Weighted convergence in Colombian regions. The role of geography and demography

Jesús Peiró-Palomino
Universitat Jaume I and INTECO

William Orlando Prieto-Bustos
Universidad Católica de Colombia

Emili Tortosa-Ausina
Universitat Jaume I, IIDL and Ivie

October 14, 2019

Abstract

We analyze per capita GDP convergence among Colombian departments between 2000 and 2016 using the distribution dynamics approach. Compared with previous studies, we provide a more complete view by including some additional information such as the asymptotic half-life of convergence, mobility indices and the continuous version of the ergodic distributions. In addition, we also extend the analysis to evaluate whether patterns could differ if weighted by either the population living in each department or their economic sizes, together with the existence and magnitude of spatial spillovers. The unweighted, unconditional analysis corroborates and supplements previous findings, especially those indicating that convergence patterns differ strongly under either pre-2008 or post-2008 trends. Both the weighted and space-conditioned analyses indicate that convergence could be much faster when these factors are introduced in the analysis. Implications are especially relevant when weighting by population, since results indicate that the number of people escaping from relative poverty would be much higher than the figure predicted by the unweighted analysis.

Keywords: Colombia, convergence, distribution dynamics, provinces, weights

JEL classification: C16, O18, O47, R11

Communications to: Emili Tortosa-Ausina, Departament d'Economia, Universitat Jaume I, Campus del Riu Sec, 12071 Castelló de la Plana, Spain. Tel.: +34 964387168, fax: +34 964728591, e-mail: `tortosa@uji.es`.

1. Introduction

Concerns about countries' wealth have triggered off a vast literature on regional growth and convergence. Conclusions as to the validity of the convergence hypothesis vary depending on methodologies, units of study (countries/regions), or sample years. The relevance of the issue has prompted a vast body of literature dealing with the topic, nicely reviewed by Islam (2003) and, more recently, Johnson and Papageorgiou (2019). Although most of this wave of research has focused on international income convergence, regional convergence has become a large area in itself.

If we also factor in the key global fact that the distribution of income in rapid-growth countries has become more unequal (Johnson and Papageorgiou, 2019), these global tendencies would indicate that examining convergence at sub-national levels seems to be as important as the country level. As indicated by Jerzmanowski (2006), over time, growth experiences differ within a country (almost) as much as they differ among countries. Indeed, in some relevant regional contexts such as the European Union, the objective of convergence has involved specific policies (the so-called "cohesion policies") and a large amount of economic resources Sala-i-Martin (1996b); Giannetti (2002); Geppert and Stephan (2008). With a much more limited budget, this is also the case of some developing countries such as Colombia, the country on which we focus, and whose high levels of income disparities are a major concern among policymakers.

Colombia is a highly unequal country with historical economic and social gaps due to disparities in human and physical capital, low-quality institutional settings and civil conflicts that have caused wealth inequities among and within regions (García and Benitez, 1998; Galvis and Meisel, 2010; Galvis-Aponte et al., 2017). In is well-known that great inequalities have an impact on redistributive tax pressures, deterring investment incentives and, ultimately, leading to a more unstable socio-political environments with detrimental effects for economic activities (see, for instance Alesina and Perotti, 1996; Alesina and Rodrik, 1994). For the Colombian case, we can distinguish different regional convergence patterns from 1960 to the mid 2000's. There was a first period of convergence from 1960 to 1980, mainly driven by transport infrastructure investments (Bonet and Meisel, 1999). This was followed by a period of divergence from 1980 to 1990, when the central region was leading economic development (Galvis et al., 2001; Acevedo, 2003a). Finally, diparities persisted from 1990 onwards, when mobility between rich and poor regions barely took place (Bonet and Meisel, 2008). The absence of economic convergence becomes a structural bottleneck to foster equal opportunities for social and economic development in the country, while it shows the poor performance of public policies in providing favorable conditions to push the lagged economies towards a sustainable growth pattern.

Given the limited performance of the more traditional neoclassical model to explain income dynamics among the Colombian regions, some authors such as Cárdenas and Pontón (1995) or Cárdenas (1993) suggested the necessity of alternatives theories able to better explain the Colombian reality. Therefore, endogenous growth models with increasing technological returns to scale based on human and physical capital spillovers postulated as better candidates to explain the evolution of income convergence. In addition, geographical comparative advantages and demographic factors might have better capacity to explain the polarization patterns found that the initial level of income. In this regard, more recent papers by Galvis et al. (2010) and Galvis-Aponte and Wilfried Hahn-De-Castro (2016) have highlighted the role of spatial dependence and neighbor effects, which can be essential for the difussion of the above-mentioned spillovers. The observed trends also reveal that fiscal policy decentralization has not been successful in closing per capita income gaps among central and peripheral regions in Colombia. In response, the new strategies for regional policy are based on a Regional Compensation Fund (RCF) to level up social and economic opportunities. The RCF is a long–term regional development policy proposal based upon human capital investments within a spatial and integrated approach designed to overcome an unequal wealth distribution (see Galvis et al., 2010).

Against this background, this paper examines the complexity of the convergence process in per capita income across the 33 Colombian departments over the period 2000–2016. Unlike previous studies that apply either σ- or β-convergence (which sometimes require strong assumptions) we follow the distribution approach developed by Quah (1993a,b), which allow data to reveal the nature of the relationship of interest by using nonparametric techniques and does not impose any assumption or restriction on the specification of the income distribution. Two are the main contributions of our paper. The first one relates to the sample and the period analized. To our knowledge, this is the first paper considering the 33 Colombian departments. The period considered is also novel–evidence for the last 15 years is nonexistent–and therefore the analysis provides a recent view of the convergence process. As a second contribution, the paper takes into account the role of demography and geography.

Considering the first element, taking into account population matters, as convergence cannot be taking place in geographic terms, but the patterns can differ when considering how many inhabitants live in each region. As indicated by Sala-i-Martin (2006), the unweighted approach is not useful if one is concerned about human welfare, since different provinces have varying population sizes, i.e the actual share of Colombian population living in poverty.

This shift to population-weighted comparisons has evident implications to the importance that we assign to the growth of the largest departments (Schultz, 1998). As for the second element, geographical features such as great mountain ranges and rain forest areas represent frictions that make connections more difficult connections, favoring the isolation of some areas. This can ultimately exacerbate regional disparities and heavily impact on the convergence process. With similar approaches to the ones we consider here, these issues have been examined for both developed (see Tortosa-Ausina et al., 2005, among others) and developing countries (see Herrerías et al., 2011, among others). In the specific context of Colombia, only (Galvis-Aponte and Wilfried Hahn-De-Castro, 2016) has partly dealt with these issues, although from a different point of view.

The results suggest that convergence in terms of GDP per capita is not taking place across Colombian departments in the analized period. In contrast, we observe a bimodal distribution of GDP per capita, with a strong polarization between poor and rich departments more compatible with the concept of *club convergence*. This pattern changes when distributions are weighted by population. For that case, the resulting distribution is clearly unimodal and sharper than the unweighted one, showing a strong convergent process when we account for demography. Similarly, geography is also a relevant element, as convergence is much more evident when departments are compared with their neighbors than with the country mean.

The rest of the paper is organised as follows. Section 2 provides a review of the literature. Section 3 explains the methodology. In Section 5 provide the results and, finally, Section 6 concludes.

Otra contribución que no hemos puesto: 33 departamentos.

2. Background and literature review

The previous literature analyzing convergence in Colombia either focusing on per capita income or other related economic or social variables is relatively large, although most of it is in Spanish—only a few studies have been written in English. In addition, some of the most relevant contributions were published a while ago, therefore missing relevant events that took place in the most recent periods. A review of the latest research on economic and social convergence in Colombia, either focusing on per capita income or other related economic or social variables, has mainly shown a polarized country, a situation that is persistent over time among departments (Galvis-Aponte et al., 2017).

Some of these studies, particularly the oldest ones, adopted σ and β-convergence approaches. This is the case of Cárdenas and Pontón (1995) (see also Cárdenas, 1993; Cárdenas et al., 1993), who evaluated per capita income convergence across departments for the 1950–1990 period, finding a robust convergent pattern. However, this result was not robust across studies, since other authors found that convergence existed in the 1950–1960 period, but not for 1960–1990 (Meisel, 1993). Similarly, Birchenall and Murcia (1997) and, to a lesser extent, Birchenall (2001), considered Quah's distribution dynamics approach, finding weaker evidence supporting convergence. In another relevant study, Bonet and Meisel (2008), also using the distribution approach and with a new database, found that there was no clear pattern towards convergence between 1975 and 2000, and that Bogotá was playing a fundamental role in this process given its size both in population and economic terms.

Besides, Bonet and Meisel (1999) found a significant negative relationship between initial income levels and growth rates and a reduction in the dispersion around the national income average from 1926 to 1960 due mainly to investment on roads and railways around the country. Nevertheless, the convergence trend changed from 1960 to1995, where it showed a polarization in per capita income levels in which Bogotá was the dominant economic force in the country. The main factors behind the polarization process were the import substitution policy implemented to protect the national industry and public consumption, which were more relevant in the capital city.

Furthermore, Rocha and Vivas (1998), Acevedo (2003b), Galvis et al. (2001), and Galvis-Aponte and Wilfried Hahn-De-Castro (2016) showed how factors such as human and physical capital, market imperfections, political stability, international trade, telecommunications infrastructure, among others, matter when explaining regional growth. In this sense, there was a new research focus on the relevance of knowledge externalities altogether with increasing returns to scale that explained why some regions grew faster than others. In this new research trend, both the endogenous and the geographic hypothesis received particular attention together with spatial dependence, spillovers and labor migration, effects that were included in econometric analyses. The results confirmed a higher concentration of economic activity, population and infrastructure in a few cities, located mostly in the central region. In contrast, peripheral regions are left behind, unable to close the regional income gap (Bonet, 2007). Also, Ardila-Rueda (2004) found that the decentralized fiscal policy has not been successful in promoting lower regional gaps. In this sense, regional public investment and regional public consumption only showed positive effects on the relative position of each region within the income distribution, but income distribution remained virtually unaltered between 1985 and 1996.

From a poverty convergence perspective, Galvis et al. (2010) found more evidence of convergence clubs where income inequalities are lower compared to the distribution of all departments around the national average. They also found a polarization trend among convergence clubs, driven by spatial factors that are creating persistent poverty traps in peripheral regions. One of the most recent applications of the distribution dynamics approach (although they also considered σ and β-convergence) to the case of the Colombia is the study by Royuela and García (2015), who have analyzed not only the evolution of per capita income convergence, but extended the analysis to well-being indicators such as life expectancy, infant mortality, educational enrolment and crime issues. Their study, focusing on the period 1975–2005, found different patterns depending on the indicator considered. Despite convergence was found for some social indicators (education, health, crime), per capita income exhibited a divergent pattern, a similar finding to Branisa and Cardozo (2009a) and Franco and Raymond (2009).[1]

3. Methodology

We consider the distribution approach proposed by Danny Quah in a series of contributions. With respect to other methods and concepts, particularly σ and β-convergence, it has the advantage of analyzing how the entire distribution of per capita income evolves. Although some contributions have already considered its application to the Colombian case (see, for the most recent, Royuela and García, 2015), we introduce certain variations in the methodology to provide more painstaking conclusions which had not been considered up to now in this context. The advantages for analyzing the entire cross-sectional distribution of per capita income are multiple and include, for instance, a better ability to detect multi-modality, polarization, or the existence of convergence clubs.

3.1. Densities estimated via kernel smoothing and local polynomials

In the first stage of the model, we report the non-parametric estimation of per capita income density functions via kernel smoothing for different years. A concentration of the probability mass would indicate convergence, while flatter densities would indicate divergence. In addition, a multiplicity of scenarios could also emerge, such as the existence of convergence/divergence clubs (Ben-David, 1994) shown by multi-modal shapes.

In our setting, where $x_{i,t}$ refers to department i's normalized per capita GDP in period t,

[1] Other contributions also considering social indicators are Branisa and Cardozo (2009b), Aguirre (2005) and Martínez (2006).

the corresponding kernel estimator will be:

$$\hat{f}(x) = \frac{1}{Nh}\sum_{i=1}^{N} K\Big(\frac{\|x - X_i\|_x}{h}\Big) \tag{1}$$

where X is departamental per capita income, N is the number of departments, x is the point of evaluation, $\|\cdot\|_x$ is a distance metric on the space of X, h is the bandwidth, and $K(x)$ is a kernel function. Our selection is the Gaussian kernel, which is both relatively straightforward to apply and fits well most contexts.[2] Regarding the choice of the bandwidth, (h), it has a much greater impact than the choice of kernel. We follow the local likelihood variant of density estimation, which allows to overcome important and well-known problems in kernel estimation (see Loader, 1996; Hjort and Jones, 1996). [3]

As shown by Loader (1996), who compare the relative efficiencies of kernel and local log-polynomial methods, they might perform better in several settings such as ours, where several types of densities (unweighted, weighted, spatially-conditioned, ergodic) are considered. Therefore, we consider changes in the local likelihood criterion as follows:

$$\sum_{i=1}^{N} \omega_i(x)\ln(f(X_i)) - N\int W\Big(\frac{u-x}{h}\Big)f(u)du \tag{2}$$

where the log-link is used, i.e., $\ln(f(x))$ is modeled by local polynomials, where W indicates that we specify a locally weighted least squares criterion for each fitting point (x), $\omega_i(x)$ refers to the localization weights, the log-link is used (i.e., $\ln(f(x))$ is modeled via local polynomials), and the term on the right is the added penalty term.[4]

3.2. How densities evolve: intra-distribution mobility

Apart from the evolution of the external shape of the distribution it is interesting to analyze its internal mobility. To do so, and considering our $x_{i,t}$ variable referring to department i's normalized per capita GDP in period t, $F_t(x)$ is the cumulative distribution of $x_{i,t}$ across departments. Associated to it there is a probability measure $\lambda_t((-\infty, x]) = F_t(x)$, $\forall x \in \mathbb{R}$, λ_t being the probability density function for each indicator across departments in period t.

We will seek for the operator, P^*, that discloses information on how the distribution of per

[2] Formally, $K(x) = (1/\sqrt{2\pi})e^{-\frac{1}{2}x^2}$. See, for more details, Härdle and Linton (1994), Silverman (1986) and, more recently, Li and Racine (2007), among others.

[3] Increasing bandwidths for data sparsity can lead to severe bias, basically because of the kernel being based on a local constant approximation which might suffer from problems in the tails, or trimming of peaks. See Loader (1999).

[4] Additional details can be found in Loader (1999) and Loader (1996).

capita GDP at time $t-1$ transforms into a different distribution at time t. For this, we will focus on a stochastic difference equation $\lambda_t = P^*(\lambda_{t-1}, u_t)$, integer t, which takes into account that $\{u_t : \text{integer } t\}$ is the sequence of disturbances of the entire distribution. In this context, P^* is the operator mapping disturbances and probability measures into probability measures, and which "encodes" the information on intra-distribution mobility. If we assume that operator P^* is time invariant, and that the stochastic difference equation is of first order (Redding, 2002), by setting null values to disturbances and iterating for $\lambda_t = P^*(\lambda_{t-1}, u_t)$ the future evolution of the distribution can be obtained, i.e., $\lambda_{t+\tau} = (P^*)^\tau \lambda_t$.

If the set of possible values of x is discretized into a finite number of classes (grids), to which we can also refer as states or intervals, e_k, $k \in \{1, \ldots, K\}$, then P^* will become a transition probability matrix such as:

$$\lambda_{t+1} = P^* \cdot \lambda_t \tag{3}$$

Accordingly, λ_t turns into a $K \times 1$ vector of probabilities that the per capita GDP of a given department is located on a given grid at time t. It is then possible to evaluate the probability of a given department moving to a higher (or lower) position on the grid. We start by discretising the set of observations into the states e_k.[5] Each p_{kl} entry in the matrix indicates the probability that a department initially in state k will transit to state l during the period (T) under analysis.

The limits between states are chosen so that all department-year observations are uniformly distributed among the cells.[6] Accordingly, each cell in the transition probability matrices is computed by counting the number of transitions out of and into each cell. Therefore, each cell's p_{kl} value is:

$$p_{kl} = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{n_{kl}^t}{n_k^t} \tag{4}$$

where n_{kl}^t is the number of departments moving during one period from state k to class l, n_k^t is the total number of provinces starting the period in state k, and T is the length of the sample period.

3.3. Ergodic distributions, transition path analysis and mobility indices

The transition probability matrices allow characterizing the ergodic or stationary distribution under current trends. To overcome the intrinsic difficulties to transition probability matri-

[5] Once each department-year observation has been classified in one of the K states, a 5×5 matrix is built (other popular dimensions are, for instance, 7×7).

[6] Other criteria for choosing the limits between states exist, including arbitrary (albeit "reasonable") choices (Kremer et al., 2001; Quah, 1993a). An alternative to avoid the discretization issue is to consider *continuous* stochastic kernels (Quah, 1996b). They, however, are not trouble-free, particularly when trying to estimate the corresponding ergodic distributions.

ces and ergodic distributions (i.e., the need to *discretize* per capita income into five states) we consider their continuous counterparts following Johnson (2000, 2005) and considering a reasonably higher number of states (1×20).

We can use the concept of asymptotic half-life of the chain ($H - L$), which refers to the time it takes to cover half of the distance to the ergodic distribution. We define the asymptotic half-life as:

$$H - L = -\frac{\ln 2}{\ln |\lambda_2|} \tag{5}$$

where $|\lambda_2|$ is the second largest eigenvalue (after 1) of the transition probability matrix, ranging between infinity (when the stationary distribution does not exist and the second eigenvalue is equal to 1) and 0 (when $\lambda_2 = 0$ and the system has already reached its stationary equilibrium).[7]

In order to *quantify* the mobility underlying each transition matrix, we also consider mobility indices such as those considered by the economic inequality literature. Specifically, we follow Shorrocks (1978), Geweke et al. (1986) and Quah (1996a), some of whose proposals evaluate the trace of the transition probability matrix, providing information on the relative magnitude of on-diagonal and off-diagonal terms. Following Quah (1996a), its expression is:

$$\mu_1(P^*) = \frac{K - \mathrm{tr}(P^*)}{K - 1} = \frac{\sum_j (1 - p_{jj})}{K - 1} \tag{6}$$

where p_{jj} is the j-diagonal entry of matrix P^*, representing the probability of remaining in state j, and K is the number of classes. Large values of μ_1 indicate more mobility (less persistence) in P^*. This concept is identical to the inverse of the harmonic mean of expected durations of remaining in a certain state.

3.4. Conditioning schemes: demography and geography

The methods presented in the previous sections provide with a full analysis of departmental per capita income dynamics. But, as indicated by **?**, using departments as units of analysis will be less useful when the issues under analysis are "How many people in Colombia live in poverty". In this section we propose a weighting scheme for the methods presented in the preceding paragraphs.

In the specific case of the population-weighted analysis, we would not be counting transitions of departments but rather of people living in each department—i.e., the unit of analysis is the *person*. As indicated by **?**, this issue has only rarely been taken into account in convergence studies applying the distribution dynamics approach, with few exceptions such as **?**,

[7] See Magrini (1999) and, more generally, Shorrocks (1978).

Kremer et al. (2001) or Jones (1997).

Regarding the expressions corresponding to the non-parametric estimation of density functions, the modified kernel estimator becomes:

$$\hat{f}_{\omega}(x) = \frac{1}{h}\sum_{i=1}^{N} \omega_i K\Big(\frac{\|x - X_i\|_x}{h}\Big) \tag{7}$$

where, depending on the type of weighting considered, ω_i corresponds to the share of Colombian population or GDP corresponding to department i. In our local likelihood approach for density estimation, the weights can be entered directly into Equation (2).

Regarding the transition probability matrices, Equation (4) takes now into account the number of people (if we weighted by population) that moves from one class to another. In this *weighted* transition probability matrix the expression corresponding to each cell will be:

$$p^{\omega}_{kl} = \frac{1}{T-1}\sum_{t=1}^{T-1}\sum_{i=1}^{n_{kl}} \frac{W^t_{ikl}}{W^t_{ik}} \tag{8}$$

where W^t_{ikl} is the population (or GDP) corresponding to department i, that moves from state k to state l in period t, and W^t_{ik} is the population (or GDP) corresponding to department i starting the period in state k.

Regarding the role of geography, as indicated by Quah (1996b), and most of the literature on spatial econometrics, it matters. Increasing returns to scale, knowledge spillovers, access to markets, labor mobility and vertical linkages between industries explain in a large extent regional income and its geographical patterns.

We conducted an analysis which compares the *state-relative* GDP per capita used in the previous sections and *neighbor-relative* per capita GDP, where we normalize each province's per capita GDP by the average per capita GDP of the neighbor provinces, excluding the province itself. Formally, the expression corresponding to the neighbour-relative per capita GDP series is:

$$x^{NR}_i = \frac{\ln y_i}{\ln \frac{1}{NE-1}(\sum_{j\in NE\setminus i} y_j)} \tag{9}$$

where NE is the number of neighbors each i province has, and nr is the super-index indicating that we are referring to the neighbor-relative per capita GDP series. The closer the values of the neighbor-relative series are to unity, the lower the disparities among neighbor provinces are and the larger the magnitude of the spillover effects.

4. Data and descriptive statistics

Two are the variables used in the anaysis, namely GDP per capita and population. Information on both variables are provided by the National Administrative Department of Statistics (DANE). We consider the period 2000–2016. In contrast to other analyses considering previous periods, our selection enables for the consideration of all 33 Colombian departments. Data on GDP per capita is measured in constant pesos of 2005.

5. Results

We provide results for all methods described in Section 3, including transition probability matrices, ergodic (stationary) distributions, mobility indices and asymptotic half-life convergence. We also report continuous counterparts (density functions) when possible, as well as results for the different conditioning schemes—GDP-weighted, population-weighted and physically-contiguous conditioned. In the case of transition probability matrices, we present tables for the different periods and sub-periods considered (2000–2016, 2000–2008 and 2008–2016), for the unweighted analysis (Table 2), GDP-weighted (Table 3), population-weighted (Table 4), and physically-contiguous conditioned (Table 5). The last three rows in each panel display information on the initial, final and ergodic distributions of (normalized) departmental per capita income.

This analysis naturally complements and expands previous applications of the distribution dynamics approach to the case of Colombian departments, including XXX, XXX and XXX. None of these approaches... In addition, we focus on a more recent period...

5.1. Unweighted distribution dynamics

Transitions for normalized departmental per capita GDP are reported in Table 2. The top panel reports results for the entire period (2000–2016), whereas the middle and bottom panels do so for each sub-period (2000–2008 and 2008–2016, respectively). Given our period of analysis is not particularly long, we consider two-year transitions (i.e., from 2000 to 2002, from 2001 to 2003, and so on) instead of more popular choices (such as five-year transitions) in order to minimize the loss of information.

For each of the matrices in Table 2, the cut-off points (upper limits) differ slightly because the period analyzed is different. Although several criteria exist, one of the most widely accepted consists in considering all observations for the analyzed period (2000–2016, 2000–2008 or 2008–2016), and divide them into five similarly-sized intervals. As a result, the numbers in brackets to the left of each matrix correspond to the number of observations (departments) starting the period in a given state (or class). In the case of the upper panel in Table 2, given we are considering two-year transitions, they sum to 495 (instead of 528), since the last two years (2015 and 2016) would be excluded (i.e., $495 = 33 \text{departments} \times 15 \text{transitions}$).

The first row of each panel displays the cut-off points that delimit the intervals (upper limits), and should be interpreted as follows: the upper limit for the first state of 0.97 implies that approximately one fifth of the total number of observations range below 97% of the average. For the other tail of the distribution, the upper-state has observations above 1.02 (102%) of the average. Although this is a relatively narrow range of variation, note that the average is unity, since our data have been normalized by the mean, after taking logs.

Inside each 5×5 matrix in Table 2, entries (cells) should be interpreted as the probability of *remaining* in a particular state after two years—since we are considering 2-year transitions. For instance, in the case of the entire 2000–2016 period (top panel in Table 2), its value would indicate that 81% of the observations starting in the lowest relative per capita GDP state (105 observations, below 0.970) would remain in that state, whereas the remaining 19% would move to states of higher relative per capita income—in this case, to state 2. This high persistence is greater when focusing on richer departments, as shown by the probability in the lower right of the matrix, which shows that 92% of the observations in the richest state remain there after two years—with 8% moving to state 4. The rest of values on the main diagonal show less persistence. Actually, the higher probability off the main diagonal, the higher mobility, whereas values on the main diagonal closer to one indicate more persistence.

Regarding the implicit mobility that we can find in Table 2, the values on the main diagonals of each matrix average to 0.784, 0.814 and 0.774 (for 2000–2016, 2000–2008 and 2008–2016, respectively), which suggests that it is during the most recent period when more changes in the relative positions have taken place. These average values represent a good starting point to measure mobility. However, we can consider more precise measures which are less frequently used in distribution dynamics studies such as the mobility indices presented in Section 3.3.

We report results for mobility indices on Table . They do not entirely corroborate what was found for the average values in the main diagonal, since μ_1 shows quite similar values for the three periods. However, apart from the absolute value found for mobility, it is important to assess its implicit trends—i.e., whether it leads to convergence, divergence or other possible outcomes.

Specifically, the last three rows in each Table 2's panels display information on the initial (2000), final (2016) and ergodic (steady-state) distributions for the selected periods. The top panel indicates that, under 2000–2016 trends, intra-distribution mobility drives probability mass to concentrate in the states of relatively high per capita income—with 69% of probability mass concentrated in states 4 and 5, and only 20% in the poorest states (1 and 2). This converging process to richer states, however, is the result of different dynamics, as shown in the central and bottom panels in the Table, since intra-distribution mobility in the first sub-period (2000–2008) leads to probability mass to concentrate strongly (75%) in state 5. In contrast, under 2008–2016 trends (Table 2.c), although convergence still existed, it was more concentrated in poorer states—with state 2 absorbing, on the long run, 26% of probability mass. Therefore, we observe that convergence took place to a large degree before 2008, whereas the last few years have witnessed more stable patterns or, in case any tendency existed, this was actually to converge to a state closer to the average...

The values corresponding to the ergodic distribution (steady state) are valid *per se*, but can be nicely complemented providing information on how fast we can reach it. This information, rarely provided in convergence analysis studies, can be obtained via the transition path analysis or asymptotic half-life of convergence. This, as indicated by Magrini (1999), refers how long it takes to cover half the distance from the ergodic distribution and the results, corresponding to applying Equation 5, are reported in Table 7. Results might, *a priori*, seem not too intuitive, given it takes a longer period to achieve the steady-state during the period leading stronger convergence (2000–2008) compared with the second period (2008–2016), of slower convergence. However, it is precisely because the ergodic distribution in Table 6.b is more extreme than in Table 6.c why it actually takes longer to reach it.

Several authors, including Bulli (2001) and Johnson (2000, 2005), among others, have highlighted the problems of considering a discrete approach in which results partly depend on how the limits among states/classes are chosen. An alternative, which we follow here, is to consider the continuous counterpart to the transition probability matrices in Table 2. The continuous counterparts to the information reported in Table 2 are displayed in Figure 2. Specifically, Figure 2.a reports densities (estimated non-parametrically) for years 2000 (solid line), 2008 (dashed line) and 2016 (dotted line). It clearly indicates that the distribution of per capita income was bi-modal in 2000, and is still bi-modal in 2016, with the probability mass becoming more apart—i.e., the rich become richer. This would confirm that the strong convergence patterns found for pre-2000 years have almost vanished, and that the convergence occurred during our sample period is more strongly related to intra-distribution dynamics (changes in the departments relative positions, or churning).

Will this polarization persist over time? The (discrete) ergodic distributions in Table 2 do not say so, since they suggested probability mass would tend to concentrate in the richer states—regardless of the trends considered (2000–2016, 2000–2008 or 2008–2016). This result is corroborated by the continuous counterpart to the ergodic distributions in Table 2, shown in Figure 6.a, which clearly shows that bi-modality will vanish, and departments will tend to converge to levels of higher relative per capita income, since probability mass will become tighter and above unity. However, the upper tail of the distribution will still be fat, indicating that, on the long run, a number of departments will still enjoy per capita income levels well above the average.

Therefore, we have complemented the existing literature for regional convergence in Colombia in several ways, by considering a more recent period, as well as some instruments that grant much more precision to the analysis—i.e., the mobility indices, transition path analysis, and the continuous approach to the ergodic distributions. The analysis in the following subsections will enrich the study further, conditioning by several relevant factors.

5.2. Conditioning

5.2.1. Weighted analysis

Results for the GDP- and population-weighted conditioned analysis are reported in Tables 3 and 4, respectively. As for the rest of the analysis (i.e., mobility indices, transition path analysis and continuous counterparts to the probability matrices), results are presented in the same tables and figures as those corresponding to the unweighted analysis.

Regarding the discrete analysis offered by transition probability matrices in Tables 3 and 4, results differ remarkably from those obtained for the unweighted analysis. Regardless of the weighting scheme (either GDP or population), and regardless of the period considered (2000–2016, 2000–2008 or 2008–2016), the intra-distribution mobility leads to ergodic distributions with the probability mass overwhelmingly concentrated in the upper states. In several cases, for instance under 2000–2016 trends, this tendency is particularly extreme, with almost 90% of the probability mass concentrated in states 4 and 5 (Tables 3.a and 4.a). This would suggest that, in the long run, most of the population (in the case of the population-weighted analysis) would escape from poverty.

The mobility indices (Table 6) and, in particular, the transition path analysis (Table 7) complement these results, although interpretation is a bit tricky. According to the asymptotic half-life of convergence in Table 7, it would take a much longer period to reach the steady-state when conditioning either by population or GDP. However, and analogously to what occurred in the unweighted case when comparing the sub-periods, this occurs because the corresponding ergodic distributions are more extreme.

The continuous counterparts to the discrete analysis offered by transition probability matrices are reported in Figure 2.b, 2.c, as well as Figures 6.b and 6.c for ergodic distributions. Results strongly corroborate those tendencies observed when discretizing the normalized per capita income space state, as for all years 2000, 2008 and 2016 bi-modality almost disappears (particularly for GDP-weighted, see Figure 2.b). Therefore, comparing years 2000 and 2016 indicates weighted convergence (either by GDP or population) has slightly intensified, although the most prominent feature is the existence of much tighter densities, indicating that in terms of either persons or GDP, discrepancies are much less marked. The importance of weighting is even more blatant when inspecting Figures 3 and 4, which provide explicit comparisons between unweighted and weighted distributions, for 2000, 2008 and 2016. In all cases it is apparent the importance of our conditioning schemes, as densities become much tighter (indicative of more convergence) when weighting either by GDP or size, and regardless of the period considered. Finally, as indicated by the ergodic distributions in Figures 6.b and 6.c, this will ultimately result in strong convergence for people and GDP, with probability mass tightly concentrated above unity, although these (weighted) steady-state distributions will become slightly bi-modal, with a cluster of people ending up slightly richer than the rest.

5.3. Conditioning: spatial analysis

The physically contiguous-conditioned (or neighbor-relative) counterparts to the previous analyses—both conditioned and unconditioned—are reported in Table 5 (transitions and ergodic distributions), as well as in Figures 5 and 6 (densities, static and ergodic, respectively). As in the preceding sections, mobility indices and transition path analysis are also reported (Tables 6 and 7).

Analogously to what was found when comparing Table 2 to Tables 3 and 4, results differ remarkably after conditioning, although several subtleties exist that deserve discussion—and are not entirely coincidental as when weighting schemes were introduced. In this case, we observe that intra-distribution mobility differs remarkably for the two sub-periods considered, being higher in during 2008–2016 (Table 5.c)—entries in the main diagonal average to 0.67, compared to 0.76 for 2000–2008 (Table 5.b). This finding is corroborated by the mobility indices in Table 6, which also indicate that persistence is lower in the second sub-period ($\mu_1^{2008-2016} = 0.680$ and $\mu_1^{2000-2008} = 0.634$). These levels of persistence are lower compared to the state-relative series, which average to 0.81 and 0.77 for the first and second sub-periods, respectively (Table 2).

The implications of disparate mobility levels are not innocuous in terms of long-term distribution, as under 2008–2016 trends probability will be more tightly concentrated above the average, yielding an almost bi-modal ergodic distribution (Table 5.c). However, although results might be partially influenced by the choice of cut-off points,[8] the overall result is that probability mass tends to concentrate more tightly in states containing values closer to the average—i.e., spatial spillovers exist for Colombian departments.

Figure 5 report the physically-contiguous counterparts to the unweighted densities (state-relative) in Figure 2. All three graphics, corresponding to the three periods, show tighter distributions for physically-contiguous compared to state-relative per capita GDP series. Therefore, regardless of the choice of cut-off points, each department's per capita GDP is much more alike to the average of its surrounding departments than to Colombia's average. This implies that, for instance, the GDP per capita in Guaviare is much similar to the average of Meta, Vichada, Guainía, Vaupés and Caquetá than to departments in the Pacific region (Cauca, Chocó, Nariño and Valle del Cauca), thereby corroborating the existence and importance of spatial spillovers. However, the tendency is more marked during the second sub-period, as shown by a much tighter density—see the dashed line in Figure 5.c compared to Figure 5.b. Therefore, the slightly unconditional convergence process is much more accelerated when factoring in the existence of spatial interactions among neighbors. Aquí algunas explicaciones nos vendrían al pelo... William?

The continuous counterpart (**?**) to the ergodic distribution in Table **??**.a is reported in Figure 6.d. It indicates that, under 2000–2016 trends, probability will become tightly concentrated in the vicinity of 1—i.e., departments per capita GDP will be very much closer to their neighbors' average than to the nation's average.[9]

When will this physically-contiguous conditioned ergodic (stationary) distribution actually be achieved? An idea is provided by the transition path analysis (asymptotic half-life of convergence) reported in the last row of Table 7 for the three periods evaluated. Some patterns emerge here. The first one is that, under 2000–2016, the steady state corresponding to neighbor-conditioned relative GDP series would be achieved faster than under either 2000–2008 or 2008–2016 trends. The second one indicates that the speed is also faster when controlling for geographic spillovers than when these do not enter the analysis—the speed is lower (more years) for the first three rows in the table. This apparently puzzling results have a twofold explanation. On the one hand, spatial spillovers already played a role by the beginning of the period and, therefore, the ergodic distribution is not too far from the initial distribution—at least when compared two the other scenarios. On the other hand, the ergodic distributions corresponding to the physically-contiguous case are less extreme and, therefore,

[8] See **?**.

[9] The tendency is even more accentuated for 2008–2016 trends, but it is not reported in order to save space.

can be achieved (hypothetically) earlier.

These results, and especially the tendency towards the stratification of provinces in different clubs, are of no minor concern to authorities, and reveal that there is still some room for policies promoting convergence in per capita GDP among Chinese provinces, because the natural tendency towards spatial agglomeration seems to be persistent. Thus, together with the explicit regional policies and the use of other central government policies to re-balance regional development (central investment projects, endowment of infrastructures, credit policy, etc.), other measures are also needed to balance the tendency towards the localisation of economic activity induced by market forces. Improvements in the accessibility and the role of market mechanisms in the interior are needed, but increasing the role assigned to official interprovincial migrations is probably necessary too.

6. Conclusions

The hypothesis of convergence—which (in its simplest form) states that countries' long-run per capita income levels are independent from initial conditions—has been widely tested for at last thirty years now. The issue became particularly important after the emergence of modern growth theory in the mid-1980s, as testing empirically the hypothesis contributed to "unlock" the mechanics of economic growth (Johnson and Papageorgiou, 2019). This critical role of the convergence hypothesis as a test for either validating or refuting alternative growth theories attracted the interest of many reputed minds in the economics profession (Islam, 2003), ultimately leading to a vast increase in the related literature—including several surveys (Durlauf and Quah, 1999; Temple, 1999; Sala-i-Martin, 1996a; De la Fuente, 1997; Islam, 2003; Johnson and Papageorgiou, 2019).

In the informative survey by Islam (2003), and in an attempt to systematize this literature, the author proposes a classification not only of the different methodologies employed to analyze macroeconomic convergence but also the ways in which it is understood. This is particularly interesting because the first distinction he considers is convergence *within* an economy vs. convergence *across* economies, since the latter (regional convergence) has become a large area in itself (Johnson and Papageorgiou, 2019). As indicated by Jerzmanowski (2006), "growth experiences differ over time within a country almost as much as they differ among countries".

In some contexts, these regional disparities have been of particular concern. It is the case of the European Union, for a variety of reasons, including the implementation of cohesion policies, expansion and further integration initiatives, and even the challenge posed by the Brexit, giving rise to the flourishing of a large body of empirical research.[10] Regional convergence, however, has been also studied in other contexts, including several developing countries. These contexts can be even more relevant, as it has become a key global fact that the distribution of income has become more unequal in rapid-growth countries. This trend is partly shared by European regions, where it has been found that convergence exists at the country level, but regional divergences persist (Geppert and Stephan, 2008).

In this study we focus in one of these other contexts, namely, Colombia. It has one of the most dynamic and fastest-growing economies in South America, but there exists a generalized consensus as to the deficiencies in the distribution of income—including the department level. Several studies have documented this reality, finding generally either weak or absence of economic convergence (depending on the period analyzed). The lack of economic convergence in Colombia becomes a structural bottleneck to foster equal opportunities for social and economic development, while simultaneously showing the poor performance of public policies in providing relevant conditions to push regional economies towards a sustainable pattern of economic growth.

We contribute to this literature in several directions. First, our database spans from 2000 to 2016, enabling us to evaluate the most recently designed and implemented convergence-enhancing public policies. Second, we use the distribution dynamics approach which has been rarely use in the case of Colombia (with the exception of...), and complementing it by considering also mobility indices (**?**), evaluating the asymptotic half-life of convergence (**?**), and following the continuous space-state approach proposed by **?**. Third, we adapt the model to control explicitly for the role of demography and geography, introducing different weighting schemes (population and GDP) as well as comparing different spatially-conditioned GDP series.

Results are multiple and can be assessed from several points of view. The unweighted results indicate that convergence has taken place, but only until 2008. Since then, the process has stagnated. Although the ergodic distribution will be tighter (indicative of convergence), the result is driven entirely by the 2000-2008 trends. These trends, however, differ remarkably when introducing the different conditioning schemes—either demography or geography. For the population-weighted analysis, convergence exists regardless of the sub-period considered, similarly to what occurs when conditioning by GDP. In all cases not only the ergodic distributions become much tighter, but the bimodality existing in 2000, 2008 and 2016 vanishes almost entirely. When taking spatial spillovers into account, (conditional) convergence also

[10]Obviously, although some related literature exists, it is still early to evaluate the effects of the Brexit from many angles. See...

accelerates.

Therefore, our results corroborate previous findings in the literature, since the weak convergence process is corroborated. However, the shift to population-weighted comparisons has obvious implications, as the pattern changes completely, indicating that population tends to concentrate in the richest departments—pointing out to some possible weaknesses in the cohesion policies. The spatial spillovers, however, were already relevant by the beginning of the analyzed period and their importance will not vanish. Given its importance, some regions' wealth might be jeopardized by its geographical proximity to regions in conflict—particularly the Northeastern regions of the country.

References

Acevedo, S. (2003a). Convergencia y crecimiento económico en colombia 1980-2000. *Ecos de Economía: A Latin American Journal of Applied Economics*, 7(17):51–78.

Acevedo, S. (2003b). Convergencia y crecimiento económico en Colombia 1980–2000. *Ecos de Economía*, 17:51–78.

Aguirre, K. (2005). Convergencia en indicadores sociales en Colombia. Una aproximación desde los enfoques tradicionales y no paramétricos. *Desarrollo y Sociedad*, 56:147–176.

Alesina, A. and Perotti, R. (1996). Income distribution, political instability, and investment. *European economic review*, 40(6):1203–1228.

Alesina, A. and Rodrik, D. (1994). Distributive politics and economic growth. *The quarterly journal of economics*, 109(2):465–490.

Ardila-Rueda, L. (2004). Gasto público y convergencia regional en colombia. *Vol. 22. No. 45. Junio, 2004. Pág.: 222-268.*

Ben-David, D. (1994). Convergence clubs and diverging economies. Discussion Paper 922, Centre for Economic Policy Research, London.

Birchenall, J. and Murcia, G. (1997). Convergencia regional: una revisión del caso Colombiano. *Desarrollo y Sociedad*, 40:273–308.

Birchenall, J. A. (2001). Income distribution, human capital and economic growth in Colombia. *Journal of Development Economics*, 66(1):271–287.

Bonet, J. (2007). Geografía económica y análisis espacial en colombia.

Bonet, J. and Meisel, A. (1999). La convergencia regional en Colombia: una visión de largo plazo, 1926–1995. *Coyuntura Económica*, 1(29):69–106.

Bonet, J. and Meisel, A. (2008). Regional economic disparities in Colombia. *Investigaciones Regionales*, 14:61–80.

Branisa, B. and Cardozo, A. (2009a). Regional Growth Convergence in Colombia Using Social Indicators. Discussion Papers 195, IAI, University of Goettingen, Goettingen.

Branisa, B. and Cardozo, A. (2009b). Revisiting the Regional Growth Convergence Debate in Colombia Using Income Indicators. Discussion Papers 194, IAI, University of Goettingen, Goettingen.

Bulli, S. (2001). Distribution dynamics and cross-country convergence: a new approach. *Scottish Journal of Political Economy*, 48:226–243.

Cárdenas, M. and Pontón, A. (1995). Growth and convergence in Colombia: 1950–1990. *Journal of Development Economics*, 47(1):5–37.

Cárdenas, M. (1993). *Crecimiento y convergencia en Colombia: 1950–1990*. Planeación y Desarrollo, Bogotá, Edición especial DNP 35 años edition.

Cárdenas, M., Pontón, A., and Trujillo, J. (1993). Convergencia y migraciones interdepartamentales en Colombia: 1959–1989. *Coyuntura Económica*, 23(1):111–137.

De la Fuente, Á. (1997). The empirics of growth and convergence: A selective review. *Journal of Economic Dynamics and Control*, 21(1):23–73.

Durlauf, S. N. and Quah, D. T. (1999). The new empirics of economic growth. In *Handbook of Macroeconomics*, volume 1A, pages 231–304. North Holland, Amsterdam.

Galvis, L. A. and Meisel, A. (2010). Fondo de compensación regional: Igualdad de oportunidades para la periferia colombiana. *Documentos de trabajo sobre economía regional*, (122):1–43.

Galvis, L. A., Meisel, A., et al. (2001). El crecimiento económico de las ciudades colombianas y sus determinantes, 1973-1998. *Coyuntura Económica*, 31(1):69–90.

Galvis, L. A., Roca, A. M., et al. (2010). Persistencia de las desigualdades regionales en colombia: Un análisis espacial. Technical report, Banco de la Republica de Colombia.

Galvis-Aponte, L. A., Galvis-Larios, W., Hahn-de Castro, L. W., and Hahn-De-Castro, L. W. (2017). Una revisión de los estudios de convergencia regional en colombia. *Documentos de Trabajo Sobre Economía Regional y Urbana; No. 264*.

Galvis-Aponte, L. A. and Wilfried Hahn-De-Castro, L. (2016). Crecimiento municipal en colombia: el papel de las externalidades espaciales, el capital humano y el capital físico. *sociedad y economía*, (31):149–174.

García, R. R. and Benitez, A. V. (1998). Crecimiento regional en colombia:¿ persiste la desigualdad? *Revista de economía del Rosario*, 1(1):67–108.

Geppert, K. and Stephan, A. (2008). Regional disparities in the European Union: Convergence and agglomeration. *Papers in Regional Science*, 87(2):193–217.

Geweke, J., Marshall, R. C., and Zarkin, G. A. (1986). Mobility indices in continuous time Markov chains. *Econometrica*, 54(6):1407–1423.

Giannetti, M. (2002). The effects of integration on regional disparities: Convergence, divergence or both? *European Economic Review*, 46:539–567.

Herrerías, M. J., Orts, V., and Tortosa-Ausina, E. (2011). Weighted convergence and regional clusters across China. *Papers in Regional Science*, 90(4):703–734.

Hjort, N. L. and Jones, M. C. (1996). Locally parametric nonparametric density estimation. *The Annals of Statistics*, 24(4):1619–1647.

Härdle, W. and Linton, O. (1994). Applied nonparametric methods. In Engle, R. and McFadden, D., editors, *Handbook of Econometrics*, volume 4. North Holland, Amsterdam.

Islam, N. (2003). What have we learnt from the convergence debate? *Journal of Economic Surveys*, 17(3):309–362.

Jerzmanowski, M. (2006). Empirics of hills, plateaus, mountains and plains: A Markov-switching approach to growth. *Journal of Development Economics*, 81(2):357–385.

Johnson, P. and Papageorgiou, C. (2019). What remains of cross-country convergence? *Journal of Economic Literature*, forthcoming.

Johnson, P. A. (2000). A nonparametric analysis of income convergence across the US states. *Economics Letters*, 69:219–223.

Johnson, P. A. (2005). A continuous state space approach to "Convergence by parts". *Economics Letters*, 86:317–321.

Jones, C. I. (1997). On the evolution of the world income distribution. *Journal of Economic Perspectives*, 11(3):19–36.

Kremer, M., Onatski, A., and Stock, J. (2001). Searching for prosperity. *Carnegie-Rochester Conference Series on Public Policy*, 55:275–303.

Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, Princeton and Oxford.

Loader, C. R. (1996). Local likelihood density estimation. *The Annals of Statistics*, 24(4):1602–1618.

Loader, C. R. (1999). *Local Regression and Likelihood*. Springer Verlag, New York.

Magrini, S. (1999). The evolution of income disparities among the regions of the European Union. *Regional Science and Urban Economics*, 29(2):257–281.

Martínez, A. (2006). Determinantes del PIB per cápita de los Departamentos Colombianos 1975–2003. Archivos de Economía 318, DNP, Bogotá.

Meisel, A. (1993). Polarización o convergencia? A propósito de Cárdenas, Pontón y Trujillo. *Coyuntura Económica*, 23(2):153–161.

Quah, D. T. (1993a). Empirical cross-section dynamics in economic growth. *European Economic Review*, 37:426–434.

Quah, D. T. (1993b). Galton's fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics*, 95(4):427–443.

Quah, D. T. (1996a). Aggregate and regional disaggregate fluctuations. *Empirical Economics*, 21:137–159.

Quah, D. T. (1996b). Regional convergence clusters across Europe. *European Economic Review*, 40:951–958.

Redding, S. (2002). Specialization dynamics. *Journal of International Economics*, 58(2):299–334.

Rocha, R. and Vivas, A. (1998). Crecimiento regional en Colombia: ¿persiste la desigualdad? *Revista de Economía del Rosario*, 1:67–108.

Royuela, V. and García, G. A. (2015). Economic and social convergence in Colombia. *Regional Studies*, 49(2):219–239.

Sala-i-Martin, X. (1996a). The classical approach to convergence analysis. *The Economic Journal*, 106(437):1019–1036.

Sala-i-Martin, X. (1996b). Regional cohesion: Evidence and theories of regional growth and convergence. *European Economic Review*, 40:1325–1352.

Sala-i-Martin, X. (2006). The world distribution of income: Falling poverty and... convergence, period. *Quarterly Journal of Economics*, 121(2):351–398.

Schultz, T. P. (1998). Inequality in the distribution of personal income in the world: How it is changing and why. *Journal of Population Economics*, 11(3):307–344.

Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica*, 46(5):1013–1024.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

Temple, J. (1999). The new growth evidence. *Journal of Economic Literature*, pages 112–156.

Tortosa-Ausina, E., Pérez, F., Mas, M., and Goerlich, F. J. (2005). Growth and convergence profiles in Spanish provinces (1965–1997). *Journal of Regional Science*, 45:147–182.

Table 1: Descriptive statistics

	Income per capita (log)			Income per capita growth (%)			Population (millions)			Population growth (%)		
Region/province	2000	2008	2016	2000–2008	2009-16	2000-16	2000	2008	2016	2000–2008	2009-16	2000-16
Andina												
Antioquia	5.40	10.80	18.30	9.10	6.80	7.90	5.20	5.90	6.50	1.30	1.20	1.30
Bogotá	8.70	17.20	27.50	8.80	6.10	7.40	6.30	7.10	7.90	1.50	1.30	1.40
Boyacá	4.60	10.60	20.10	11.10	8.30	9.70	1.20	1.20	1.20	0.20	0.10	0.20
Caldas	3.60	8.20	13.10	10.60	6.10	8.40	0.90	0.90	0.90	0.10	0.10	0.10
Cundinamarca	5.20	10.20	16.70	8.60	6.40	7.50	2.00	2.30	2.70	1.80	1.50	1.70
Huila	4.10	8.20	13.00	9.10	6.10	7.60	0.90	1.00	1.10	1.40	1.20	1.30
Norte de Santander	3.00	6.50	10.70	9.90	6.50	8.20	1.10	1.20	1.30	0.80	0.80	0.80
Quindío	4.00	6.70	12.00	6.80	7.60	7.20	0.50	0.50	0.50	0.50	0.50	0.50
Risaralda	3.80	8.00	13.70	9.80	6.90	8.40	0.80	0.90	0.90	0.60	0.50	0.60
Santander	6.20	17.60	31.10	13.90	7.60	10.70	1.90	1.90	2.00	0.50	0.50	0.50
Tolima	3.70	8.20	13.10	10.50	6.00	8.20	1.30	1.30	1.40	0.30	0.20	0.30
Mean	4.80	10.20	17.20	9.80	6.70	8.30	2.00	2.20	2.40	0.80	0.70	0.80
Standard Deviation	1.50	3.60	6.30	1.70	0.70	1.00	1.80	2.00	2.30	0.50	0.40	0.50
Caribe												
Atlántico	4.40	8.50	14.60	8.70	6.90	7.80	2.00	2.20	2.40	1.40	1.20	1.30
Bolivar	3.90	9.80	17.70	12.30	7.10	9.70	1.70	1.90	2.10	0.90	1.10	1.00
Cesar	3.30	10.00	15.30	14.70	5.60	10.10	0.80	0.90	1.00	1.30	1.20	1.30
Córdoba	3.00	5.80	8.80	9.10	5.30	7.20	1.30	1.50	1.70	1.50	1.50	1.50
La Guajira	3.50	8.10	7.90	11.50	1.10	6.30	0.50	0.70	0.90	4.20	3.20	3.70
Magdalena	2.40	5.30	8.70	10.20	6.30	8.20	1.10	1.10	1.20	0.60	0.90	0.80
San Andrés	4.90	9.80	16.90	9.20	7.60	8.10	0.06	0.07	0.07	0.80	0.80	0.80
Sucre	2.20	4.60	8.10	9.50	7.20	8.40	0.70	0.70	0.80	0.90	0.90	0.90
Mean	3.50	7.70	12.10	10.60	5.90	8.20	1.00	1.10	1.30	1.40	1.40	1.40
Standard Deviation	0.80	2.00	3.80	1.90	2.00	1.20	0.60	0.60	0.70	1.00	0.70	0.80
Pacífica												
Chocó	1.60	3.60	7.30	10.60	10.00	10.30	0.40	0.40	0.50	0.80	0.90	0.90
Valle del Cauca	5.70	11.20	17.90	8.70	6.10	7.40	3.90	4.20	4.60	1.00	1.00	1.00
Cauca	2.30	5.20	11.00	10.90	9.60	10.30	1.20	1.20	1.30	0.80	0.80	0.80
Nariño	2.10	4.50	7.80	9.60	7.30	8.40	1.40	1.50	1.70	1.20	1.20	1.20
Mean	2.90	6.10	11.00	9.90	8.20	9.10	1.70	1.90	2.00	0.90	1.00	1.00
Standard Deviation	1.60	2.90	4.20	0.90	1.60	1.20	1.30	1.40	1.50	0.10	0.10	0.10
Orinoquía												
Meta	5.80	17.40	26.20	15.30	7.60	11.50	0.60	0.80	0.90	2.20	2.00	2.10
Vichada	2.70	5.10	6.20	8.20	2.60	5.40	0.04	0.06	0.07	2.60	2.40	2.50
Casanare	25.70	28.30	30.10	2.10	2.20	2.10	0.20	0.30	0.30	2.10	1.80	2.00
Arauca	8.70	23.10	15.90	15.90	-4.10	5.90	0.20	0.20	0.20	1.40	1.10	1.20
Mean	10.70	18.50	19.60	10.40	2.10	6.20	0.30	0.30	0.40	2.10	1.80	2.00
Standard Deviation	8.80	8.60	9.30	5.70	4.20	3.40	0.20	0.20	0.30	0.40	0.40	0.40
Amazonia												
Amazonas	2.60	4.70	7.90	7.60	6.70	7.20	0.06	0.07	0.07	1.10	1.10	1.30
Caquetá	2.30	4.70	8.50	9.00	7.80	8.40	0.30	0.40	0.40	1.10	1.20	1.20
Guainía	2.60	4.20	6.60	6.90	6.00	6.40	0.03	0.03	0.04	2.00	1.60	1.80
Guaviare	2.70	4.40	6.40	6.80	5.50	6.10	0.08	0.10	0.10	1.40	1.40	1.40
Putumayo	2.60	5.80	8.30	11.30	6.00	8.60	0.20	0.30	0.30	1.00	1.10	1.00
Vaupés	1.90	2.90	5.20	5.50	7.30	6.40	0.03	0.04	0.04	1.40	1.00	1.20
Mean	2.40	4.40	7.20	7.90	6.50	7.20	0.10	0.10	0.10	1.40	1.20	1.30
Standard Deviation	0.20	0.80	1.10	1.90	0.80	1.00	0.10	0.10	0.10	0.30	0.20	0.20
Full sample												
Mean	4.50	9.10	13.70	9.70	6.10	7.90	1.20	1.30	1.40	9.70	6.10	7.90
Standard Deviation	4.10	5.60	6.80	2.20	2.70	3.10	1.40	1.50	1.70	2.20	2.70	3.10

Table 2: Transition probability matrix and ergodic distribution, per capita income (GDP/N), unweighted, 2-year transitions, limits all years

	Upper limit, all years:				
(Number of observations)	0.970	0.988	1.005	1.023	Max.
(105)	0.81	0.19	0.00	0.00	0.00
(92)	0.20	0.68	0.12	0.00	0.00
(103)	0.00	0.10	0.73	0.17	0.00
(99)	0.00	0.00	0.14	0.78	0.08
(96)	0.00	0.00	0.00	0.08	0.92
Initial distribution (2000)	0.18	0.24	0.21	0.15	0.21
Final distribution (2016)	0.15	0.24	0.18	0.21	0.21
Ergodic distribution	0.10	0.10	0.12	0.28	0.41

a) 2000–2016

	Upper limit, all years:				
(Number of observations)	0.970	0.989	1.006	1.024	Max.
(45)	0.89	0.11	0.00	0.00	0.00
(47)	0.18	0.72	0.09	0.00	0.00
(47)	0.00	0.06	0.74	0.20	0.00
(44)	0.00	0.00	0.14	0.78	0.07
(48)	0.00	0.00	0.00	0.06	0.94
Initial distribution (2000)	0.18	0.24	0.21	0.18	0.18
Final distribution (2008)	0.24	0.18	0.18	0.24	0.15
Ergodic distribution	0.04	0.02	0.05	0.14	0.75

b) 2000–2008

	Upper limit, all years:				
(Number of observations)	0.970	0.988	1.004	1.020	Max.
(50)	0.80	0.20	0.00	0.00	0.00
(43)	0.14	0.74	0.12	0.00	0.00
(47)	0.00	0.09	0.74	0.17	0.00
(46)	0.00	0.05	0.10	0.71	0.14
(45)	0.00	0.00	0.00	0.12	0.88
Initial distribution (2008)	0.15	0.24	0.15	0.24	0.21
Final distribution (2016)	0.09	0.33	0.27	0.18	0.12
Ergodic distribution	0.19	0.26	0.18	0.20	0.16

c) 2008–2016

Notes: The variable of analysis is $x_{it} = \ln y_{it}/\ln \bar{y}_t$, where y_{it} is the per capita GDP of the province (in constant 1952 prices). The 5-year (or quinquennial) transition refers to the movement of x_{it} from one of the five states in period t to another (including staying in the same) state in period $t+5$. The transition matrices presented in this Table are estimated by averaging the observed 5-year transitions of *provinces* during the periods 1952–2005 (top panel), 1952–1978 (middle panel), and 1978–2005 (bottom panel). The transition matrices and ergodic distributions displayed in each panel are based on five states, whose upper limits (the "grid") are chosen to yield a virtually uniform distribution over the observed sample. In order to facilitate comparisons, these cut-off points were calculated using the entire 1952–2005 sample (totalling 28 provinces $\times$ 54 years $=$ 1512 observations), i.e., the top panel, and remained unchanged throughout the entire analysis. The numbers in parentheses on the left are the numbers of observations beginning from a particular state. The cells are arranged in ascending order, with the upper left cell in each matrix showing transitions from the poorest to the poorest. The way the variable of analysis x_{it} is computed allows an economically meaningful interpretation of each state to be made, i.e., observations in state one are those with GDP per capita lower than the 91.5% of the national average, as indicated by its cut-off point. The ergodic distributions are computed following Kremer et al. (2001).

Table 3: Transition probability matrix and ergodic distribution, per capita income (GDP/N), GDP-weighted, 2-year transitions, limits all years

	Upper limit, all years:				
(Share of GDP)	0.970	0.988	1.005	1.023	Max.
(0.03)	0.81	0.19	0.00	0.00	0.00
(0.06)	0.16	0.73	0.11	0.00	0.00
(0.12)	0.00	0.09	0.78	0.13	0.00
(0.28)	0.00	0.00	0.09	0.83	0.08
(0.51)	0.00	0.00	0.00	0.06	0.94
Initial distribution (2000)	0.04	0.06	0.10	0.15	0.50
Final distribution (2016)	0.01	0.07	0.08	0.19	0.52
Ergodic distribution	0.01	0.04	0.07	0.29	0.59

a) 2000–2016

	Upper limit, all years:				
(Share of GDP)	0.970	0.989	1.006	1.024	Max.
(0.03)	0.86	0.14	0.00	0.00	0.00
(0.06)	0.12	0.78	0.10	0.00	0.00
(0.12)	0.00	0.07	0.77	0.17	0.00
(0.20)	0.00	0.00	0.09	0.77	0.13
(0.59)	0.00	0.00	0.00	0.05	0.95
Initial distribution (2000)	0.04	0.06	0.10	0.29	0.36
Final distribution (2008)	0.03	0.07	0.07	0.33	0.36
Ergodic distribution	0.01	0.03	0.11	0.24	0.62

b) 2000–2008

	Upper limit, all years:				
(Share of GDP)	0.970	0.988	1.004	1.020	Max.
(0.03)	0.74	0.26	0.00	0.00	0.00
(0.08)	0.13	0.79	0.07	0.00	0.00
(0.11)	0.00	0.11	0.75	0.14	0.00
(0.29)	0.00	0.01	0.12	0.74	0.14
(0.49)	0.00	0.00	0.00	0.13	0.87
Initial distribution (2008)	0.03	0.05	0.09	0.19	0.50
Final distribution (2016)	0.01	0.07	0.06	0.21	0.52
Ergodic distribution	0.02	0.10	0.11	0.26	0.52

c) 2008–2016

Notes: Table **??**'s notes also apply here with the exception that the transition matrices are estimated by averaging the observed 5-year transitions of *GDP* (i.e., the GDP of each province that moves from one state to another) during the periods 1952–2005 (top panel), 1952–1978 (middle panel), and 1978–2005 (bottom panel). Therefore, the numbers in parentheses on the left are the percentage of GDP beginning from a particular state; these percentages were calculated taking into account the GDP of each province beginning from a particular state, and the sum of the numbers in parentheses in Table **??**.a represents 100%.

Table 4: Transition probability matrix and ergodic distribution, per capita income (GDP/N), population-weighted, 2-year transitions, limits all years

(Share of population)	Upper limit, all years: 0.970	0.989	1.006	1.024	Max.
(0.07)	0.81	0.19	0.00	0.00	0.00
(0.12)	0.16	0.73	0.11	0.00	0.00
(0.16)	0.00	0.09	0.77	0.13	0.00
(0.28)	0.00	0.00	0.09	0.83	0.08
(0.37)	0.00	0.00	0.00	0.06	0.94
Initial distribution (2000)	0.09	0.11	0.14	0.16	0.38
Final distribution (2016)	0.01	0.14	0.12	0.21	0.41
Ergodic distribution	0.02	0.05	0.08	0.31	0.53

a) 2000–2016

(Share of population)	Upper limit, all years: 0.970	0.989	1.006	1.024	Max.
(0.07)	0.86	0.14	0.00	0.00	0.00
(0.10)	0.12	0.78	0.10	0.00	0.00
(0.16)	0.00	0.07	0.77	0.16	0.00
(0.21)	0.00	0.00	0.10	0.77	0.13
(0.45)	0.00	0.00	0.00	0.05	0.95
Initial distribution (2000)	0.09	0.11	0.14	0.30	0.24
Final distribution (2008)	0.06	0.13	0.10	0.35	0.25
Ergodic distribution	0.01	0.03	0.13	0.26	0.57

b) 2000–2008

(Share of population)	Upper limit, all years: 0.970	0.988	1.004	1.020	Max.
(0.06)	0.74	0.26	0.00	0.00	0.00
(0.14)	0.14	0.79	0.07	0.00	0.00
(0.15)	0.00	0.11	0.74	0.15	0.00
(0.30)	0.00	0.01	0.12	0.74	0.14
(0.35)	0.00	0.00	0.00	0.13	0.87
Initial distribution (2008)	0.06	0.10	0.13	0.21	0.38
Final distribution (2016)	0.01	0.14	0.10	0.23	0.41
Ergodic distribution	0.03	0.12	0.12	0.27	0.46

c) 2008–2016

Notes: Table **??**'s notes also apply here with the exception that the transition matrices are estimated by averaging the observed 5-year transitions of *people* (i.e., the population of each province that moves from one state to another) during the periods 1952–2005 (top panel), 1952–1978 (middle panel), and 1978–2005 (bottom panel). Therefore, the numbers in parentheses on the left are the percentage of population beginning from a particular state; these percentages were calculated taking into account the population of each province beginning from a particular state, and the sum of the numbers in parentheses in Table **??**.a represents 100%.

Table 5: Transition probability matrix and ergodic distribution, per capita income (GDP/N), physically contiguous-conditioned, 2-year transitions, limits all years

	Upper limit, all years:				
(Number of observations)	0.977	0.988	0.999	1.008	Max.
(101)	0.80	0.18	0.01	0.01	0.00
(101)	0.18	0.69	0.13	0.00	0.00
(95)	0.00	0.14	0.65	0.19	0.02
(98)	0.00	0.00	0.17	0.73	0.10
(100)	0.00	0.00	0.01	0.13	0.86
Initial distribution (2000)	0.21	0.18	0.24	0.15	0.21
Final distribution (2016)	0.18	0.15	0.24	0.27	0.15
Ergodic distribution	0.16	0.18	0.20	0.26	0.20

a) 2000–2016

	Upper limit, all years:				
(Number of observations)	0.979	0.989	0.999	1.010	Max.
(46)	0.88	0.10	0.00	0.02	0.00
(46)	0.13	0.73	0.14	0.00	0.00
(50)	0.00	0.12	0.61	0.22	0.05
(43)	0.00	0.00	0.15	0.73	0.11
(46)	0.00	0.00	0.00	0.15	0.85
Initial distribution (2000)	0.24	0.18	0.21	0.18	0.18
Final distribution (2008)	0.24	0.18	0.09	0.27	0.21
Ergodic distribution	0.23	0.11	0.14	0.33	0.18

b) 2000–2008

	Upper limit, all years:				
(Number of observations)	0.977	0.987	0.999	1.007	Max.
(73)	0.83	0.17	0.00	0.00	0.00
(18)	0.29	0.38	0.28	0.06	0.00
(23)	0.00	0.17	0.49	0.34	0.00
(32)	0.00	0.06	0.18	0.69	0.06
(85)	0.00	0.00	0.02	0.00	0.98
Initial distribution (2008)	0.18	0.15	0.24	0.24	0.18
Final distribution (2016)	0.18	0.24	0.09	0.24	0.24
Ergodic distribution	0.21	0.09	0.11	0.09	0.50

c) 2008–2016

Notes: Table **??**'s notes also apply here with the exception that the variable of analysis is the neighbour-relative GDP per capita series of province i in period t, x_{it}^{NR}, as defined in Equation (9). The 5-year (or quinquennial) transition refers to the movement of x_{it}^{NR} from one of the five states in period t to another (including staying in the same) state in period $t+5$. Therefore, the transition matrices presented in this Table are estimated by averaging the observed 5-year transitions of *provinces* during the periods of 1952–2005 (top panel), 1952–1978 (middle panel), and 1978–2005 (bottom panel).

Table 6: Mobility indices (μ_1)[a], 2-year transitions

Transition matrix	2000–2008	2008–2016	2000–2016
Unweighted	0.624	0.624	0.623
GDP-weighted	0.569	0.605	0.548
Population-weighted	0.574	0.606	0.549
Physically contiguous-conditioned	0.634	0.680	0.631

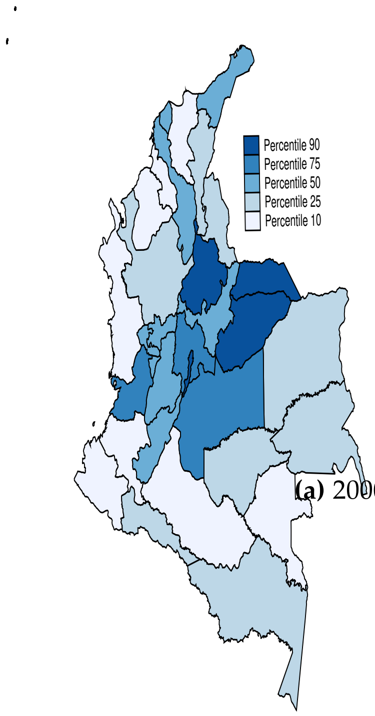
[a] The values refer to the μ_1 index, as defined in Equation (6), which summarises the mobility information in each transition probability matrix in one number so as to facilitate comparisons across them.

Table 7: Transition path analysis (asymptotic half-life of convergence, $H - L$)[a], 2-year transitions

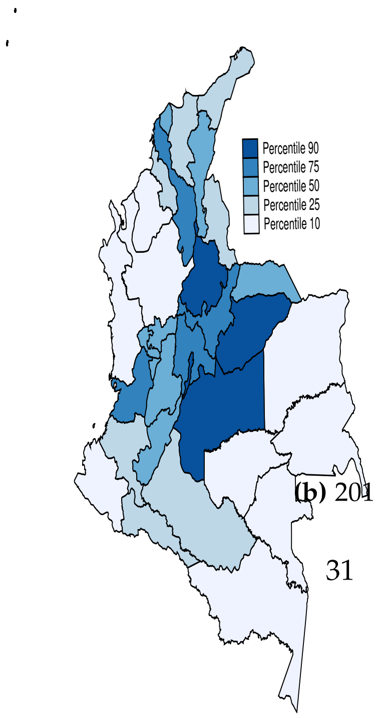
Transition matrix	2000–2008	2008–2016	2000–2016
Unweighted	72.024	31.588	47.874
GDP-weighted	138.502	65.334	45.343
Population-weighted	158.355	65.998	46.854
Physically contiguous-conditioned	24.629	24.068	16.501

[a] The values indicate the speed at which the ergodic or steady-state distribution is approached. Specifically, they refer to the concept of the asymptotic half-life of the chain, $H - L$, which is how long it takes to cover half the distance from the stationary distribution. Since we are using 2-year transitions, these numbers should be multiplied by 2 in order to have them in years.

Figure 1: GDP per capita

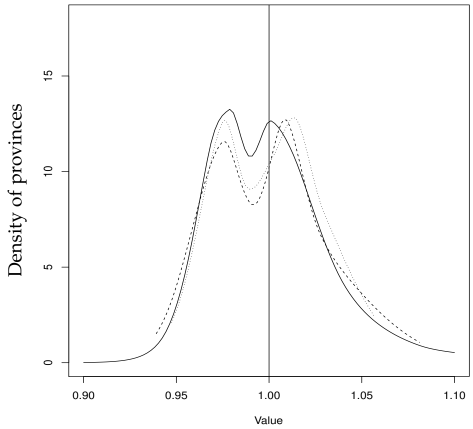


(a) 2000

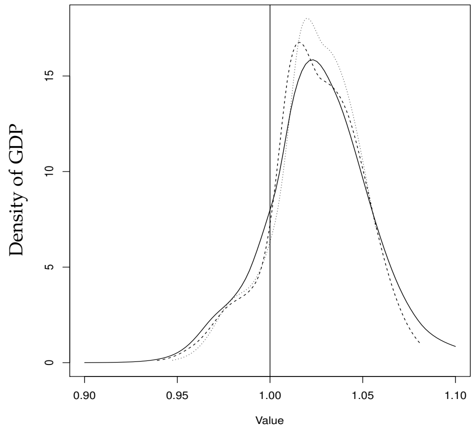


(b) 2016

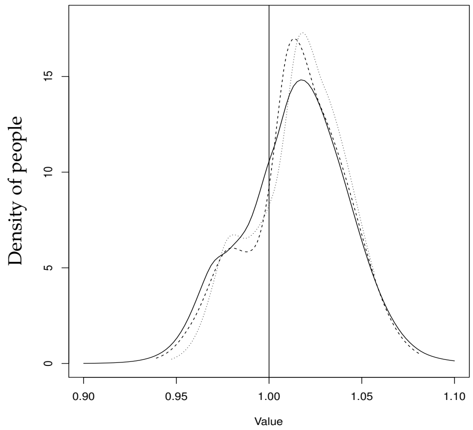
Figure 2: GDP/N, densities, 2000 vs. 2008 vs. 2016



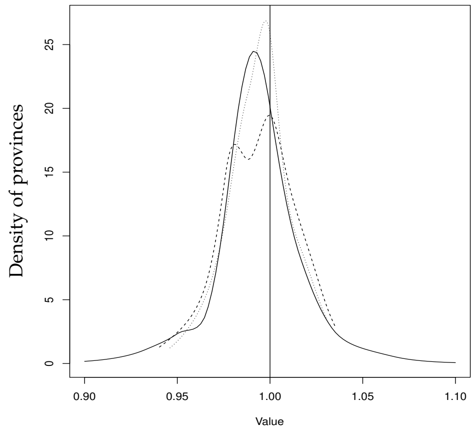
(a) Unweighted



(b) GDP-weighted



(c) Population-weighted

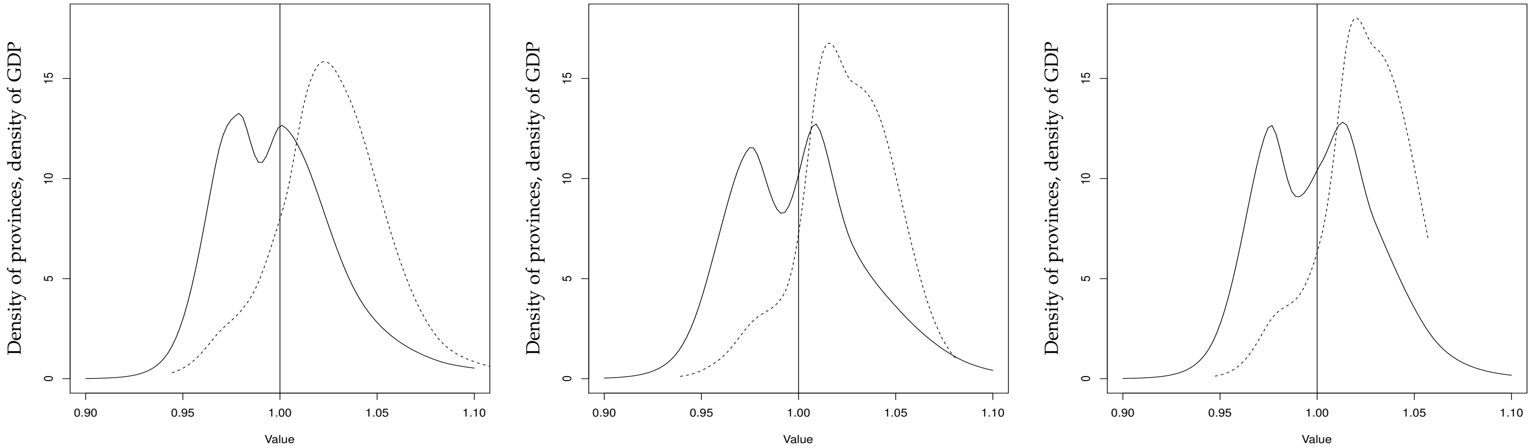


(d) Physically contiguous-conditioned

2000 —— 2008 ------ 2016 ï¿½ï¿½ï¿½ï¿½ï¿½ï¿½ï¿½ï¿½

Notes: All figures contain densities estimated using local likelihood density estimation. The vertical line represents the average, which is the unity because we have normalised the variable of analysis, $x_{it} = \ln y_{it} / \ln \bar{y}_t$, where y_{it} is the per capita GDP of the province (in constant 1952 prices).

Figure 3: GDP/N, densities, unweighted vs. GDP-weighted

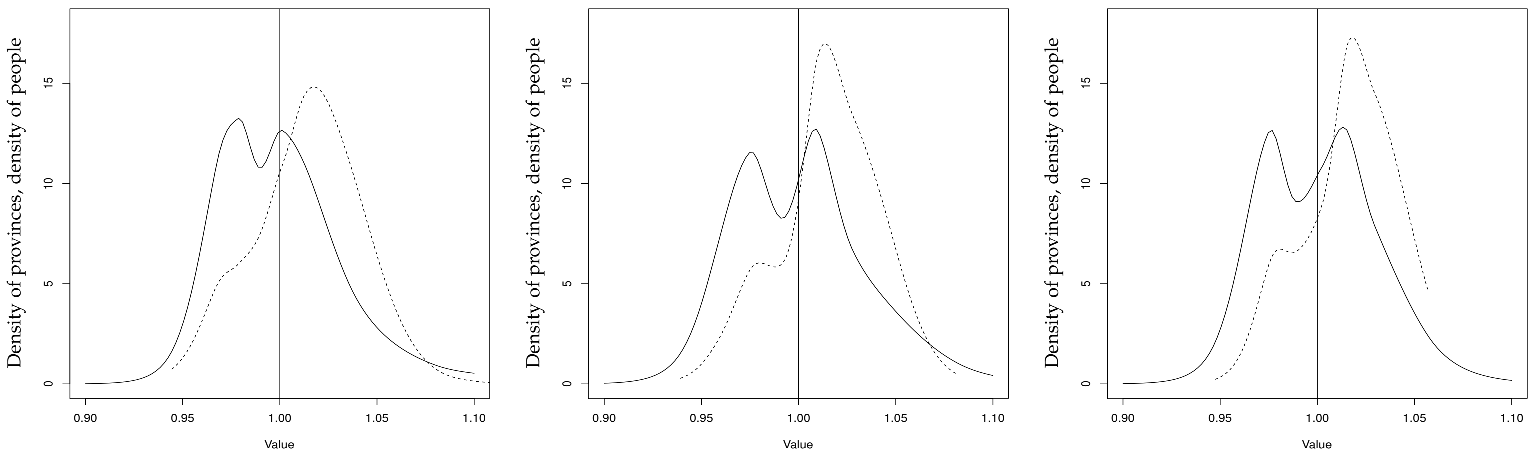


(a) 2000 **(b)** 2008 **(c)** 2016

Unweighted —— GDP-weighted ------

Notes: All figures contain densities estimated using local likelihood density estimation for the years 2000, 2008 and 2016. The vertical line represents the average, which is unity because we have normalised the variable of analysis, $x_{it} = \ln y_{it} / \ln \bar{y}_t$, where y_{it} is the per capita GDP of the province (in constant 2000 prices). The solid line is the unweighted density of the x_{it}, whereas the dashed line refers to the GDP-weighted density.

Figure 4: GDP/N, densities, unweighted vs. population-weighted

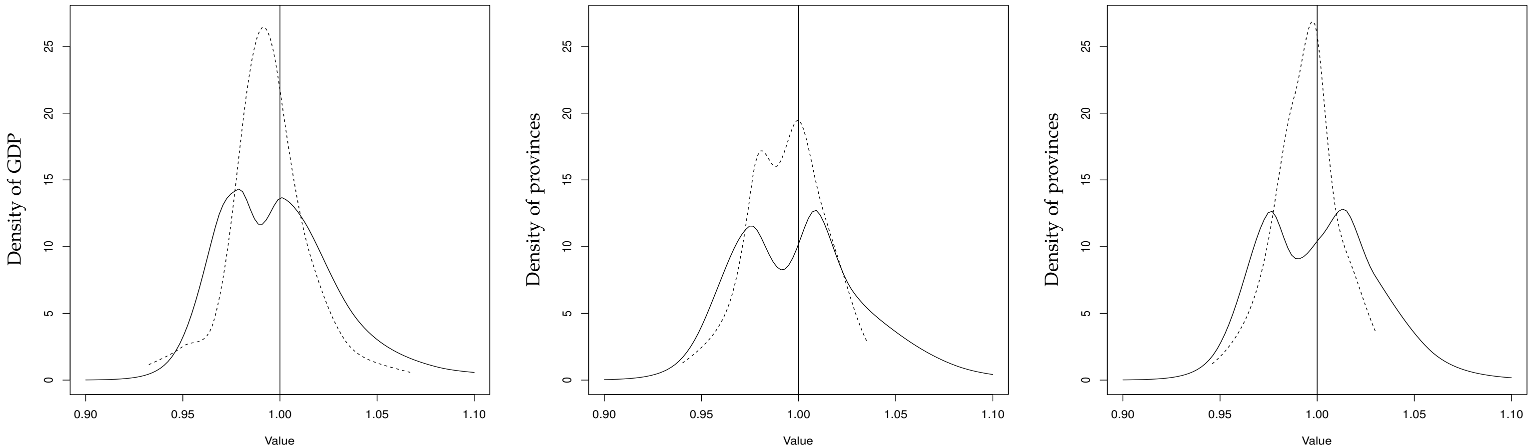


(a) 2000 **(b)** 2008 **(c)** 2016

Unweighted —— Population-weighted ------

Notes: All figures contain densities estimated using local likelihood density estimation for the years 1952, 1978 and 2005. The vertical line represents the average, which is unity because we have normalised the variable of analysis, $x_{it} = \ln y_{it}/\ln \bar{y}_t$, where y_{it} is the per capita GDP of the province (in constant 1952 prices). The solid line is the unweighted density of the x_{it}, whereas the dashed line refers to the population-weighted density.

Figure 5: GDP/N, densities, unweighted vs. physically contiguous-conditioned

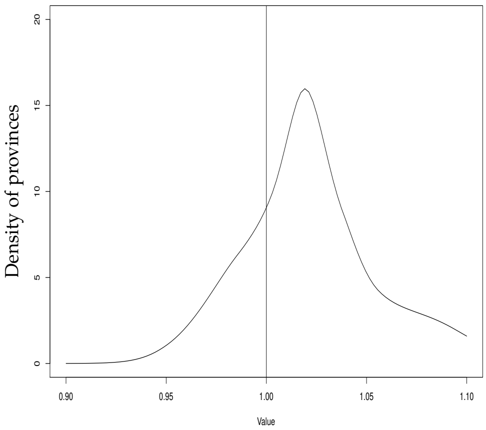


(a) 2000 **(b)** 2008 **(c)** 2016

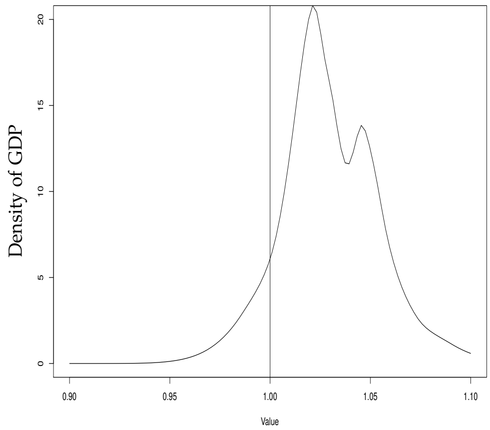
Unweighted —— Physically contiguous-conditioned ------

Notes: All figures contain densities estimated using local likelihood density estimation for the years 1952, 1978 and 2005. The vertical line represents the average, which is unity because we have normalised the variable of analysis, $x_{it} = \ln y_{it}/\ln \bar{y}_t$, where y_{it} is the per capita GDP of the province (in constant 1952 prices). The solid line is the unweighted density of the x_{it}, whereas the dashed line refers to the neighbour-relative GDP per capita series of province i in period t, x_{it}^{NR}, as defined in Equation (9).

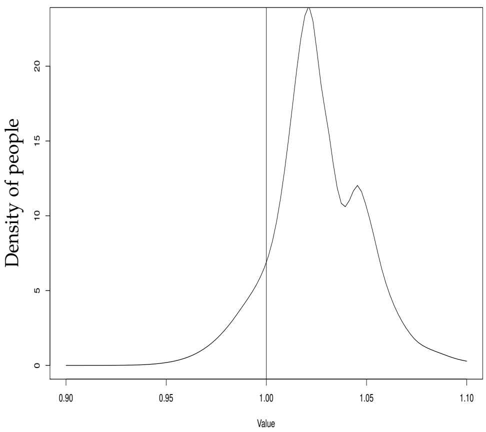
Figure 6: GDP/N, ergodic distributions, 2000–2016, 2-year transitions



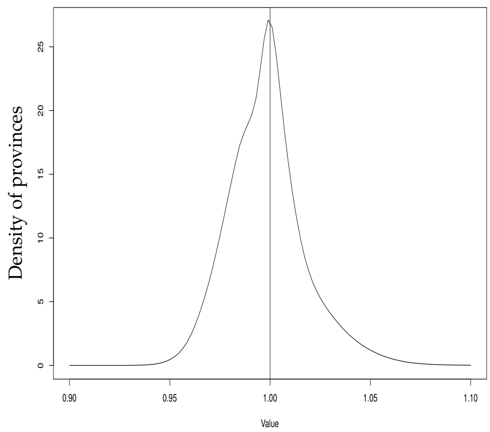
(a) Unweighted



(b) GDP-weighted



(c) Population-weighted



(d) Physically contiguous-conditioned

Notes: All figures contain ergodic densities estimated using local likelihood density estimation. The vertical line represents the average, which is unity because we have normalised the variable of analysis, $x_{it} = \ln y_{it} / \ln \bar{y}_t$, where y_{it} is the per capita GDP of the province (in constant 1952 prices). The solid line in each subfigure represents the ergodic densities under 1952–1978 trends, whereas the dashed lines are the ergodic densities under 1978–2005 trends. The scale of the vertical axes is not displayed in full in order to facilitate comparison of the densities.