the following periods: 1970-80, 1980-91, 1991-2000. Growth rates are therefore expressed as the decade percentage change and not an average annual one. Thus, for each micro-region there are three sets of observations in our panel, one for each of the three periods. Although the CBR and CDR are measured using a single Census, we argue that these variables reflect individual behavior along the previous decade, given, for example, that CBR uses lifetime fertility in its calculation. We than transform these representative annual statistics into decade growth rates.

There are a number of problematic issues regarding data availability. One problem is that there are no measures of GDP available in 1991. To deal with this problem, we construct 1991 measures by interpolating the 1985 and 1996 values assuming a constant growth rate. Another complication is that although there is data by municipality for the residential capital stock, there is no data on the total stock of physical capital. Faced with this dilemma, we use residential capital as a proxy for the total physical capital stock. Whereas this may be a source of mismeasurement, Firme and Simão Filho (2014) actually claim that this proxy delivers more consistent results.[14]

Another important issue in constructing these two panels is that many municipalities were created during the 30 years of analysis. More to the point, some municipalities were split into multiple municipalities in the period. Thus, it is necessary to construct the data for these newly established areas prior to their creations. Towards this goal, we work

[13] el-Badry correction was applied when the share of missing parities were above 2% (el Badry, 1961).
[14] See Lima and Silveira Neto (2016) for detailed discussion.

with Minimum Comparable Areas of 1970 with the strategy of using the data for the unsplit municipality for the earlier years.[15] After making these adjustments and dropping municipalities that are missing data for some years, we arrive at a balanced panel for 528 micro-regions containing 1584 observations.

Fig. 3 depicts the geographical distribution of population among micro-regions in 1970. It is readily apparent that most of Brazil's population is concentrated in the coastal (east) area of the country. The North and Central-West regions are characterized by large but sparsely populated micro-regions mainly due to geographic features such as the Amazon forest and swamplands. In light of this fact, we exclude these two regions from our main sample reducing our balanced panel to 440 micro-regions.[16]

Table 1 presents the means and standard deviations of the main variables of interest across Brazilian micro-regions with standard deviation indicated in parentheses. Over the thirty-year period, average GDP per capita has increased substantially, especially in the 1970s when the so called Brazilian economic miracle occurred. During the 1980s, average GDP per capita decreased with the debt crisis and macroeconomic instability. It recovered mildly in the 1990s associated with greater openness and price stability. Turning to the standard deviations, we see that income inequality between micro-regions has been rising over time, and is positively correlated with average GDP per capita. The average capital stock per worker variable mirrors very closely the GDP per capita variable, especially in terms of averages. The average per worker capital has not increased by as much however over the thirty year period. The years of schooling variable shows a more stable increase than the two other economic variables with no decline in the 1990s. The disparity in years of schooling across micro-regions has increased over the entire thirty-year period, except in the 1990s.

Turning to the demographic variables, the average population growth rate shows a steady secular decline whereas both the participation rate and the working age population show steady secular increases. Although the standard deviation of the population growth rate declines across micro-regions, the disparities in the participation rate and the share of the working age population show no clear trends. The standard deviation of the participation rates at first decreases and then increases. For the share of the working age population, the disparity across micro-regions first increases and then decreases. The average NRI of micro-regions follows a pattern similar to the population growth rate with a steady decline

[15]We use Minimum Comparable Areas identified by Ehrl (2017).

[16]Most results do not change with the full sample, however, specification tests are weaker.

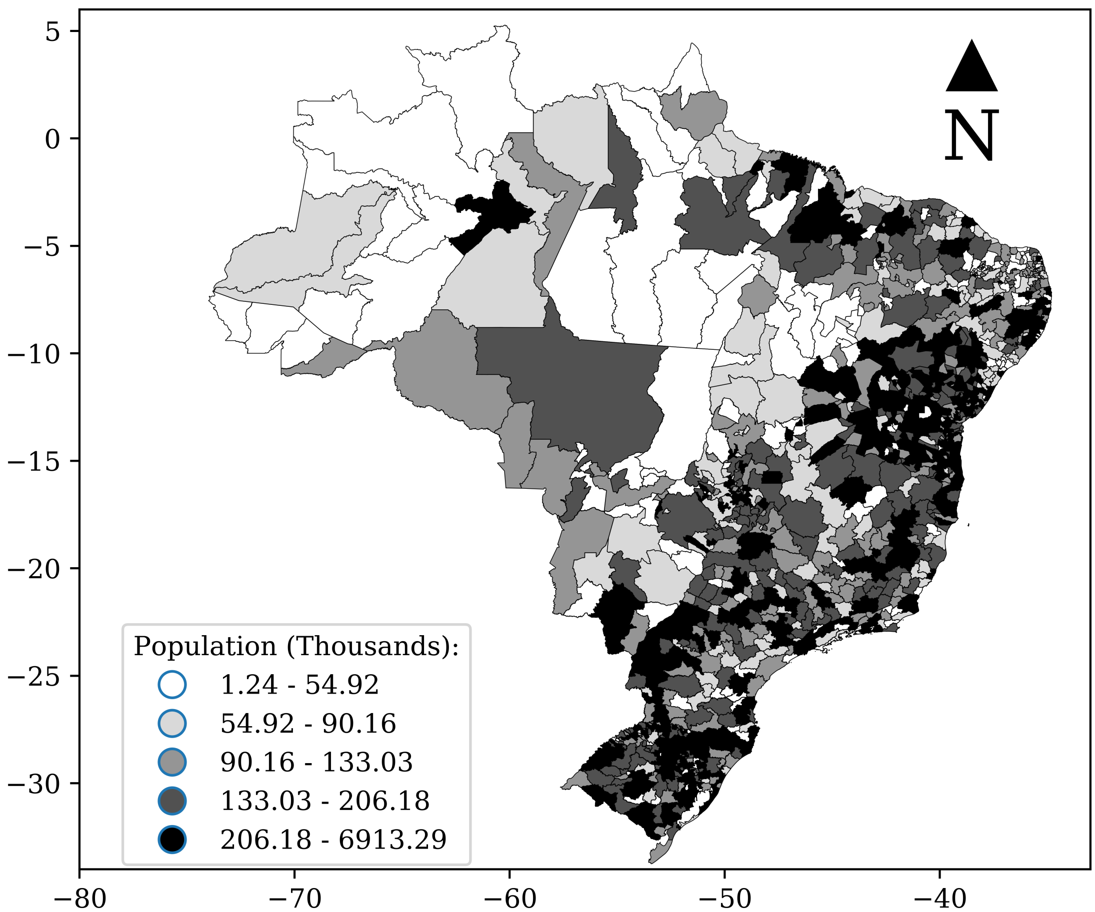

Figure 3: Population by micro-regions in 1970

Table 1: Mean and Standard Deviation for Brazilian micro-regions Variables

	1970	1980	1991	2000	Total
Per Capita GDP	3.437	6.643	6.152	7.083	6.061
	(2.997)	(4.699)	(3.932)	(4.325)	(4.295)
Per Worker Capital	12.32	14.28	13.41	15.91	14.19
	(8.532)	(8.411)	(8.165)	(7.990)	(8.334)
Years of Schooling	2.325	3.414	4.707	5.754	4.306
	(1.311)	(1.574)	(1.703)	(1.652)	(2.032)
Population Growth	–	0.242	0.196	0.135	0.152
	–	(0.195)	(0.129)	(0.0871)	(0.150)
Participation Rate	0.582	0.630	0.659	0.706	0.653
	(0.0537)	(0.0446)	(0.0432)	(0.0486)	(0.0646)
Working-Age Pop.	0.550	0.582	0.608	0.649	0.604
	(0.0462)	(0.0572)	(0.0489)	(0.0379)	(0.0596)
Natural Rate of Increase	0.0108	0.00890	0.00696	0.00421	0.00725
	(0.00653)	(0.00602)	(0.00397)	(0.00384)	(0.00556)
Net Migration	–	0.126	0.102	0.0858	0.0837
	–	(0.214)	(0.119)	(0.0735)	(0.134)
Crude Birth Rate	39.81	32.84	24.03	20.41	27.78
	(8.666)	(7.016)	(5.185)	(3.545)	(9.457)
Crude Death Rate	28.97	23.95	17.07	16.20	20.54
	(5.550)	(4.604)	(4.604)	(3.983)	(6.789)

Means weighted by population. Standard deviations in parenthesis.

throughout. However, the disparity drops in the 1970s and 1980s and remains relatively constant thereafter. The time path of the average NRI is explained by the continuous decline in the CBR that is larger in absolute value than the continuous decline in the average CDR. Finally, average net migration shows a continuous steady decline contributing more to the reduction in population growth compared with the NRI.

To provide a more comprehensive picture of the economic and demographic trends, Fig. 4 plots population growth, NRI, CBR, CDR, share of working-age population and years of schooling for micro-regions. The middle line drawn in each plot represents the average whereas the other two lines represent the 10% and 90% percentiles. Panels (a)-(d) make it clear that Brazil is at the final stages of its demographic transition: CBRs have declined to such a low value relative to CDRs that population growth rates are near the replacement rate. At the same time, this transition has been accompanied by rising education levels (panel e), consistent with a quantity-quality argument, and by increasing shares of the working age population (panel f), implying the potential of a larger demographic dividend.

In sum, the dynamics of the Brazilian regional economies over the last thirty years of the 20^{th} century are characterized by initially strong economic growth followed by a drawn out recession and stabilization. They are also characterized by latter stages of the demographic transition. These transitions have brought about an increases in educational levels and demographic dividends. The process of development is not uniform across regions, however, as there are large differences in both economic and demographic variables in this period across micro-regions.

This brings us to the question of how much of the growth of Brazilian regions is accounted for by changes in demographic variables and how much is accounted for by non-demographic variables. Answering this question is the main objective of the next section.

4 Effect of Demographic Dividend in Brazil

This section first discusses the empirical strategy adopted in this paper and then reports the results of that strategy applied to the regression equations derived in Section 2. The estimation of all three models is challenging given that all regressors are potentially endogenous and given that the lagged dependent variable is correlated with the fixed effect term (Nickell, 1981).

The typical approach to mitigate these potential biases is to use either a Difference

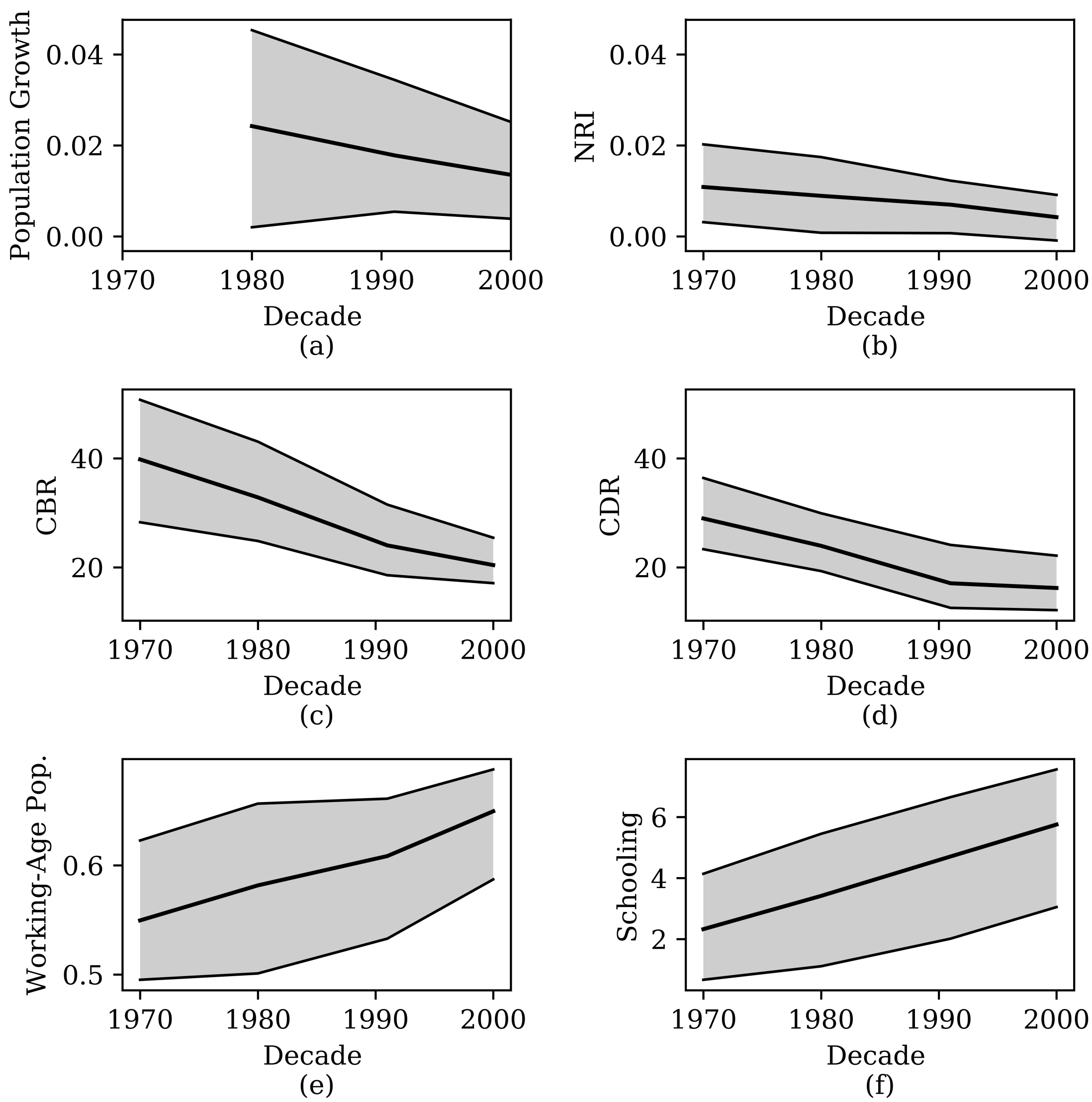

Figure 4: Evolution of Demographic Variables and education in Brazilian Regions

Note: Data from Brazilian Census. The middle line is the average and the lower and upper bounds are the 10^{th} and the 90^{th} percentiles.

GMM (Diff-GMM, hereafter) or a System GMM (Sys-GMM, hereafter).[17] Neither, however, is ideally suited for our purpose as the former performs poorly in short panels, (which is our case), and the latter has questionable extra moment conditions, (which in our case entails assuming that all micro-regions are in their steady states).[18] For completeness, we carry out both estimation procedures, but limit our discussion in the main text to the results from a two-step Sys-GMM that corrects for arbitrary heteroscedasticity and arbitrary within meso-regions correlations.[19] We also report Within-Groups estimators in the main text in order to highlight the nature of the biases, and thus gains from using a GMM approach. The results pertaining to the Diff-GMM as well as several other estimation procedures are contained in the appendix.

Our reason for focusing on the two-step Sys-GMM estimation is based on the relevance of the instruments. The relevance of the instruments are based on three statistics: the Hansen-J, Kleibergen-Paap LM and Kleibergen-Paap rk Wald F. The first tests the validity of overidentification instruments, the second tests for underindentification and the third indicates the strength of the instruments.[20] In the case of the Diff-GMM estimation, these statistics draw into question the validity of the instruments. In this sense, the two-step Sys GMM is our preferred estimation procedure.

We begin by applying the analysis of Cuaresma et al. (2014) to the Brazilian micro-regions. More specifically, following Cuaresma et al. (2014), we estimate the model, namely, Eq. (4), three times using the Sys-GMM, corresponding to the first three columns of Table 2. The first column omits the schooling variables as controls. The second adds the change in schooling over the period as a control whereas the third adds the lagged level of schooling as a control. Recall that Cuaresma et al. (2014) arrive at the conclusion that there is a secondary demographic dividend associated with education by first showing that the coefficients on the demographic variables in an estimation that does not control for education imply an effect larger than a pure translation one. This is the point of Column

[17] Cuaresma et al. (2014) uses the Sys-GMM in a similar application. However, other empirical growth papers, such as Acemoglu et al. (2008) and Cervellati et al. (2014) employ the Diff-GMM.

[18] Hauk and Wacziarg (2009) claim that Sys-GMM may still be a good estimator in practice for short panels.

[19] Meso-regions, which is a regional division in Brazil defined by the IBGE, are larger than micro-regions but smaller than states. We cluster errors by meso-regions in order to capture common regional shocks among micro-regions.

[20] Although Kleibergen-Paap statistics are not designed for GMM methods, we follow the suggestion of Bazzi and Clemens (2013) and use these statistics as a check for the validity of our instruments. See their paper and references therein for a discussion and explanation of these tests.

(1). Then they show that this extra effect disappears from the equation once the level and the change in education are accounted for. This is the point of Columns (2) and (3).[21]

For each regression column, we report the p-values of the Hansen-J test for the validity of the overidentifying restrictions. We also report the p-values for the null hypothesis tests regarding evidence of a pure accounting (translation) effect associated with the demographic change. Recall from Eq. (4) that the coefficients on the growth of the labor force rate and the growth of the population must be 1 and -1, respectively, and the coefficients on the lag of the participation rate and the share of the working-age population must be equal to 1 minus the coefficient on the lagged GDP per capita control if the demographic dividend only has a pure accounting effect. In Table 2 the row labeled "Growth p-value" is the p-value from the test in which $-\beta_7 = \beta_8 = 1$ in Eq. (4); the row labeled "Level p-value" is the p-value from the test that $\beta_3 = \beta_4 = 1 - \beta_1$; and the row labeled "All p-value" is the p-value test in which the null hypothesis is the previous equalities for all coefficients jointly.

Column (1) in Table 2 shows that demographic variables have an important role in explaining economic growth when educational variables are not accounted for. Faster population growth negatively impacts growth of GDP per capita whereas a larger share of the working-age population and faster growth of the labor force positively impact it. All three effects are significant at the 5% level. Turning to the tests for an accounting effect, namely, Columns (2) and (3), although we cannot conclude that the growth related demographic variables, i.e., $\Delta \ln L_{it}$ and $\Delta \ln N_{it}$, have productivity effects, we can conclude that the level related demographic variables, i.e., $\ln p_{it-1}$ and $\ln wa_{it-1}$, do.

It is important to note that the p-value for the Hansen-J test associated with Column (1) is very low suggesting that the moment conditions used in this estimation are not valid. Therefore, one must be very careful in interpreting the coefficients reported and drawing any conclusions. The Hansen-J test result is not at all surprising, however. Recall that, when not included as a regressor, the effect of education is partially captured in the error term. As theory predicts an active relationship between education and demographic variables, one would expect invalid instruments since the lagged demographic variables are used as instruments.[22] In fact, when we include education in the model, Column (2), the

[21] The results of the regression pertaining to the lagged GDP per capita and growth of capital per worker variables are omitted to save space.

[22] This is not the only possible reason to reject the null in the Hansen-J test. As emphasized by Davidson and MacKinnon (2004, p. 368), "a [Hansen-J] test may reject the null hypothesis for more than one reason. Perhaps the model is misspecified. [...] Perhaps, one or more of the instruments is invalid because it is

Table 2: Demographic Dividend for Brazilian Regions: Eq. (4)

	Sys-GMM			Within Groups		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln p_{it-1}$	0.73	0.89	0.62	0.37	0.82^{**}	0.80^{**}
	(0.69)	(0.51)	(0.45)	(0.30)	(0.28)	(0.28)
$\ln wa_{it-1}$	2.06^{**}	1.77^{***}	1.33^{**}	0.57^{*}	1.14^{***}	1.09^{***}
	(0.64)	(0.41)	(0.45)	(0.29)	(0.26)	(0.30)
$\Delta \ln L_{it}$	2.07^{**}	0.12	0.15	0.65^{**}	0.63^{**}	0.64^{***}
	(0.73)	(0.55)	(0.54)	(0.22)	(0.19)	(0.19)
$\Delta \ln N_{it}$	-1.69^{*}	-0.25	-0.45	-0.32	-0.37	-0.38
	(0.73)	(0.35)	(0.38)	(0.25)	(0.22)	(0.22)
Δs_{it}		0.66^{***}	0.69^{***}		0.30^{***}	0.32^{***}
		(0.13)	(0.12)		(0.04)	(0.06)
s_{it-1}			0.08			0.02
			(0.05)			(0.06)
Observations	1320	1320	1320	1320	1320	1320
Micro-Regions	440	440	440	440	440	440
Hansen-J (p-value)	0.00	0.22	0.42			
Growth (p-value)	0.28	0.10	0.28	0.00	0.01	0.01
Level (p-value)	0.01	0.00	0.17	0.07	0.60	0.66
All (p-value)	0.05	0.00	0.12	0.01	0.01	0.01

Robust standard errors clustered by meso-regions in parentheses. Sys-GMM meas that the coefficients are estimated using a two-step efficient Sys-GMM correcting for arbitrary heteroscedasticity and arbitrary within meso-regions correlation. Within groups means that coefficients are estimated using a OLS controling for fixed effects. $\ln y_{it-1}$ and $\dot{\bar{k}}_{it}/\bar{k}_{it}$ are omitted to save space. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Hansen-J test does not reject that the moment conditions are valid. In this case, only the working age share of the population is significant at the 5% level. However, we reject the null that level demographic variables have only accounting effects, and all demographic variables have only accounting effects jointly. This conclusion no longer holds when we add lagged education in Column (3): even though the lagged share of the working-age population still positively affects economic growth, we cannot reject the null that both level and growth demographic variables have only accounting effects.

Columns (4)-(6) repeat the analysis, but use a OLS estimation of the fixed effects model to estimate Eq. (4). Recall that this estimation is potentially biased given the correlation of the lagged dependent variable with the fixed-effect components, and in most cases will be downward in nature. Importantly, Columns (5) and (6) show that the demographic variables do not lose their significance once the level and growth of education are introduced into the regression equation. Therefore, the bias inherent in this procedure is sufficiently large to make one believe that demographic variables have more than a pure accounting effect on the growth of GDP per capita. These results show the importance of using a GMM approach to mitigate the Nickell bias.

Table 2 clearly shows that the accumulation of human capital via schooling is important for understanding regional growth performances. This is consistent with the findings of Acemoglu and Dell (2010) and Gennaioli et al. (2014) that point to human capital differences as the fundamental determinant of differences in regional development.[23] The inclusion of education in the regressions reduces the importance of demographic variables to the point that there is only a translation effect. Moreover, the instruments are valid after education is included suggesting a correlation between demographic and education. In this sense, the demographic dividend extends beyond a pure translation effect. This echos the findings of Cuaresma et al. (2014). Why the demographic dividend is related to human capital growth is another matter. One possibility is the quantity-quality trade-off where fertility declines are related to improvements in the quality of offspring. Another is the quality of migrants that either move in or move out of the region. We investigate these possibility in the next table.

Next, we consider our second model, which decomposes population growth into its natural rate of increase component and its net migration component. These results are

correlated with the error term."

[23] Silveira-Neto and Azzoni (2006) and Lima and Silveira Neto (2016) highlight the importance of human capital accumulation for the development of Brazilian regions.

shown in Table 3. The first three columns repeat the Sys-GMM analysis of Table 2. Starting with Column (1) we see that the negative effect of population growth found in Table 2 is mainly driven by the NRI; net migration has a statistically insignificant effect on growth of per capita GDP. The effect of the lagged share of the working-age population is positive and significant at the 5% level. Importantly, we can reject the nulls that demographic variables only have accounting effects. However, when the accumulation of human capital is added, Column (2), none of the three growth related demographic variables are statistically significant while the effect of the lagged share of the working-age population becomes larger (with an increase in statistical significance as well). Additionally, we can now reject the null hypothesis that the level demographic variables only have accounting effects, although we cannot reject the null pertaining to the growth demographic variables or all demographic variables. Once we introduce lagged education as a control in Column (3), the lagged share of the working age population loses some of its explanatory power. This mirrors the finding in Table 2. Importantly, we cannot reject the null that all demographic variables have only accounting effects.[24] The Within-Groups estimation application is shown in Columns (4)-(6). The results are similar to the results reported in Table 2. Within-Groups estimators are biased and would lead to the conclusion that there is a productivity effect associated with the demographic dividend.

Turning to our third model shown in Table 4, which decomposes the NRI into its CBR and CDR components, Column (1) indicates that only the CBR and CDR variables are statistically significant; the age structure and the change in the labor force each loses its significance once the NRI is broken down into its two components. This is most likely on account that fertility is a direct determinant of age structure. Interestingly, the CDR affects growth negatively. This is inconsistent with the translation effect, which suggests the opposite effect as an increase in the CDR leads to a decline in population growth. The null that the growth demographic variables only have accounting effects is easily rejected. However, these findings change with the introduction of the accumulation of human capital variable in Column (2). In particular, the lagged share of the working-age population now becomes statistically significant at the 5% level while the CBR and CDR lose their significance. Still, we are unable to reject the null that level demographic variables only

[24]We refrain from rejecting the null hypothesis that All demographic variables have only accounting effects when the p-values is 0.05 for three reasons. (1) level and growth tests point to not rejecting this hypothesis for the groups of variables; (2) it makes the results consistent with Table 2; (3) the p-value is sensitive to the model specification as discussed in the appendix. Therefore, we do not believe there is enough evidence to reject the null.

Table 3: Demographic Dividend for Brazilian Regions: Eq. (5)

	Sys-GMM			Within Groups		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln p_{it-1}$	0.51	0.95	0.76	0.42	0.84**	0.82**
	(0.60)	(0.48)	(0.45)	(0.31)	(0.28)	(0.29)
$\ln wa_{it-1}$	1.64*	1.73***	1.53**	0.62*	1.16***	1.07***
	(0.70)	(0.51)	(0.50)	(0.29)	(0.26)	(0.30)
$\Delta \ln L_{it}$	1.48*	0.22	0.14	0.60**	0.60**	0.60**
	(0.75)	(0.57)	(0.49)	(0.21)	(0.18)	(0.18)
nm_{it}	-1.01	-0.17	-0.26	-0.26	-0.33	-0.33
	(0.65)	(0.35)	(0.35)	(0.24)	(0.20)	(0.21)
nr_{it}	-1.86**	-0.04	-0.13	-0.65	-0.58	-0.62
	(0.59)	(0.52)	(0.46)	(0.35)	(0.33)	(0.34)
Δs_{it}		0.69***	0.70***		0.29***	0.32***
		(0.16)	(0.14)		(0.04)	(0.06)
s_{it-1}			0.03			0.04
			(0.05)			(0.06)
Observations	1320	1320	1320	1320	1320	1320
Micro-Regions	440	440	440	440	440	440
Hansen-J (p-value)	0.00	0.12	0.18			
Growth (p-value)	0.20	0.06	0.13	0.00	0.00	0.00
Level (p-value)	0.10	0.00	0.07	0.11	0.60	0.73
All (p-value)	0.04	0.00	0.05	0.00	0.00	0.00

Robust standard errors clustered by meso-regions in parentheses. Sys-GMM meas that the coefficients are estimated using a two-step efficient Sys-GMM correcting for arbitrary heteroscedasticity and arbitrary within meso-regions correlation. Within groups means that coefficients are estimated using a OLS controling for fixed effects. $\ln y_{it-1}$ and $\dot{\bar{k}}_{it}/\bar{k}_{it}$ are omitted to save space. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

have accounting effects. Adding lagged education, Column (3), weakens the effects of the demographic variables, especially the lagged share of working-age population. Importantly, we cannot reject the null that demographic change had only accounting affects on the development process of Brazilian regions. Once again the bias associated with the Within-Groups estimator presented in Columns (4)-(6) would lead to an opposite conclusion.

The conclusion that follows from these exercises is that changes in the age structure that accompany the demographic transition do not explain the economic performance of Brazilian regions beyond a pure accounting effect. Nonetheless, this change is likely to be correlated with the accumulation of human capital, which turns out to be crucial for explaining the disparities in regional development. In this sense, there is a second demographic dividend. The negative effect of population growth – and more importantly, fertility – disappears after controlling for education. This suggests that the gains from declining fertility are associated with more educated offspring, which is in line with the quantity vs quality theory of fertility (Becker and Lewis, 1973). It is also possible that human capital accumulation is causing fertility to decline, since more educated parents have higher opportunity costs of rearing children (Becker, 1981). Net migration, whether we do or do not control for education, has the same effect suggesting that there is only a pure translation effect, or something more complicated whereby there are offsetting effects.[25]

Regardless of the direction of causation, the results strongly suggest that education is a fundamental growth-inducing component of the demographic transition. The estimations that make use of the CDR and CBR data suggest important behavioral changes accompanying the demographic transition. The estimations in their entirety provide strong evidence of a second demographic dividend that exists in the form of education. These findings are certainly in line with a number of unified theories of growth such as Galor and Weil (2000) where the slowdown of population growth is accompanied by investment in education, which, in turn, implies growth of income per capita. Interestingly, we do not find evidence of the Nelson-Phelps theory that a region's TFP growth depends on its human capital, although the introduction of lagged education strengthens the effects of human capital accumulation on growth of per capita GDP.[26]

To provide some perspective of the importance of demographic changes for economic

[25] On the positive side, there is brain drain. On the negative there are migrants who are forced to leave their homes due to natural disasters such as drought.

[26] Lagged education may be insignificant in explaining growth in the case of Brazil regions as TFP growth explains only one fifth of GDP growth (Pinheiro, 1990; Abreu Pessoa et al., 2008). This is in contrast to cross-country studies where TFP growth accounts for nearly half of GDP growth.

Table 4: Demographic Dividend for Brazilian Regions: Eq. (6)

	Sys-GMM			Within Groups		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln p_{it-1}$	-0.78	0.57	0.47	0.44	0.77**	0.75**
	(0.48)	(0.69)	(0.51)	(0.27)	(0.26)	(0.27)
$\ln wa_{it-1}$	0.73	1.39*	1.21*	0.88**	1.23***	1.13***
	(0.39)	(0.54)	(0.51)	(0.30)	(0.27)	(0.32)
$\Delta \ln L_{it}$	-0.90	-0.09	-0.04	0.58**	0.59***	0.59***
	(0.73)	(0.65)	(0.56)	(0.18)	(0.17)	(0.17)
nm_{it}	0.94	0.10	-0.04	-0.19	-0.26	-0.26
	(0.54)	(0.49)	(0.47)	(0.21)	(0.19)	(0.19)
cbr_{it}	-1.52**	-0.30	-0.44	-1.19***	-0.97**	-1.02**
	(0.48)	(0.67)	(0.55)	(0.34)	(0.34)	(0.34)
cdr_{it}	-1.98*	-0.83	-0.64	-0.66	-0.34	-0.30
	(0.87)	(0.78)	(0.71)	(0.41)	(0.41)	(0.42)
Δs_{it}		0.55*	0.56***		0.23***	0.27***
		(0.21)	(0.16)		(0.05)	(0.07)
s_{it-1}			0.04			0.04
			(0.04)			(0.06)
Observations	1320	1320	1320	1320	1320	1320
Micro-Regions	440	440	440	440	440	440
Hansen-J (p-value)	0.07	0.14	0.22			
Growth (p-value)	0.00	0.09	0.21	0.00	0.00	0.00
Level (p-value)	0.05	0.14	0.26	0.11	0.38	0.49
All (p-value)	0.00	0.00	0.05	0.00	0.00	0.00

Robust standard errors clustered by meso-regions in parentheses. Sys-GMM meas that the coefficients are estimated using a two-step efficient Sys-GMM correcting for arbitrary heteroscedasticity and arbitrary within meso-regions correlation. Within groups means that coefficients are estimated using a OLS controling for fixed effects. $\ln y_{it-1}$ and $\dot{\bar{k}}_{it}/\bar{k}_{it}$ are omitted to save space. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

performance we end this section with a simple exercise that is based on the regression results. In particular, we wish to determine how important the translation effect alone is for understanding regional performances. Effectively, this ignores the variations in human capital accumulation that is associated with demographic variables. For this purpose, we focus on the poorest and richest macro-regions in Brazil, the Northeast and Southeast, respectively, which are clearly at very different stages of demographic transition particularly in terms of their shares of the working-age population. Whereas in the Northeast only 55% of population were of working age in 1991, this number was 64% in the Southeast. Let us suppose that demographic variables only have accounting effects so that the difference in the growth of labor force and growth of the population, $\dot{L}/L - \dot{N}/N$, affects the growth of GDP per capita one for one. Keeping labor productivity, physical capital and human capital accumulation constant across the two regions, how much of the gap in per capita GDP would be eliminated if the Northeast had had the same pattern of $\dot{L}/L - \dot{N}/N$ as Southeast? The per capita GDP gaps between 1980 and 2000 predicted by the counterfactual are plotted in Fig. 5. For comparison, the actual gaps are plotted. One can see that the gap would decrease from 3.5 to 3.2 in 1980, a roughly 9% decline. In 2000, the reduction in the gap would be roughly 5%. These are small reductions. We conclude from this that the pure accounting effect of the demographic dividend is of small consequence in accounting for differences in regional living standards. To the extent that demographic variables contribute significantly to regional disparities, it must be through human capital accumulation.

5 Conclusion

Exploiting heterogeneity across Brazilian micro-regions over the 1970-2000 period, this paper has attempted to determine if there is a demographic dividend that extends beyond a pure accounting effect. Using a Sys-GMM approach, it finds evidence of a pure accounting effect, but only after controlling for human capital. Therefore, in the case of Brazilian micro-regions, there is a second demographic dividend, which is associated with education. This second dividend is the far more important of the two dividends in terms of growth. Indeed, in a counterfactual exercise, we show that the accounting effect is responsible for less that 10% of the income gap between the poorest and richest regions in Brazil. Our findings echo those of Cuaresma et al. (2014).

To the extent that demographic variables matter for economic growth, it is through

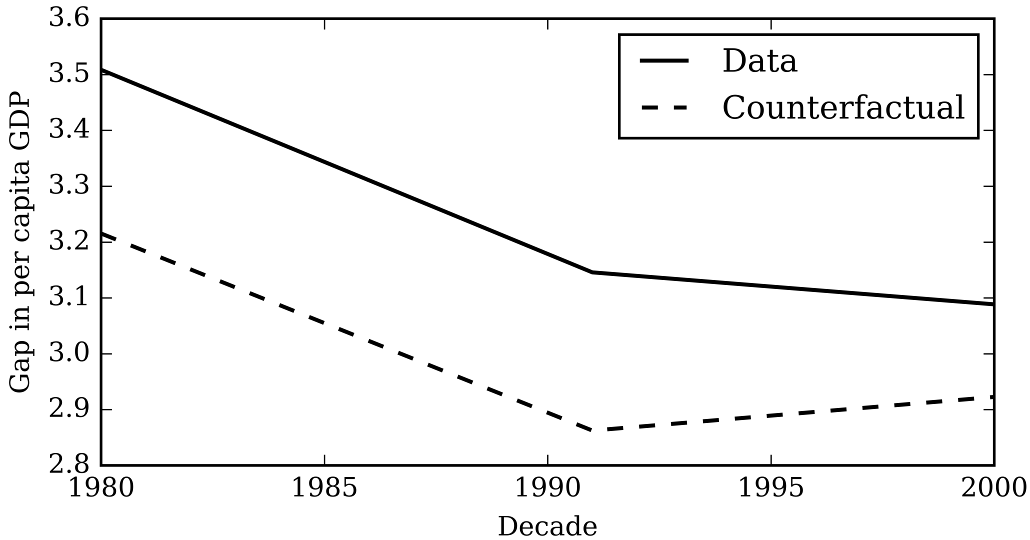

Figure 5: Counterfactual the gap of GDP per capita between Southeast and Northeast regions

their effect on human capital accumulation. This is what the estimations reveal. Demographic variables, both in growth rates and in levels, affect economic growth beyond a pure translation effect only when education is not accounted for. This effect is probably due to bias from the omission of education as a regressor.

We emphasize that we are not claiming that demographic variables cause human capital accumulation. This is not something that we address in these tests. It may be that the causation is from education to demographic variables. All we can conclude from these exercises is that education is a fundamental growth-inducing component of the demographic transition.

In addition to testing for causation, there are a number of future areas of research to pursue in light of the paper's findings. One possible extension is to consider refinements in our measures of changes in the labor force and share of the working age population, such as female labor force participation. As an increase in the labor force participation rate may have a different effect depending on whether it is driven more by women or men.

A Robustness

In this appendix, we discuss the sensitivity of the results reported in the main body of the paper. It is informative to see if our results change if an alternative estimation strategy is used. Tables 5 to 7 report results from a variety of different estimation methods for the regressions that control for lagged schooling and human capital accumulation. Namely, we report the counterpart to the results pertaining to Column (3) of Tables 2 and 3, and Column (6) in Table 3. Column (1) presents a pooled OLS whereas Column (2) presents a Within-Groups OLS (Fixed Effects). In Columns (3) and (4) we present results for the two-step efficient Diff-GMM correcting for arbitrary heteroscedasticity and arbitrary within meso-regions correlation. Column (4) differs from Column (3) in that it collapses the instrument set. The remaining columns use the two-step efficient Sys-GMM. Columns (5) uses the baseline specification whereas Column (6) collapses the instruments. Column (7) treats lagged variables as endogenous and Column (8) considers only the first lag of lagged variables as instruments. Information in the tables include the translation tests, the estimation method, how lagged variables are treated, the number of instruments, whether the instruments are collapsed or not, the p-value of Hansen-J test of validity of instruments, the p-value of Kleibergen-Paap LM test of underidentification and the Kleibergen-Paap rk F statistic.[27]

The reason for estimating the equation using pooled OLS and within-group OLS is motivated by Bond et al. (2001), who noted that the coefficient for the lagged dependent variable is upward biased in the pooled OLS estimation and downward biased in the within-groups OLS estimation. Thus, if the Difference GMM or System GMM coefficient is close or below the within-groups estimate, then the bias caused by the persistence of the times series may be important.

Looking at Table 5, we find that for both the Sys-GMM and Diff-GMM, the coefficient of lagged GDP per capita is between the within-groups and the OLS estimates, and all of them, except in Column (7), are significant. Regarding the quality of our instruments, the Diff-GMMs present low p-values for both the Hansen-J test and the Kleibergen-Paap LM test suggesting that moment conditions are not valid and that only lagged instruments underidentify the endogenous variables. Also, the Kleibergen-Paap rk F statistic is below unit suggesting weak instruments. For the Sys-GMM, Hansen-J tests do not reject the null

[27] Kleibergen-Paap statistics are obtained from two-step least squares IV estimation method using lagged variables in a GMM style. See Bazzi and Clemens (2013) for details.

of valid moments condition at the level of 18% (at least) and, excluding the case where lagged variables are treated as endogenous, the Kleibergen-Paap LM test rejects the null hypothesis of underidentification. Columns (5), (6) and (8) present similar Kleibergen-Paap rk F statistic, although none is above the usual rule of thumb of 5. All three columns have same qualitative and very similar quantitative results with the exception that the lagged share of working-age population is not statistically significant at the 5% level when only the first lag of lagged variables are used as instruments. We conclude that the result present in Column (3) of Table 2 is robust to different specifications that satisfies the specification tests.

Most the analysis of Table 5 is repeated in Table 6, where the results of Eq. (5) are presented. The main difference is that even for the Sys-GMM p-values for Hansen-J tests are lower. However, we fail to reject the moment condition are valid at the 18% level for Column (5) and at the 13% level for Column (8). Again the main specifications – the ones with higher Kleibergen-Paap rk F statistics – present similar quantitative results confirming the robustness of the results presented in Column (3) of Table 3.

The analysis for Table 7, where the robustness of Column (6) in Table 3 is presented, mirrors the ones for previous tables. Columns (5), (6) and (8) present Kleibergen-Paap rk F statistics above 5 and the Hansen-J tests do not reject the null of valid instrument at least at the level of 16%. Again the results in these columns are similar and we conclude the results presented in Section 4 are robust.

Table 5: Robustness: Column (3) Table 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln y_{it-1}$	0.69***	-0.00	0.55*	0.66*	0.54***	0.55***	0.21	0.53***
	(0.04)	(0.06)	(0.25)	(0.33)	(0.08)	(0.09)	(0.19)	(0.08)
$\ln p_{it-1}$	0.27*	0.80**	2.35	2.51	0.62	0.63	2.45	0.39
	(0.12)	(0.28)	(1.52)	(1.68)	(0.45)	(0.56)	(2.31)	(0.57)
$\ln wa_{it-1}$	0.69**	1.09***	2.90	3.00	1.33**	1.28*	3.41*	1.00
	(0.23)	(0.30)	(2.21)	(3.10)	(0.45)	(0.55)	(1.55)	(0.53)
s_{it-1}	0.07**	0.02	-0.13	0.03	0.08	0.07	0.03	0.10
	(0.02)	(0.06)	(0.43)	(0.62)	(0.05)	(0.07)	(0.14)	(0.06)
$\Delta \ln k_{it}$	0.23***	0.06	-0.14	-0.58	0.07	0.03	-0.64	0.05
	(0.05)	(0.06)	(0.52)	(0.63)	(0.20)	(0.26)	(0.77)	(0.21)
$\Delta \ln L_{it}$	0.67***	0.64***	1.46	1.75	0.15	0.20	1.73	-0.10
	(0.11)	(0.19)	(1.51)	(1.76)	(0.54)	(0.64)	(1.71)	(0.67)
$\Delta \ln N_{it}$	-0.62***	-0.38	-1.98	-2.21	-0.45	-0.44	-1.10	-0.34
	(0.12)	(0.22)	(1.31)	(1.43)	(0.38)	(0.46)	(0.80)	(0.43)
Δs_{it}	0.35***	0.32***	0.54	0.76	0.69***	0.66***	0.64***	0.72***
	(0.04)	(0.06)	(0.45)	(0.75)	(0.12)	(0.13)	(0.18)	(0.14)
Observations	1320	1320	880	880	1320	1320	1320	1320
Microregions	440	440	440	440	440	440	440	440
Growth (p-value)	0.01	0.01	0.42	0.42	0.28	0.44	0.81	0.26
Level (p-value)	0.25	0.66	0.39	0.31	0.17	0.34	0.08	0.54
All (p-value)	0.01	0.01	0.17	0.13	0.12	0.29	0.09	0.22
Method	OLS	WG	Diff.	Diff.	Sys.	Sys.	Sys.	Sys.
Lagged Var.			Pred.	Pred.	Pred.	Pred.	End.	Lag 1
Instruments			14	12	27	23	19	25
Collapsed			No	Yes	No	Yes	No	No
Hansen-J p-value			0.01	.	0.42	0.18	0.24	0.36
Kleibergen-Paap LM			0.36	0.35	0.00	0.00	0.41	0.00
Kleibergen-Paap rk F			0.68	0.65	4.12	4.38	0.64	4.38

Robust standard errors clustered by meso-regions in parentheses. Method indicates which method is used: OLS, fixed effects (WG), Diff-GMM (Diff.), Sys-GMM (Sys.). Lagged Var. indicates if lagged variables are treated as predetermined (Pred.), endogenous (End.) or if only one lag is considered (Lag 1). Instruments indicates the number of instruments and Collapsed whether they are collapsed or not. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6: Robustness: Column (3) Table 3

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln y_{it-1}$	0.69***	-0.01	0.52	0.53	0.59***	0.56***	0.20	0.58***
	(0.04)	(0.06)	(0.33)	(0.41)	(0.08)	(0.09)	(0.19)	(0.09)
$\ln p_{it-1}$	0.22	0.82**	1.41	3.15	0.76	0.87	2.45	0.65
	(0.11)	(0.29)	(1.74)	(2.08)	(0.45)	(0.63)	(1.97)	(0.48)
$\ln wa_{it-1}$	0.93***	1.07***	1.05	4.22	1.53**	1.42**	3.79*	1.25*
	(0.23)	(0.30)	(3.25)	(4.21)	(0.50)	(0.52)	(1.48)	(0.56)
s_{it-1}	0.06**	0.04	0.08	-0.33	0.03	0.05	0.02	0.05
	(0.02)	(0.06)	(0.83)	(0.94)	(0.05)	(0.06)	(0.12)	(0.06)
$\Delta \ln k_{it}$	0.23***	0.05	0.31	-0.72	0.02	0.01	-0.53	0.03
	(0.05)	(0.06)	(0.61)	(0.79)	(0.22)	(0.26)	(0.66)	(0.23)
$\Delta \ln L_{it}$	0.72***	0.60**	0.78	2.53	0.14	0.54	1.84	0.10
	(0.11)	(0.18)	(1.56)	(2.10)	(0.49)	(0.75)	(1.67)	(0.51)
nr_{it}	-0.31	-0.62	-1.15	-1.64	-0.13	-0.37	-0.65	-0.11
	(0.16)	(0.34)	(1.15)	(1.33)	(0.46)	(0.59)	(0.90)	(0.45)
nm_{it}	-0.72***	-0.33	-1.48	-2.57	-0.26	-0.52	-0.94	-0.26
	(0.12)	(0.21)	(1.10)	(1.52)	(0.35)	(0.47)	(0.97)	(0.35)
Δs_{it}	0.36***	0.32***	0.81	0.38	0.70***	0.62***	0.60**	0.69***
	(0.04)	(0.06)	(0.88)	(1.02)	(0.14)	(0.14)	(0.22)	(0.14)
Observations	1320	1320	880	880	1320	1320	1320	1320
Microregions	440	440	440	440	440	440	440	440
Growth (p-value)	0.00	0.00	0.37	0.69	0.13	0.64	0.22	0.12
Level (p-value)	0.03	0.73	0.53	0.24	0.07	0.17	0.07	0.33
All (p-value)	0.00	0.00	0.14	0.21	0.05	0.37	0.03	0.14
Method	OLS	WG	Diff.	Diff.	Sys.	Sys.	Sys.	Sys.
Lagged Var.			Pred.	Pred.	Pred.	Pred.	End.	Lag 1
Instruments			17	14	32	26	24	30
Collapsed			No	Yes	No	Yes	No	No
Hansen-J p-value			0.00	0.04	0.18	0.06	0.28	0.13
Kleibergen-Paap LM			0.24	0.08	0.00	0.00	0.16	0.00
Kleibergen-Paap rk F			0.94	0.97	5.58	5.69	1.29	5.39

Robust standard errors clustered by meso-regions in parentheses. Method indicates which method is used: OLS, fixed effects (WG), Diff-GMM (Diff.), Sys-GMM (Sys.). Lagged Var. indicates if lagged variables are treated as predetermined (Pred.), endogenous (End.) or if only one lag is considered (Lag 1). Instruments indicates the number of instruments and Collapsed whether they are collapsed or not. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 7: Robustness: Column (6) Table 3

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln y_{it-1}$	0.66***	-0.02	0.36	0.79*	0.56***	0.53***	0.34**	0.57***
	(0.04)	(0.06)	(0.23)	(0.32)	(0.08)	(0.09)	(0.13)	(0.09)
$\ln p_{it-1}$	0.09	0.75**	2.05	0.91	0.47	0.39	-0.02	0.40
	(0.12)	(0.27)	(1.07)	(1.29)	(0.51)	(0.64)	(1.07)	(0.52)
$\ln wa_{it-1}$	0.64**	1.13***	2.62	0.14	1.21*	1.21*	1.23	1.02
	(0.24)	(0.32)	(1.78)	(2.14)	(0.51)	(0.49)	(0.70)	(0.53)
s_{it-1}	0.05*	0.04	-0.33	0.51	0.04	0.06	0.12	0.04
	(0.02)	(0.06)	(0.47)	(0.59)	(0.04)	(0.06)	(0.07)	(0.05)
$\Delta \ln k_{it}$	0.23***	0.10	-0.05	0.21	0.02	0.10	0.10	0.03
	(0.05)	(0.05)	(0.33)	(0.42)	(0.20)	(0.23)	(0.28)	(0.19)
$\Delta \ln L_{it}$	0.51***	0.59***	1.26	0.45	-0.04	0.26	-0.43	-0.06
	(0.10)	(0.17)	(0.89)	(1.13)	(0.56)	(0.63)	(0.84)	(0.56)
cbr_{it}	-0.68***	-1.02**	-1.25	-1.77	-0.44	-0.61	-0.80	-0.45
	(0.18)	(0.34)	(1.42)	(1.21)	(0.55)	(0.54)	(0.50)	(0.53)
cdr_{it}	-0.56**	-0.30	0.67	1.37	-0.64	-0.52	-1.87	-0.73
	(0.21)	(0.42)	(1.31)	(1.06)	(0.71)	(0.85)	(1.09)	(0.71)
nm_{it}	-0.41***	-0.26	-1.55	-1.32	-0.04	-0.28	0.36	-0.03
	(0.12)	(0.19)	(0.89)	(0.95)	(0.47)	(0.51)	(0.52)	(0.48)
Δs_{it}	0.28***	0.27***	0.30	1.24	0.56***	0.49**	0.44*	0.53**
	(0.05)	(0.07)	(0.54)	(0.64)	(0.16)	(0.15)	(0.21)	(0.17)
Observations	1320	1320	880	880	1320	1320	1320	1320
Microregions	440	440	440	440	440	440	440	440
Growth (p-value)	0.00	0.00	0.74	0.36	0.21	0.52	0.05	0.17
Level (p-value)	0.08	0.49	0.33	0.64	0.26	0.24	0.29	0.47
All (p-value)	0.00	0.00	0.37	0.15	0.05	0.23	0.01	0.08
Method	OLS	WG	Diff.	Diff.	Sys.	Sys.	Sys.	Sys.
Lagged Var.			Pred.	Pred.	Pred.	Pred.	End.	Lag 1
Instruments			20	16	37	29	29	35
Collapsed			No	Yes	No	Yes	No	No
Hansen-J p-value			0.00	0.03	0.22	0.16	0.24	0.19
Kleibergen-Paap LM			0.17	0.09	0.00	0.00	0.35	0.00
Kleibergen-Paap rk F			0.98	1.24	6.16	6.35	1.12	6.49

Robust standard errors clustered by meso-regions in parentheses. Method indicates which method is used: OLS, fixed effects (WG), Diff-GMM (Diff.), Sys-GMM (Sys.). Lagged Var. indicates if lagged variables are treated as predetermined (Pred.), endogenous (End.) or if only one lag is considered (Lag 1). Instruments indicates the number of instruments and Collapsed whether they are collapsed or not. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

References

Abreu Pessoa, S., de Holanda Barbosa Filho, F., and Veloso, F. A. (2008). Evolução da produtividade total dos fatores na economia brasileira com ênfase no capital humano -1992-2006. Proceedings of the 36th Brazilian economics meeting, ANPEC - Brazilian Association of Graduate Programs in Economics.

Acemoglu, D. and Dell, M. (2010). Productivity differences between and within countries. *American Economic Journal: Macroeconomics*, 2(1):169–88.

Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2008). Income and democracy. *American Economic Review*, 98(3):808–42.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, 106(2):407–443.

Barros, R., Firpo, S., Guedes, R., and Leite, P. (2015). Demographic changes and poverty in Brazil. Discussion Papers 96, Instituto de Pesquisa Econômica Aplicada - IPEA.

Bazzi, S. and Clemens, M. A. (2013). Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth. *American Economic Journal: Macroeconomics*, 5(2):152–86.

Becker, G. S. (1981). Altruism in the family and selfishness in the market place. *Economica*, 48(189):1–15.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of political Economy*, 81(2, Part 2):S279–S288.

Benhabib, J. and Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary economics*, 34(2):143–173.

Bloom, D. E., Canning, D., and Malaney, P. N. (2000). Population dynamics and economic growth in asia. *Population and development review*, 26:257–290.

Bloom, D. E. and Williamson, J. G. (1998). Demographic transitions and economic miracles in emerging asia. *The World Bank Economic Review*, 12(3):419–455.

Bond, S., Hoeffler, A., and Temple, J. (2001). GMM estimation of empirical growth models.

Caselli, F., Esquivel, G., and Lefort, F. (1996). Reopening the convergence debate: a new look at cross-country growth empirics. *Journal of economic growth*, 1(3):363–389.

Cervellati, M., Jung, F., Sunde, U., and Vischer, T. (2014). Income and democracy: Comment. *American Economic Review*, 104(2):707–19.

Cuaresma, J. C., Lutz, W., and Sanderson, W. (2014). Is the demographic dividend an education dividend? *Demography*, 51(1):299–315.

Davidson, R. and MacKinnon, J. G. (2004). *Econometric theory and methods*, volume 5. Oxford University Press New York.

Dix-Carneiro, R. (2014). Trade liberalization and labor market dynamics. *Econometrica*, 82(3):825–885.

Ehrl, P. (2017). Minimum comparable areas for the period 1872-2010: an aggregation of Brazilian municipalities. *Estudos Econômicos (São Paulo)*, 47(1):215–229.

el Badry, M. A. (1961). Failure of enumerators to make entries of zero: errors in recording childless cases in population censuses. *Journal of the American Statistical Association*, 56(296):909–924.

Firme, V. d. A. C. and Simão Filho, J. (2014). Análise do crescimento econômico dos municípios de Minas Gerais via modelo MRW (1992) com capital humano, condições de saúde e fatores espaciais, 1991-2000. *Economia Aplicada*, 18(4):679–716.

Galor, O. and Weil, D. N. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American economic review*, 90(4):806–828.

Gennaioli, N., La Porta, R., De Silanes, F. L., and Shleifer, A. (2014). Growth in regions. *Journal of Economic growth*, 19(3):259–309.

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The quarterly journal of economics*, 114(1):83–116.

Hauk, W. R. and Wacziarg, R. (2009). A monte carlo study of growth regressions. *Journal of Economic Growth*, 14(2):103–147.

Kelley, A. C. and Schmidt, R. M. (2005). Evolution of recent economic-demographic modeling: A synthesis. *Journal of Population Economics*, 18(2):275–300.

Kuhn, M. and Prettner, K. (2018). Population age structure and consumption growth: evidence from national transfer accounts. *Journal of Population Economics*, 31(1):135–153.

Lee, R. and Mason, A. (2006). What is the demographic dividend? *Finance and Development*, 43(3):16.

Lima, R. C. D. A. and Silveira Neto, R. D. M. (2016). Physical and human capital and Brazilian regional growth: a spatial econometric approach for the period 1970–2010. *Regional Studies*, 50(10):1688–1701.

Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A contribution to the empirics of economic growth. *The quarterly journal of economics*, 107(2):407–437.

Nelson, R. R. and Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *The American economic review*, 56(1/2):69–75.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.

Pinheiro, A. C. (1990). Measuring and explaining total factor productivity growth: Brazilian manufacturing in the seventies. Discussion Papers 189, Instituto de Pesquisa Econômica Aplicada - IPEA.

Potter, J. E., Schmertmann, C. P., and Cavenaghi, S. M. (2002). Fertility and development: Evidence from Brazil. *Demography*, 39(4):739–761.

Rentería, E., Souto, G., Mejía-Guevara, I., and Patxot, C. (2016). The effect of education on the demographic dividend. *Population and Development Review*, 42(4):651–671.

Silveira-Neto, R. and Azzoni, C. R. (2006). Location and regional income disparity dynamics: The Brazilian case. *Papers in Regional Science*, 85(4):599–613.

United Nations (1983). Manual x: Indirect techniques for demographic estimation. Technical report.

United Nations (2012). *World Population Prospects: The 2010 Revision, Supplementary tabulation*. UN.

Wei, Z. and Hao, R. (2010). Demographic structure and economic growth: Evidence from China. *Journal of Comparative Economics*, 38(4):472–491.

Williamson, J. G. (2013). Demographic dividends revisited. *Asian Development Review*, 30(2):1–25.

Zaiceva, A. and Zimmermann, K. (2016). Migration and the demographic shift. volume 1 of *Handbook of the Economics of Population Aging*, pages 119 – 177. North-Holland.