Insper | IBMEC São Paulo

Income Convergence Clubs for Brazilian Municipalities: A Non-Parametric Analysis

Márcio Laurini

Eduardo Andrade

Pedro L. Valls Pereira

Insper Working Paper

WPE: 050/2004

Income Convergence Clubs for Brazilian Municipalities: a Non-Parametric Analysis[1]

Márcio Laurini Ibmec/SP
Eduardo Andrade Ibmec/SP
Pedro L. Valls Pereira Ibmec/SP

Abstract

In this article we analyze the evolution of relative per capita income distribution of Brazilian municipalities over the period 1970-1996. Our analyses are based on non-parametric methodologies and do not assume probability distributions or functional forms for the data. We have carried out two convergence tests - a test for Sigma convergence based on the Bootstrap principle and a Beta convergence test using Smoothing Splines for the growth regressions. The results obtained demonstrate the need to model the dynamics of income for Brazilian municipalities as a process of convergence clubs, using the methodology of transition matrices and stochastic kernels. The results show the formation of two convergence clubs, a low income club formed by the municipalities of the North and Northeast regions, and another high income club formed by the municipalities of the Center-West, Southeast and South regions. The formation of convergence clubs is confirmed by a bootstrap test for multimodality.

Key Words: Convergence Clubs, Non-Parametric Methods, Distribution Dynamics.

[1]Ibmec Business School, Rua Maestro Cardim 1170, 11o. andar, 01323-001, São Paulo, Brasil; Tel: (55)(11)3175-2311; Fax: (55)(11)3175-2315; E-mail: (marciopl@ibmec.br). The authors wish to thank Patrícia Palomo and Samuel Kinoshita for their excellent academic assistance. The last author wishes to acknowledge the partial financial assistance of CNPq and PRONEX.

1 Introduction

The hypothesis of per capita income convergence may be summarized as a progressively diminishing trend over time in the differences in relative incomes between rich and poor economies. Convergence is one of the principal predictions of the neoclassical growth model proposed by Solow (1956) and Swan (1956), being a consequence of the assumption of diminishing returns for factors of production. This implies that the productivity of capital is greater in relatively poorer economies, leading to a higher rate of growth in economies with a lower capital stock, and to income convergence in the long-run. Due to a greater homogeneity in technological and behavioral parameters, caused by the absence of barriers to the mobility of capital and labor within a single country, the convergence between the incomes of municipalities within a single country would be even more likely.

Traditionally, tests for convergence and income distribution modeling are based on the assumption that the distribution of data is known, for example that data follows a normal distribution, while in tests of Beta convergence it is assumed that the relation between the growth rate and the logarithm of initial income is linear. Our analysis shows that the assumption of linearity in the growth regression may hide divergent relationships for some relative income bands.

The convergence tests based on cross-section regressions, such as the use of growth regressions that express the growth rate as a function of initial income, have been criticized by Quah (1993) on the grounds that modeling a conditional mean may be inadequate for analyzing the hypothesis of convergence. The first problem with this regression is the assumption that the estimated coefficient is the same for all economies. The second problem is known as 'Galton's Fallacy', as pointed out by Friedman (1992) and Quah (1993), who show that the negative coefficient encountered in growth regressions may be a symptom of regression to the mean rather than implying convergence.

Relaxing these assumptions of linearity and a given distribution, we test for convergence and model the dynamics of relative income for Brazilian municipalities using non-parametric methods. We have carried out Sigma convergence tests using the traditional statistics of the Coefficient of Variation and the Theil index, which measure the dispersion between incomes, obtaining the distributions of these estimators using Bootstrap methods. The Beta convergence test that uses the non-parametric Smoothing Spline estimator, relaxes the linearity imposed by estimation using ordinary least squares, and we derive a convergence test based on the first derivative of this estimator. This test shows that the hypothesis of convergence, represented by a negative relationship between the growth rate and initial income, is not valid for all levels of initial income, showing that there are signs of divergence for the relative incomes of Brazilian municipalities. This result is consistent with the bimodality obtained in the non-parametric density estimation using a Kernel Function for income for the years 1970 and 1996. This bimodality, which may be interpreted as the formation of income convergence clubs as proposed by Quah (1996), is tested statistically through a test of multimodality that uses bootstrap methods.

We model the evolution of relative income distribution for Brazilian municipalities using the Distribution Dynamics methodologies proposed by Quah (1996), which model the evolution of income as a Markov process. The advantage of this methodology is that it formulates a law of movement for the entire distribution of incomes between the periods under analysis, allowing us to model the existence of convergence clubs in the data. This Markov process for relative incomes is modeled as a discrete formulation that uses transition matrices, and as a continuous formulation, known as a Stochastic Kernel, which avoids the problems associated with the discretization of the transition process in the estimation of transition matrices.

Our analysis shows that there is evidence for the formation of two convergence clubs, one consisting of the richer municipalities in the Southeast, South and Center-West regions, and another consisting of the relatively poorer municipalities of the Northeast and North regions, and that the hypothesis of convergence to the same income level are rejected by the data.

The database consists of per capita incomes for 3,781 Brazilian municipalities for the years 1970 and 1996, constructed on the basis of income and population data obtained, respectively, from the IPEA and IBGE[2]. This article is organized in the following way: in Section 2 we describe a number of previous studies on convergence in Brazil. In the next section, we carry out a test of sigma convergence using bootstrap methods. Then , in Section 4 we test the hypothesis of beta convergence in a non-parametric fashion using Smoothing Splines, and in Section 5 we estimate densities using kernel functions and test for the presence of bimodality. Section 6 contains the estimated distribution dynamics, while Section 7 presents our conclusions.

2 Previous Studies

Previous studies of income convergence in Brazil used income data at state level almost exclusively, due to the difficulty of obtaining such data for municipalities. The studies by Ferreira & Diniz (1995), and Schwartsman (1996) found $\beta-$convergence in per capita incomes for Brazilian states for the period 1970-85. Azzoni (2001) has criticized this result, pointing out that the period 1970-85 used in these studies was a period of very strong convergence and reduction in income inequalities, but that these convergence dynamics were not subsequently maintained, and has also demonstrated some problems with the construction of the data used in the study.

In reply to Azzoni (2001) criticisms, Ferreira (1998) estimated Markov transition matrices for the state GDP per capita data for years from 1970 to 1995. The results of the ergodic distributions (long-term distribution of per capita incomes) estimated by Ferreira (1998) demonstrate a trend towards concentration in the middle income categories and the disappearance of income categories above 120% of the national mean, with little alteration in the income distributions of the poor and very poor categories.

Using measures of spatial association, Mossi *et al.* (2003) arrived at results that pointed to the polarization of incomes with a strong spatial component. The low income cluster consisted principally of the states of the Northeast region (states of PI, CE, RN, PB and BA), while the states of the South and Southeast region (RJ, SP, PR and MG) formed the high income cluster. Mossi *et al.* (2003) use stochastic transition matrices in their analysis of the evolution of state per capita incomes. The results of their estimation of transition matrices show a high persistence in the extreme categories (they divide their sample into 5 income categories, analyzing the period 1939-98). The estimation of stochastic kernels by Mossi *et al.* (2003) shows the same characteristics of high persistence in both the spatially conditional and the spatially unconditional analyses. The principal results confirm the fact that the dynamics of income distribution are heavily influenced by regional factors, and that there are two income convergence clusters, a low income cluster formed by the states of the Northeast region and a high income cluster formed by the states of the Southeast and South regions.

Using traditional growth regressions estimated by ordinary least squares and quantile regression, Andrade *et al.* (2002) are unable to reject the hypothesis of Beta Convergence for Brazil and for separate regions using the same municipal incomes database as the one in our study. Ribeiro & Pôrto Júnior (2002) study convergence for municipalities in the Southern region for the period 1970-91, finding signs of the formation of convergence clubs within this region, as well as for Brazilian states for the period 1985-98, and demonstrating a trend among Brazilian states towards stratification of income into three groups, a group of poor states, consisting of 26.9% of all states, an average income group consisting of 52% of all states, and a group of rich states, consisting of 11.4% of all states.

By comparison with previous studies, our article uses municipal income data and replaces the parametric sigma and beta convergence tests with non-parametric methodologies, finding more robust results in favor of the hypothesis of formation of convergence clubs within Brazil. The modeling of distribution dynamics that we have used allows us to capture the law of movement of relative per capita income without the problems associated with the discretization of Markov processes, while the results of the process of formation of two convergence clubs are confirmed statistically by a test of multimodality.

[2]The Appendix constains the methodology used in constrution of the database.

3 Sigma Convergence

A simple definition of the process of sigma convergence is that of convergence to a single income point, which may be understood as a continuous dynamic of reduction of the differences in incomes between economies, implying lower dispersion and inequality of incomes.

In order to analyze the dispersion between relative incomes, two measures that are frequently used in the literature to test sigma convergence are the Theil Index and the Coefficient of Variation, which measure the degree of inequality existing in the data. The traditional methods of verifying Sigma convergence with these inequality indicators take the form of constructing a time series with the index values measured for each year, and verifying through a linear regression against time whether there is a significant trend towards the reduction of inequalities, as would be shown by a negative parameter in this regression.

Since we are only using data for the years 1970 and 1996 in our analysis, we tested statistically for a reduction in income inequalities through coefficients of variation and Theil indices estimated for the years 1970 and 1996 by obtaining the distribution of these estimators using the bootstrap method and by constructing confidence intervals for the estimated values.

The bootstrap method treats the available sample as the population, and through repeated resampling of this sample, obtains the distribution of estimators or statistics of the test. Given the need for only weakly restrictive regularity conditions, the bootstrap method allows accurate approximations to distributions in finite samples. The bootstrap method is also advantageous in that it avoids the need for mathematical derivations requiring long computing times where these are excessively complex. Applying the bootstrap method to the Theil Index and Coefficient of Variation, we may test whether the reduction in these estimators is statistically significant, without needing to assume a priori that the data derive from a given distribution.

The use of bootstrap methods for inequality indices was originally introduced by Mills & Zandvakili (1997), with their use justified on the grounds that the inequality indices were non-linear functions of income and hence, the asymptotic properties of these estimators might not be accurate and their properties in finite samples unknown. In addition, since some of the inequality estimators are functions that are limited on the interval [0,1], e.g. the Theil and Gini indices, the confidence intervals obtained using traditional asymptotic theory might not respect these theoretical limits of the estimator.

Table 1 shows the confidence intervals obtained using bootstrap methods for the Theil Index and the Coefficient of Variation for municipal per capita income data for every Brazilian municipality in the years 1970 and 1996. The confidence intervals were obtained using the non-parametric BC_a percentile (Bias Corrected and Accelerated) bootstrap method. This method requires fewer replications of the bootstrap in order to approximate the distributions of estimators correctly and more accurately, and according to Efron & Tibshirani (1993), is also invariant with regard to transformations in the estimators. Table 1 contains the values corresponding to the 0.01, 0.025, 0.05, 0.95, 0.975 and 0.99 percentile points of the distributions obtained using the bootstrap method, which allow the construction of confidence intervals.

The tests of sigma convergence show that there has been a reduction in municipal per capita income inequalities for all regions except the North region, where there was an increase in Theil index and in the Coefficient of Variation. However, the reduction in inequality corresponding to the hypothesis of sigma convergence is only statistically valid for the South region, for which we reject at the 1% significance level the null hypothesis that both the Theil index and the coefficient of variation are the same for the period 1970-96. Note that the 1996 confidence intervals for the South region do not fit the confidence intervals of the two indicators for this region for 1970. For the other regions in which there were reductions in inequality, we are unable to reject on statistical grounds the null hypothesis that the indicators are the same.

One of the necessary conditions for the validity of the results from the bootstrap procedure is that the samples derive from an independent process, although the analyses in the subsequent sections show that there may be a regional factor in income distributions. In order to control this effect, which would represent a violation of the independence requirement for bootstrap methods, we carried out a procedure known as a Stratified Bootstrap method. In this procedure, we resample for every municipality in Brazil with the constraint that the number of municipalities in each region that are included in the each resampling remains constant, which is equivalent to resampling within each region and calculating the result for the whole country.

The distributions obtained for the coefficient of variation and the Theil index for Brazil using a stratified bootstrap approach are shown in Figure 1. The vertical lines mark the values for a confidence interval at the 5% significance level. Table 2 contains the upper and lower values for the confidence intervals obtained by this method and show that the result obtained using the bootstrap method without stratification is maintained. In spite of a reduction in the values calculated for the Theil index and the coefficient of variation, we cannot reject the hypothesis that they are statistically equal between 1970 and 1996.

4 Beta Convergence

The hypothesis of Beta Convergence may be seen as the existence of a negative relationship between the growth rate and the value of initial income, caused by the presence of diminishing returns in the production function[3] used in the growth models of Solow (1956) and Swan (1956). Beta Convergence is a necessary but not a sufficient factor for the existence of sigma convergence, since exogenous shocks in growth rates could increase dispersion between incomes.

The growth regression consists of estimating the following equation:

$$\left(\frac{1}{T}\right) log \left(\frac{Y_{iT}}{Y_{i0}}\right) = \alpha + \beta log\left(Y_{i0}\right) + \mu_{it} \tag{1}$$

where Y_{it} and Y_{i0} are the incomes for the period T and the initial period respectively, T is the number of periods, α and β are constants and μ_{it} is the mean error in the growth rate between the times 0 and T. The hypothesis of Beta Convergence is given by a negative value for β in this equation.

The problem with this approach is that the formation of convergence clubs cannot be captured by a parametric estimation using least squares, since this imposes the same rate of convergence on all levels of income. According to Quah (1996), the concept of convergence clubs is equivalent to the disappearance of intermediate income categories, and the emergence of two attractors for income, a high income and a low income one. This behavior would be visualized by the existence of two distinct peaks in the empirical density of the data.

In order to capture the formation of convergence clubs, we then use non-parametric methods that allow us to estimate the β parameter equivalent to each level of initial income. In order to model the relationship between the rate of growth and initial income without adopting a defined functional form, we used the non-parametric regression technique known as *Smoothing Spline*, which may be defined as the solution to the problem of minimizing the following function:

$$S_\lambda(g) = \sum_{i=1}^{n} \left(Y_i - g(x_i)\right)^2 + \lambda \int \left(g''(x)\right)^2 dx \tag{2}$$

where g can be any curve, x is the data set and λ is a smoothness of adjustment parameter that controls the trade-off between the minimization of the residual and the roughness of the adjustment. According to Hardle (1990), this minimization problem has a single solution $\widehat{m}_\lambda(x)$, given by a cubic polynomial called a *cubic spline*. One of the advantages of this interpolation method is that it produces the first derivatives of the function $\widehat{m}_\lambda(x)$ directly. The first derivative may be interpreted as the measure of response of the dependent variable Y to a change in the explanatory variable x, in an analogous way to the parameters of a linear regression. The sign of the estimated derivative will be our convergence indicator, for which negative values of the derivative indicate income convergence and positive values income divergence.

We determine the smoothing parameter λ using the Generalized Cross Validation criterion, and use Wahba (1983) formulation to construct the confidence intervals for the Smoothing Spline,

[3]Barro & Sala-i Martin (1992) derive the growth regression used in the tests of Beta Convergence.

which, by visualizing this model as a Bayesian model, determined that the confidence intervals for the spline were given by:

$$\widehat{S}(x_i) \pm z_{\alpha/2}\sqrt{\sigma^2 \cdot a_{ii}(\lambda)} \tag{3}$$

where $\widehat{S}_\lambda(x_i)$ are the values predicted by the spline, $z_{\alpha/2}$ is the corresponding value in a normal distribution at the desired confidence level for the given confidence interval, and $a_{ii}(\lambda)$ are the elements of the diagonal of the matrix of leverages $A(\lambda)$ defined by smoothing spline[4]. The variance σ^2 of the smoothing spline is defined as:

$$\widehat{\sigma^2} = \frac{e`e}{tr(A(\lambda))} \tag{4}$$

where e is the error associated with each value in the sample and tr is the trace of the matrix $A(\lambda)$ of leverages. The confidence interval for the derivative of the smoothing spline is obtained from the predicted values for the spline using the delta method. In constructing the confidence intervals, we have assumed a 5% significance level.

The existence of divergence for some income categories would be demonstrated by a positive relationship between the rate of growth and initial income for these incomes, which could be measured by the first derivative of the smoothing spline. The formation of convergence clubs would be given by the existence of a divergence category, corresponding to positive values for the first derivative of the spline for intermediate values of income. This could be interpreted as the disappearance of municipalities with intermediate incomes, which would become part of the group of high income municipalities.

Figures 2, 3, 4, 5, 6 and 7 show the Smoothing Splines and the estimated first derivatives, with the rate of income growth between 1970 and 1996 as the dependent variable and the log of income in 1970 as the explanatory variable for every municipality in Brazil, as well as for municipalities of separate regions, together with the associated confidence intervals. These graphs show that within each region we did not find signs of divergence, demonstrating that the formation of convergence clubs is due to the shift in relative incomes between regions, and not between municipalities within each region.

The non-parametric regression between the rate of growth and initial income in the form of a Smoothing Spline for every municipality in Brazil (Figure 2) shows that for incomes between approximately 5-6 times the logarithm of per capita GDP, and for incomes exceeding 7 times the logarithm of per capita GDP, the relationship between growth rates and initial income is a curve with a negative slope, as expected for the hypothesis of Beta convergence. However, if we observe the logarithm of incomes in 1970 between 6.3552 and 6.7640[5], the adjusted curve, and in particular, the derivative of the Spline show the presence of divergence, indicated by a derivative with positive values that are statistically different from zero. The divergence category is consistent with the values of income that tend to disappear with the formation of convergence clubs, as will be confirmed in sections 5 and 6 by the non-parametric estimations of density and the modeling of distribution dynamics. The results of the Smoothing Splines applied separately to each region do not indicate the presence of divergence, suggesting that the formation of convergence clubs within Brazil has been caused by a uniform shift of relative per capita incomes within each region.

The assumption of a linear relationship between initial income and the growth rate that has been used in traditional tests of convergence has proven itself inadequate. The assumption of the same rate of divergence for all levels of initial income suffers from the problem of reversion to the mean and does not reveal the existence of divergence categories relative to specific levels of income. Non-parametric estimation using Smoothing Splines shows that intermediate incomes are diverging, which we may interpret as the disappearance of these intermediate income categories with the formation of two convergence clubs.

[4]For more details on the components of the Smoothing Spline estimator, see the documentation for the ModReg software package at http://www.r-project.org

[5]Corresponding to per capita incomes of US$ 575.72 and US$ 866.15 in 1970.

5 Non-Parametric Densities

One way of analyzing the distribution of relative per capita incomes is through the visualization of probability density functions. We have estimated the probability density function in a non-parametric form with densities estimated using kernel functions.

A kernel is defined as a continuous, limited and symmetric function, with the property that its indefinite integral is equal to unity:

$$\int K(u)du = 1. \tag{5}$$

This property allows us to construct an estimator for the density as the density function for a scalar Z at the point z_0 may be approximated by:

$$f(z_0) = \lim_{h \to 0} \frac{1}{2h} P(z \in (z_0 - h, z_0 + h)), \tag{6}$$

with an estimator for $\widehat{f}(z_0, z)$ given by:

$$\widehat{f}(z_0, z) = \frac{\#(z \in (z_0 - h, z_0 + h))}{2hn}. \tag{7}$$

Using these properties, the typical form of a density estimator per kernel is given by:

$$\widehat{f}(z_0, z) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{z_0 - z_i}{h}\right) \tag{8}$$

which uses a kernel function $K(u)$:

$$K(u) = \frac{1}{2} I_{(-1,1)}(u) \tag{9}$$

where I is the indicator function. The Gaussian kernel used here is defined as:

$$2\pi^{-1/2} exp\left(-u^2\right). \tag{10}$$

A fundamental factor in the use of density estimators that use a kernel is the choice of the parameter h. This parameter, known as the bandwidth parameter, determines the weighting given to the points $z_i \neq z_0$. The parameter h controls the neighborhood of points used in the estimation of $\widehat{f}(z_0, z)$. Lower values of h lead to a lower number of points used in the estimation of the density around point z_0, with the result that the estimated density for the data is not as smooth. In defining the parameter h, we used Silverman's rule[6] for a Gaussian kernel, which corresponds to 0.9 times the minimum of the standard deviation of the data and the difference between the lowest and highest quartile for the data, multiplied by the sample size, with the result raised to the power of 1/5.

Figure 8 shows the evolution of estimated densities using a Gaussian kernel for the natural logarithm of relative per capita incomes for the years 1970 and 1996. Relative income is constructed by dividing the value of municipal per capita income by the mean per capita income for all the municipalities in the same year, and taking the natural logarithm of this value. In this normalization process, a zero value on the horizontal axis indicates per capita income equal to the national mean, while a value of 0.69 is equivalent to double the national mean, and so on. In this way, the relative income at any point on this axis is the natural logarithm of income relative to the mean for Brazil in the same year.

We may observe the formation of two modes in the distribution for the sequence of densities for relative incomes of Brazilian municipalities, which Quah (1996) has termed 'Twin Peaks'. For 1970, we may observe the start of bimodality. In this year, the upper and lower peaks of

[6] The properties of this rule may be seen in Silverman (1986), pages 48-49.

the distribution correspond to -0.84 and -0.05 of the logarithm of relative income (respective per capita incomes of 0.43 and 0.94 times the Brazilian mean for that year).

For the year 1996, the positions of the lower and upper peaks shift to -1.08 and 0.06 (corresponding to relative incomes of 0.33 and 1.07 of the Brazilian mean in 1996). As shown in Figure 8, the peaks become more pronounced and further apart in 1996 relative to 1970, suggesting the formation of two convergence clubs for the relative incomes of Brazilian municipalities: one group formed of rich municipalities, and another composed of poor ones.

Figure 9 shows the densities obtained from the kernel for the relative incomes of municipalities in the five Brazilian regions for the years 1970 and 1996. These are, in decreasing order of relative income, the Southeast, South, Center-West, North and Northeast regions[7]. The two main messages of Figure 9 are that there are no signs of the formation of convergence clubs within each region, since all the regional densities are unimodal. More importantly, the difference in relative incomes between the poorer regions (North and Northeast) and the richer ones (Southeast, South and Center-West) are increasing over time. By comparison with 1970, the densities of the richer regions have shifted to the right, becoming richer in relative terms. The peaks for the Southeast, South and Center-West have shifted from 0.058, 0.048 and -0.446 in 1970 (i.e. 1.06, 1.05 and 0.64 times the national mean), respectively, to 0.23, 0.029 and -0.1625 in 1996 (equivalent to 1.27, 1.03 and 0.85 times mean Brazilian income). The opposite occurred with incomes in the poorer North and Northeast regions, which shifted from -0.713 and -1.139 (0.49 and 0.32 times the national mean) respectively, to -0.891 and -1.17 in 1996 (0.41 and 0.31 times the mean income for the year).

The estimated densities for Brazil (Figure 8) and for the regions (Figure 9) indicate that municipalities in the North and Northeast regions are largely responsible for forming the lower income peak, while municipalities in the Center-West, Southeast and South regions form the higher income peak. In this way, the poorer municipalities have in general become poorer in relative terms, while the richer municipalities have become richer in relative terms, which is equivalent to the definition of convergence club formation advocated by Quah (1996).

5.1 Tests of Multimodality

In order to verify whether the multimodality that exists in the non-parametrically adjusted density is statistically significant, we used a test of multimodality[8] based on the bootstrap principle proposed by Silverman (1981), using the algorithm described by Efron & Tibshirani (1993). Since the density adjusted by a kernel approach does not take on a functional form or distribution, this test of multimodality is based on finding through bootstrap a test distribution for the hypothesis of m modes against m+1 modes.

Since the number of modes found in the density function estimated using the kernel is a function of the bandwidth used, Silverman (1981) proposes the use of the difference between the bandwidth that constrains m modes in the data and the bandwidth that determines m+1 modes as a test statistic. Since the number of modes is a non-increasing function of the chosen bandwidth, the test of multimodality is based on using an adjusted density distribution for the data that is a function of the minimum bandwidth required for inducing the null hypothesis of m modes.

In accordance with Efron & Tibshirani (1993), we defined the adjusted density as $\widehat{f}(t, h_1)$, using equation 8, where t is the sample size and h_1 the bandwidth required to induce the number of modes assumed in the null hypothesis. Kernel estimations artificially increases the variance of the estimation, for which reason it is necessary to adjust it in such a way as to make it equivalent to that of the sample, defining a new density $\widehat{g}(t, h_1)$, so that the test statistic is constructed using the value estimated for h_1. A higher value of h_1 indicates that a greater degree of smoothing is

[7] The number of municipalities analyzed in each region is: Southeast (1393), South (671), Center-West (226), North (160), and Northeast (1331), in accordance with Table 6 presented in the Appendix.

[8] This test of multimodality was used by Bianchi (1997) to test the hypothesis of income convergence for a group of 119 countries between the years of 1970 and 1989. Bianchi (1997) rejects the hypothesis of convergence in favor of the formation of convergence clubs.

required to induce m modes in the density function by comparison with the value adjusted by Silverman's criterion for the estimation bandwidth.

The test of the hypothesis based on bootstrap replications is obtained by holding constant the estimated value for h_1 (minimum bandwidth for inducing m modes). The significance level for the test is obtained through the probability that in the n replications of the bootstrap, the minimum value h_1^* for inducing m modes in each replication is greater than the value h_1 obtained from the observed data. The significance level obtained via the bootstrap method is given by:

$$SL = Prob_{g(t^*,h_1)}\{h_1^* > h_1\} \tag{11}$$

In order to obtain equation (11) using replications with the same variance as the original data, we used the smooth bootstrap of Efron & Tibshirani (1993)[9]. The significance levels were obtained with 2000 bootstrap replications.

The results of the multimodality tests (Table 3) applied to the logarithm of the relative incomes for every municipality in Brazil in 1996 show that we have obtained an empirical significance level of 0.0474 for the null hypothesis of 1 mode, indicating that we would only refrain from rejecting it in 4% of the replications. This suggest that we can reject the hypothesis at the 5% significance level that the distribution of relative incomes is unimodal, in favor of the alternative hypothesis of bimodality. When we assume a null hypothesis of 2 modes, the empirical significance level is 0.7611, indicating that we should not reject it and suggesting the formation of two convergence clubs for municipal incomes in Brazil.

The tests of multimodality also confirm the evidence shown by graphs of adjusted densities for the five regions of the country that each region is converging towards a unimodal distribution. The lowest significance level for the null hypothesis of unimodality in 1996 was obtained for the Northeast, with a value of 0.2148, which in visual terms presents a more heterogeneous density. This result is consistent with an interpretation that the formation of two convergence clubs within Brazil is due to a shift in relative incomes in the North and Northeast regions to lower levels, and to higher income levels in the Center-West, Southeast and South regions, with each region shifting while maintaining a single peak.

6 Distribution Dynamics

Since the hypothesis of unimodal convergence is rejected by non-parametric methods, we now use the methodology of Distribution Dynamics to model the evolution of the relative distribution of per capita incomes for Brazilian municipalities. This approach models directly the evolution of relative income distributions as a first order Markov process[10].

The modeling of Distribution Dynamics assumes that the density distribution ϕ_t has evolved in accordance with the following equation:

$$\phi_{t+1} = M \cdot \phi_t, \tag{12}$$

where M is an operator that maps the transition between the income distributions for the periods t and t+1. Since the density distribution ϕ for the period t only depends on the density ϕ for the immediately previous period, this is a first order Markov process. In order to capture the dynamics of relative incomes between 1970 and 1996, we require an operator M that determines the evolution from graph (a) to graph (b) in Figure 8. Equation (12) may be seen as analogous to a first order autoregression in which we replace points by complete distributions.

The operator M may be constructed either by assuming that distribution ϕ_t has a finite number of states, using the model known as Markov's Transition Matrices, or by avoiding discretization and modeling M as a continuous variable, in what is known as a Stochastic Kernel. The application of Markov transition matrices is carried out in Section 6.1, and continuous modeling using Stochastic Kernels is carried out in Section 6.2.

[9]See page 231.

[10]This methodology was popularized through the work of Quah (1996, 1998).

6.1 Discrete Modeling - Markov Transition Matrices

We shall assume that the probability of variable s_t taking on a particular value j depends only on its past value s_{t-1} according to the following equation:

$$P\{s_t = j | s_{t-1} = i, s_{t-2} = k, ...\} = P\{s_t = j | s_{t-1} = i\} = P_{ij} \tag{13}$$

This process is described as a first order Markov chain with n-states, where P_{ij} indicates the probability that state i will be followed by state j. As:

$$P_{i1} + P_{i2} + \ldots + P_{in} = 1 \tag{14}$$

we may construct the so-called transition matrix, where line i and column j give the probability that state i will be followed by state j:

$$P = \begin{bmatrix} P_{11} & P_{12} & ... & P_{1n} \\ P_{21} & P_{22} & ... & P_{2n} \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ P_{n1} & P_{n2} & ... & P_{nn} \end{bmatrix} \tag{15}$$

The use of Markov chains to model the evolution of the distribution of relative incomes between Brazilian municipalities reposes on the idea that each state of this matrix represents a category of relative income. The transition matrix estimated for the relative per capita incomes of the municipalities was constructed in such a way as to have nine states. We determined that the nine categories of income would be limited by the following vector of relative incomes with regard to the mean value of national[11] per capita income for the year under analysis: {0-0.25, 0.25-0.5, 0.5-0.75, 0.75-1, 1-1.25, 1.25-1.5, 1.5-1.75, 1.75-2, 2-∞}[12].

The probability p_{ij} measures the proportion of municipalities in regime i during the previous period that migrate to regime j in the current period. According to Geweke & Zarkin (1986), the maximum likelihood estimator for the transition probability p_{ij} is given by:

$$\widehat{p_{ij}} = \frac{\sum m_{ij}}{\sum m_i}, \tag{16}$$

where $\sum m_{ij}$ is the number of municipalities that were in income category i in the previous period and have migrated to income category j in the current period, and $\sum m_i$ is the total of municipalities that were in income category i in the previous period.

A transition matrix defined in this way presents some interesting characteristics in the study of mobility. The first is that, given the transitions estimated for the period, the probabilities of transition for n periods ahead may be forecast by the transition matrix multiplied by itself n times, in accordance with Hamilton (1994). The second relevant characteristic is the fact that the estimated transition probabilities point to the relative long-term distributions of income, known as an ergodic distribution.

The ergodic distribution may be found if we note that since the transition matrix requires that each row sums to unity, one of the eigenvalues of this matrix must necessarily have a value equal to unity. If the other eigenvalues are within the unit circle, the transition matrix is said to be ergodic and thus possesses an unconditional distribution. This unconditional distribution vector, which in our case will represent the long-term distributions of relative income, is the eigenvector associated with the unitary eigenvalue of the transition matrix.

Table 4 shows the transition matrix estimated for the relative incomes data for Brazilian municipalities. The first column contains the number of municipalities in each income category in 1970. The matrix formed by rows 2-10 and columns 3-11 contains the matrix of probabilities of transition between income categories, while the last row of the matrix contains the ergodic distributions imposed by the estimated transition matrices.

[11]See discussion below on the problems of ad hoc choice of the discretization of the number of natural states.

[12]The same discretization process is used when we analyze the transition of municipalities by region.

The estimated values show that there is a relatively high mobility for the intermediate income categories. This means that, on the one hand, the very poor groups tend to remain very poor, with the probabilities of remaining at the previous level of income for the two lowest income categories equal to 0.39 and 0.56. On the other hand, the very rich groups tend to remain very rich, with a probability of 0.5 of remaining in the highest category of income. Moreover, the intermediate categories between these values are more likely to migrate to higher or lower levels of income than to remain at the same level of income, confirming the trend for middle income categories to disappear.

The ergodic distributions of income for Brazil and for each individual region are shown in Table 5. The long-term distribution shows that Brazil has two peaks, one of income between 0.25-0.5 times the mean income for the year, including 0.246 of the total of municipalities in the country, and another peak, consisting of 0.118 of municipalities with incomes between 0.75-1 times the mean income, pointing to the formation of two convergence clubs. The distribution of relative income in each region is basically unimodal. This may be seen from the fact that for the Northeast and North regions, the long-term distribution is concentrated in the income category between 0.25-0.5 of the mean income, while for the Center-West, Southeast and South regions, the long-term distributions are concentrated among incomes greater than 0.75 times the mean national income.

The result obtained suggests that the hypothesis of convergence to a single point is rejected when we use Markov transition matrices. The evidence captured by these matrices suggests that we may consider the existence of 2 income peaks, one for the poorest municipalities with incomes of between 0.25-0.5 of the mean national per capita income and another for municipalities with incomes of over 0.75 times national per capita GDP. This evidence points to the existence of two convergence clubs for relative per capita income among Brazilian municipalities.

The results obtained using transition matrices formed by the discretization of the number of states are subject to two serious problems. The first is that the number of intervals in the matrix and the limit values for each interval are determined in an ad hoc way by the researcher, which may significantly alter the results obtained. The second problem is that the discretization process may eliminate the property of Markovian dependence that exists in the data, as Bulli (2001) has pointed out, with loss of information inherent to the discretization process, a phenomenon known in the literature on Markov chains as the problem of *Alliasing*. The solution to this problem consists in carrying out a continuous analysis of transition, which avoids discretization through the use of conditional densities that are estimated non-parametrically and known as Stochastic Kernels. We shall do this in the following section.

6.2 Stochastic Kernels

In order to avoid the problems associated with discretization in the estimation of transition matrices, we may estimate directly a continuous transition function between relative per capita incomes in 1970 and 1996 in a non-parametric form. This continuous transition function receives the name of Stochastic Kernel and is basically an estimate of a conditional bivariate density function for which we condition the function to the values of income in the initial year. In formal terms, a stochastic kernel is defined as:

Definition: Let $M_{(u,v)}$ and (R, $\Re$) be measurable spaces. A **Stochastic Kernel** on (M,$\Re$) is a function $M_{(u,v)(y,A)} \times (R, \Re) \rightarrow [0,1]$ such that:

a. for each y $\in$ R , $M(u,v)(y,A)$ is a probability measure on (R,$\Re$);

b. for each A $\in \Re$, $M(u,v)$ is a measurable function on $\Re$;

c. for each A $\in \Re$, u(A)=$\int M_{(u,v)}(y,A)\, dv(y)$ is valid.

Conditions (a) and (b) ensure that the stochastic kernel is a well-defined mapping for the probability spaces $M_{(u,v)}$ e (R, $\Re$). The principal concept of a stochastic kernel lies in condition (c). Given an initial period t for a given income y, there is a fraction dv(y) of economies with income close to y. For the period t+n, part of the economies contained in dv(y) will shift to a subset A$\subseteq$ R.. Normalizing this fraction of economies by the total number of economies, we have the definition of the stochastic kernel $M(u,v)(y,A)$. The integral $\int M_{(u,v)}(y,A)\, dv(y)$ represents the total number of economies that will be present in the subset A of economies in the period t+n regardless of initial income. In this integral $M(u, v)(y, A)$ represents the total of economies that have migrated from y to A, while dv(y) is the weighting associated with each M(·) given by the marginal distribution of y[13]. In this way, the stochastic kernel may be visualized as the continuous form of the transition matrix, in which we have a continuum of rows and columns - i.e. we form a continuum of states.

The estimation of the stochastic kernel is carried out by obtaining empirical measures for the elements of the integral $\int M_{(u,v)}(y, A)\, dv(y)$. The term $\int M_{(u,v)}(y, A)\, dv(y)$ is obtained by estimating non-parametrically the joint density of relative incomes for the periods t and t+n, using a bivariate kernel, with this joint density becoming a stochastic kernel when we normalize it using the marginal distribution in t, which is the empirical counterpart of dv(y). An important property is the fact that the transition probability is independent of the transition period t, corresponding to a stationary transition density, as per Quah (1996) and Bulli (2001).

Figure 10 shows the estimated Stochastic Kernel for Brazilian municipalities, while Figure 11 contains the Stochastic Kernels estimated for each region separately[14]. Our interpretation of Figure 10 is as follows: the transition probability associated with the change in an income interval in period t to another income interval in period t+n may be visualized by calculating the probability defined by the stochastic kernel at the intersection of the income interval on the t axis with the interval on the t+n axis. This projection is analogous to the cell formed by the crossing of income intervals in the rows and columns of the transition matrix.

Figure 10 suggests that there are two pronounced peaks in the stochastic kernel, showing that there are two regions of income concentration and confirming the formation of two convergence clubs, one of which captures the convergence among the poorer incomes, and the other peak, the convergence towards higher incomes. This figure also points to the fact that the transition probability estimated for the intermediate income regions is very low. This result confirms the one found above in the discrete analysis, that there is a tendency for intermediate income regions to disappear. Consistent with the results shown in sections 5 and 5.1, the hypothesis of convergence to a single point is not valid for relative per capita incomes of Brazilian municipalities.

The separate estimation of stochastic kernels for each region of the country (Figure 11) aims to analyze whether for each region there is a single point of convergence, equivalent to only one point of the stochastic kernel, or whether there is some other trend towards the formation of convergence clubs within each region.

Figure 11 suggests that in none of the stochastic kernels estimated for each region is there any significant trend towards the formation of convergence clubs. This result is important in that it shows that the formation of two convergence clubs for Brazilian municipalities as a whole is caused by the fact that the regions are moving towards one or other point of convergence. The other possible hypothesis is that some of the municipalities in each region are moving towards one or other of the clubs, which would be demonstrated by the existence of more than one peak in the stochastic kernels estimated for each region, something that does not occur according to Figure 11. This figure also shows the shift in incomes in the North and Northeast regions to lower levels of relative income, which may be seen by noting that there are higher probabilities of migration to regions with lower relative incomes in 1996. This result contrasts with those obtained for the Center-West, Southeast and South regions, which show a higher probability of transition to higher income levels in 1996.

7 Conclusions

The analyses carried out in this article point to the formation of two income convergence clubs for Brazilian municipalities during the period 1970-96: a low income club formed by the Northeast and

[13] Quah (1998) on pages 75-77 formalizes the other necessary conditions.

[14] The Stochastic Kernels were estimated in the same way as the univariate kernel, using a Gaussian kernel function, with the bandwidths calculated using Silverman's rule. In order to facilitate visualization, the figures are concentrated in regions with higher transition probabilities.

North regions and another higher income club formed by the municipalities of the Center-West, Southeast and South regions. The evidence shows that the differential in relative income between these two convergence clubs became even greater for the year 1996 relative to 1970 values.

The estimation of growth regressions using Smoothing Splines succeeded in identifying the formation of convergence clubs. This analysis provides a direct answer to the basic hypothesis of convergence club formation, i.e. that income levels tend to disappear over time, as municipalities that were in an intermediate income category migrate to higher or lower levels of relative income. By relaxing the assumption of the same rate of convergence for all municipalities, which is assumed in the estimation by ordinary least squares, the estimation of a non-parametric regression using Smoothing Splines permits the identification of convergence clubs. These estimations also suggest that for each individual region there is no divergence. This result corroborates the hypothesis that the formation of clubs is due to a uniform distancing of relative income in the North and Northeast regions relative to the values for the Center-West, Southeast and South regions.

The existence of two peaks in the distribution of relative incomes for Brazilian municipalities was showed by the estimation of densities using a kernel function, with this bimodality confirmed by a test for bimodality using bootstrap methods. When this test was applied separately for each region, it showed that it was not possible to reject the null hypothesis of unimodality, confirming that the process of convergence club formation is due to a uniform shift in relative income distribution for the Brazilian regions.

In order to capture the process of convergence club formation for Brazilian municipalities, we model the evolution over time of relative incomes in Brazil as a first order Markov process, using the methodology of Distribution Dynamics advocated by Quah (1996). We estimate this process as a discrete formulation using transition matrices, and in continuous form using the method of Stochastic Kernels. The results suggest that a suitable dynamic for relative incomes of Brazilian municipalities is the formation of two convergence clubs rather than the convergence process forecast by the neoclassical growth model.

The non-parametric methods used showed themselves capable of overcoming existing problems in traditional estimations of convergence and are consistent with the results of two income clusters found by Mossi *et al.* (2003), as well as with the results obtained by Ferreira *et al.* (2003). These results suggest that the main areas of relative poverty within Brazil are primarily located in the Northeast region, followed by the North region, and that the other regions have higher levels of income. The results presented here show that the new models of economic growth that include poverty traps, such as those of Becker *et al.* (1990), may be adequate for representing the dynamics of relative incomes in Brazilian municipalities.

References

Andrade, E., Laurini, M.P., Madalozzo, R., & Valls, P.L. 2002. Convergence Across Municipalities Using Quantile Regression. *Ibmec Working Papers*, **14**.

Azzoni, C.R. 2001. Economic Growth and Regional Income Inequality in Brazil. *The Annals of Regional Science*, **35**, 133–152.

Barro, R., & Sala-i Martin, X. 1992. Convergence. *Journal of Political Economy*, **100:2**, 223–251.

Becker, G., Murphy, K. M., & R., Tamura. 1990. Human Capital, Fertility and Economic Growth. *Journal of Political Economy*, **98:5**, 12–37.

Bianchi, M. 1997. Testing for Convergence - Evidence from Non Parametric Multimodality Tests. *Journal Of Applied Econometrics*, **12:4**, 393–410.

Bulli, S. 2001. Distribution Dynamics and Cross-Country Convergence: A New Approach. *Scottish Journal of Political Economy*, **48:2**, 226–243.

Efron, B., & Tibshirani, R, J. 1993. *An Introduction to Bootstrap.* Nova York: Chapman Hall.

Ferreira, A.F.H. 1998. Concentração Regional e Dispersão das Rendas Per Capita Estaduais: Um Comentário. *Texto para Discussão Cedeplar-UFMG.*, **121**.

Ferreira, A.F.H, Lanjouw, P., & Neri, M. 2003. A Robust Poverty Profile for Brasil Using Multiple Data Sources. *Revista Brasileira de Economia*, **57:1**, 59–92.

Ferreira, A, H., & Diniz, C. C. 1995. Convergência Entre Rendas Per Capita Estaduais no Brasil. *Revista de Economia Política*, **15:4**, 38–56.

Friedman, M. 1992. Do Old Fallacies Ever Die ? *Journal of Economic Literature*, **30**, 2129–2132.

Geweke, J. Marshall, R. C., & Zarkin, G. 1986. Mobility Indices in Continuos Time Markov Chains. *Econometrica*, **54**, 1407–1423.

Hamilton, J. D. 1994. *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.

Hardle, W. 1990. *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.

Mills, J., & Zandvakili, S. 1997. Statistical Inference via Bootstrapping for Measures of Inequality. *Journal of Applied Econometrics*, **12**, 133–150.

Mossi, M.B., Aroca, P., Fernandez, I. J., & Azzoni, C.R. 2003. Growth Dynamics and Space in Brazil. *Internation Regional Science Review*, **26**, 393–418.

Quah, D. 1993. Galton's Fallacy and the Tests of the Convergence Hypothesis. *Scandinavian Journal of Economics*, **95**, 427–443.

Quah, D. 1996. Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *The Economic Journal*, **106**, 1045–1055.

Quah, D. 1998. TsRf Reference Manual. *Internet adress: http://econ.lse.ac.uk/staff/dquah/d/tsrf/0t.pdf*.

Ribeiro, E. P., & Pôrto Júnior, S. S. 2002. Crescimento e Convergência - Uma Análise Empírica para a Região Sul. *Anais do Encontro de Economia da Região Sul*, **2000**.

Schwartsman, A. 1996. Convergence Across Brazilian States. *Manuscript*.

Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.

Silverman, B.W. 1981. Using Kernel Density Estimates to Investigate Multimodality. *Journal of The Royal Statistical Society, Series B*, **43**, 97–99.

Solow, R. 1956. A Contribution to the Theory of Economic Growth. *Quaterly Journal of Economics*, **70**, 65–94.

Swan, T. W. 1956. Economic Growth and Capital Accumulation. *Economic Record*, **32**, 334–361.

Wahba, G. 1983. Bayesian Confidence Intervals for the Cross Validated Smoothing Splines. *Journal of the Royal Statistical Society, Series B*, **45**, 133–150.

Appendix - Tables, Figures and Database

Tables

Table 1: Bootstrap Confidence Intervals - Theil Index 1970 and 1996

Region	Theil							CV						
	Value	0.01	0.025	0.05	0.95	0.975	0.99	Value	0.01	0.025	0.05	0.95	0.975	0.99
Brazil 70	0.3550	0.3119	0.3173	0.3233	0.4032	0.4147	0.4251	1.1598	1.0017	1.0022	1.0406	1.3736	1.4090	1.4417
Brazil 96	0.3249	0.2971	0.2999	0.3030	0.3702	0.3813	0.3922	1.0074	0.8846	0.8960	0.9062	1.2405	1.2927	1.3056
North 70	0.1428	0.1106	0.1156	0.1194	0.1756	0.1827	0.1936	0.5786	0.5001	0.5132	0.5232	0.6591	0.6762	0.6893
North 96	0.1688	0.1205	0.1265	0.1330	0.2156	0.2234	0.2322	0.6588	0.5411	0.5564	0.5749	0.7585	0.7782	0.7912
Northeast 70	0.2095	0.1644	0.1684	0.1729	0.3068	0.3190	0.3350	0.8968	0.6656	0.6878	0.7126	1.2855	1.3338	1.3398
Northeast 96	0.1934	0.1575	0.1627	0.1674	0.2286	0.2335	0.2426	0.7899	0.6838	0.6975	0.7104	0.8912	0.9195	0.9365
Center 70	0.1624	0.1103	0.1174	0.1232	0.2339	0.2492	0.2711	0.6917	0.5131	0.5382	0.5613	0.8826	0.9146	0.9289
Center 96	0.1429	0.1033	0.1062	0.1102	0.2362	0.2583	0.2584	0.6226	0.4678	0.4783	0.4818	0.8968	0.9315	0.9315
Southeast 70	0.3245	0.2624	0.2705	0.2774	0.4025	0.4200	0.4330	1.1198	0.8951	0.9278	0.9560	1.3538	1.4153	1.4925
Southeast 96	0.2459	0.2049	0.2095	0.2143	0.3195	0.3388	0.3559	0.9059	0.7355	0.7483	0.7606	1.1900	1.2667	1.3432
South 70	0.1495	0.1154	0.1176	0.1196	0.2277	0.2528	0.2570	0.6791	0.5144	0.5214	0.5272	0.9989	1.0206	1.0218
South 96	0.0868	0.0736	0.0758	0.0773	0.0981	0.1023	0.1039	0.4451	0.4015	0.4089	0.4143	0.4782	0.4851	0.4927

Table 2: Stratified Bootstrap - Theil Index and Coefficientof Variation: 1970-1996

Startified Bootstrap	0.01	0.025	0.05	0.95	0.975	0.99
Theil 1970	0.3022	0.3124	0.3173	0.3964	0.4073	0.4250
Theil 1996	0.2890	0.2947	0.2991	0.3553	0.3621	0.3812
Coeficient of Variation 1970	0.9468	0.9902	1.0076	1.3086	1.3409	1.3955
Coeficient of Variation 1996	0.8375	0.8594	0.8736	1.1534	1.1844	1.2678

Table 3: Tests of Multimodality

	SL 1 mode	SL 2 modes	h - Silverman	h - 1 mode	h - 2 modes
Brazil- 1970	0.8710	*	0.1319	0.1498	0.1289
Brazil - 1996	0.0474	0.7611	0.1398	0.3603	0.1047
North - 1970	0.4882	*	0.1528	0.1466	0.0898
North - 1996	0.4563	*	0.1503	0.1412	0.0821
Northeast - 1970	0.2148	*	0.1042	0.3424	0.0684
Northeast - 1996	0.2498	*	0.0878	0.1879	0.1722
Center 1970	0.6926	*	0.1340	0.1467	0.1062
Center 1996	0.6991	*	0.1505	0.1341	0.2889
Southeast - 1970	0.7156	*	0.1399	0.1788	0.1540
Southeast - 1996	0.7836	*	0.1320	0.1264	0.1091
South- 1970	0.5237	*	0.1168	0.1881	0.0857
South - 1996	0.7841	*	0.0898	0.0855	0.0459

Table 4: Transition Matrix - Brazil

Income		-$\infty - 0.25$	0.25-0.5	0.5-0.75	0.75-1	1-1.25	1.25-1.5	1.5-1.75	1.75-2	2-∞
277	-$\infty - 0.25$	0.39	0.49	0.08	0.01	0.00	0.01	0.00	0.00	0.01
1011	0.25-0.5	0.21	0.56	0.13	0.05	0.02	0.01	0.00	0.01	0.01
659	0.5-0.75	0.04	0.30	0.25	0.21	0.10	0.05	0.02	0.01	0.02
536	0.75-1	0.02	0.12	0.19	0.27	0.17	0.09	0.05	0.03	0.06
368	1-1.25	0.00	0.06	0.08	0.22	0.24	0.15	0.10	0.05	0.09
266	1.25-1.5	0.01	0.03	0.03	0.15	0.20	0.21	0.15	0.09	0.12
200	1.5-1.75	0.00	0.02	0.04	0.10	0.19	0.16	0.16	0.09	0.22
125	1.75-2	0.01	0.02	0.01	0.06	0.18	0.21	0.17	0.14	0.22
337	2-∞	0.00	0.01	0.01	0.04	0.04	0.14	0.13	0.12	0.50
Erg. Dist.		0.100	0.246	0.106	0.118	0.106	0.091	0.069	0.046	0.116

Table 5: Ergodic Distributions

Income	-$\infty - 0.25$	0.25-0.5	0.5-0.75	0.75-1	1-1.25	1.25-1.5	1.5-1.75	1.75-2	2-∞
North	0.124	0.519	0.260	0.075	0.015	0.004	0.003	0.000	0.000
Northeast	0.295	0.587	0.081	0.017	0.004	0.006	0.005	0.001	0.003
Center	0.001	0.014	0.056	0.092	0.233	0.144	0.180	0.092	0.187
Southeast	0.003	0.042	0.085	0.137	0.129	0.131	0.106	0.080	0.287
South	0.001	0.016	0.069	0.252	0.246	0.176	0.107	0.062	0.070
Brazil	0.100	0.246	0.106	0.118	0.106	0.091	0.069	0.046	0.116

Figures

Figure 1: Stratified Bootstrap - Theil Index and Coefficient of Variation

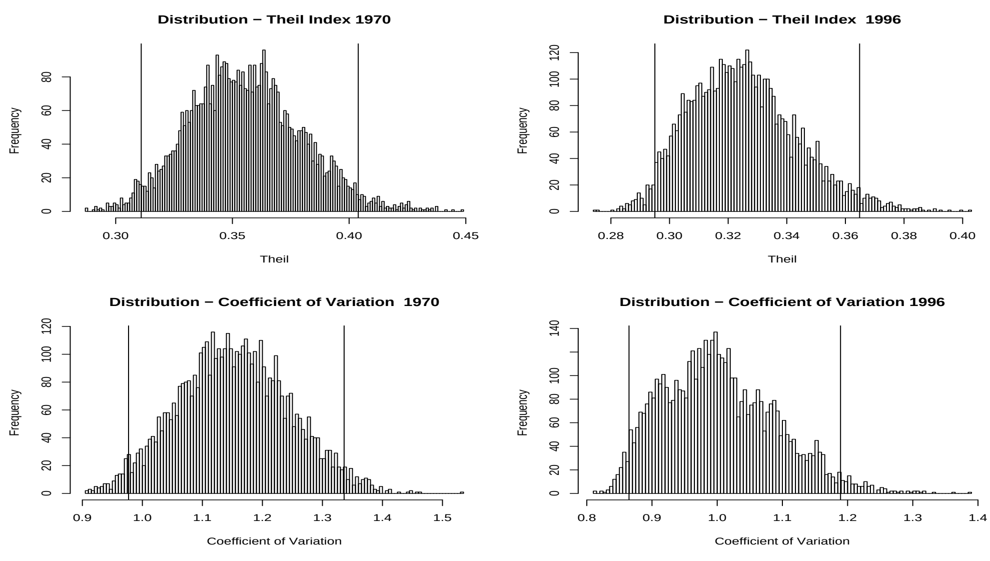


Figure 2: Growth Regression - Smoothing Spline -Brazil

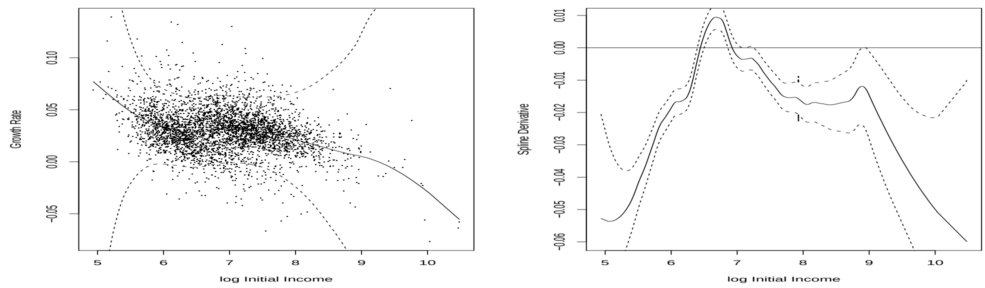


Figure 3: Growth Regression - Smoothing Spline - North

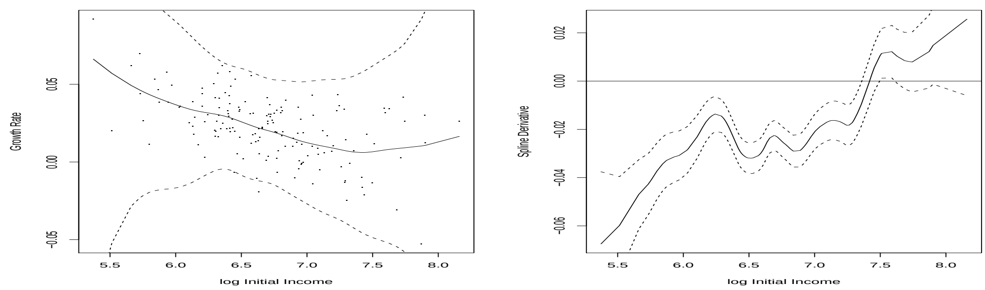


Figure 4: : Growth Regression - Smoothing Spline - Northeast

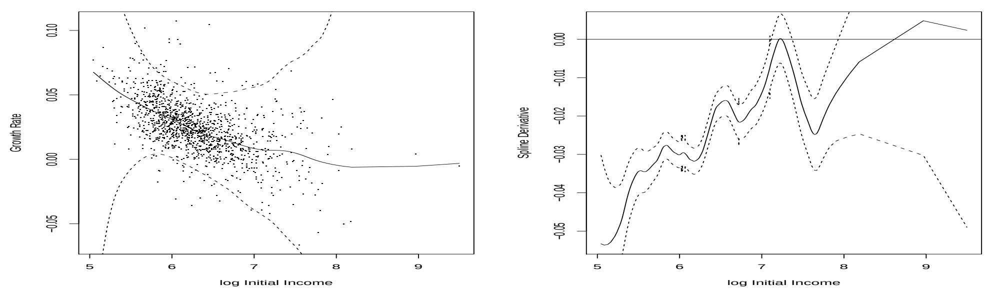


Figure 5: Growth Regression - Smoothing Spline - Center-West

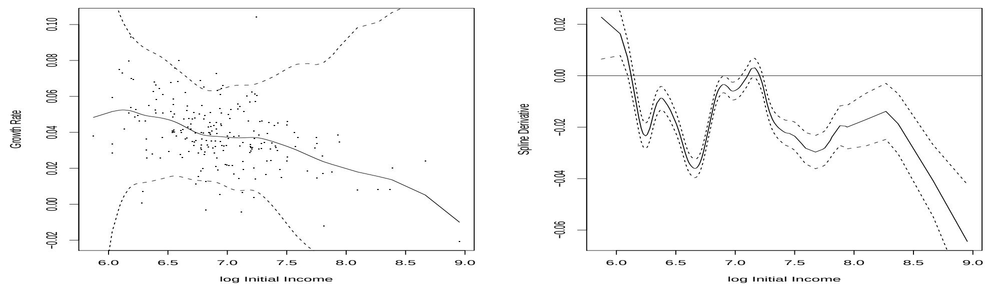


Figure 6: Growth Regression - Smoothing Spline - Southeast

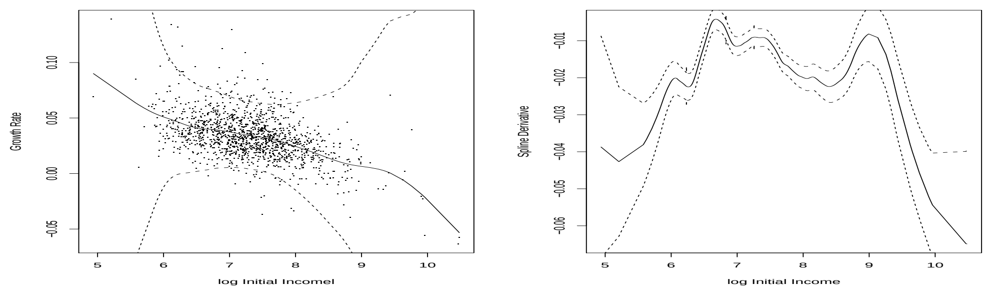


Figure 7: Growth Regression - Smoothing Spline - South

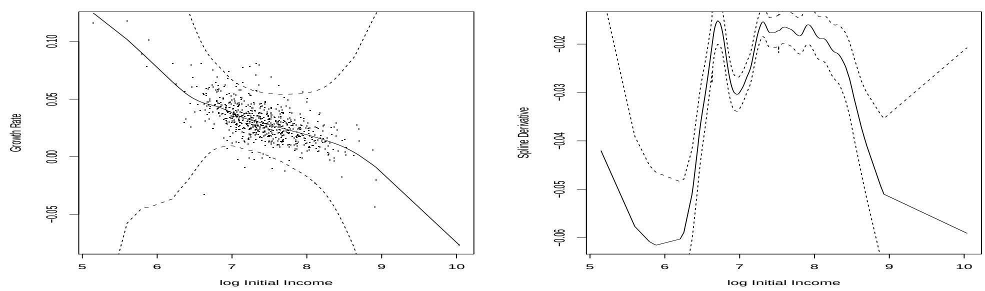
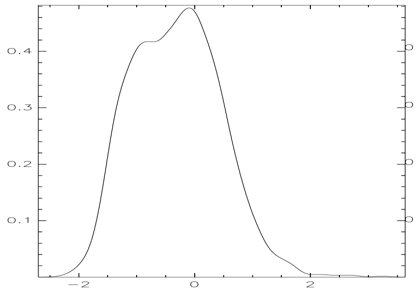
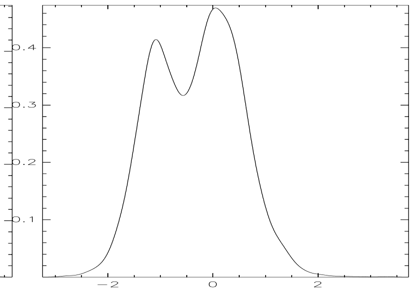


Figure 8: Sequence of Densities - Brazil - 1970 and 1996



(a) Density - 1970



(b) Density - 1996

Figure 9: Densities - Brazilian Regions - 1970 and 1996

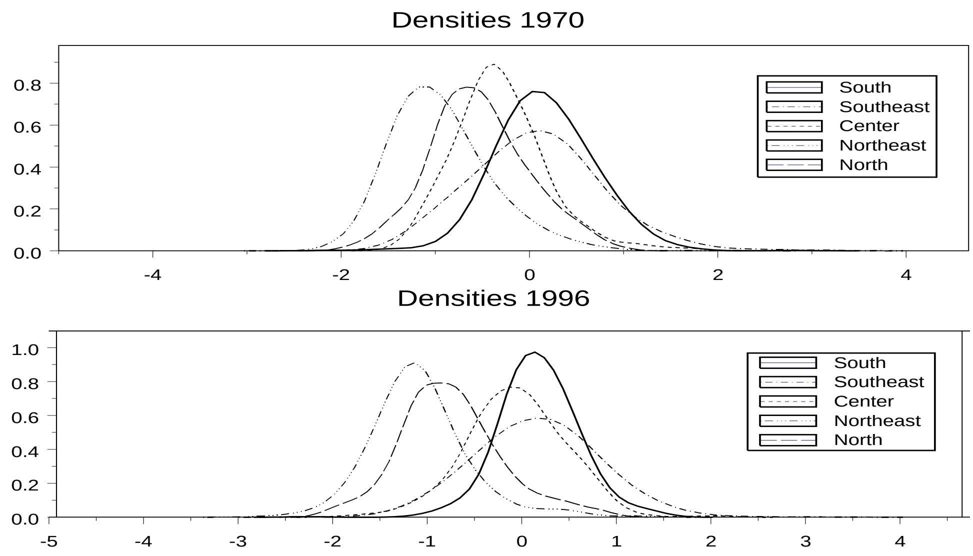


Figure 10: Stochastic Kernel - Brazil (1970-1996)

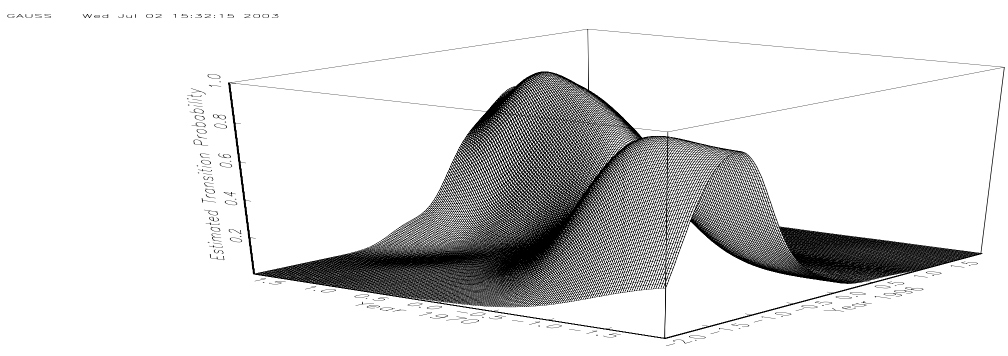
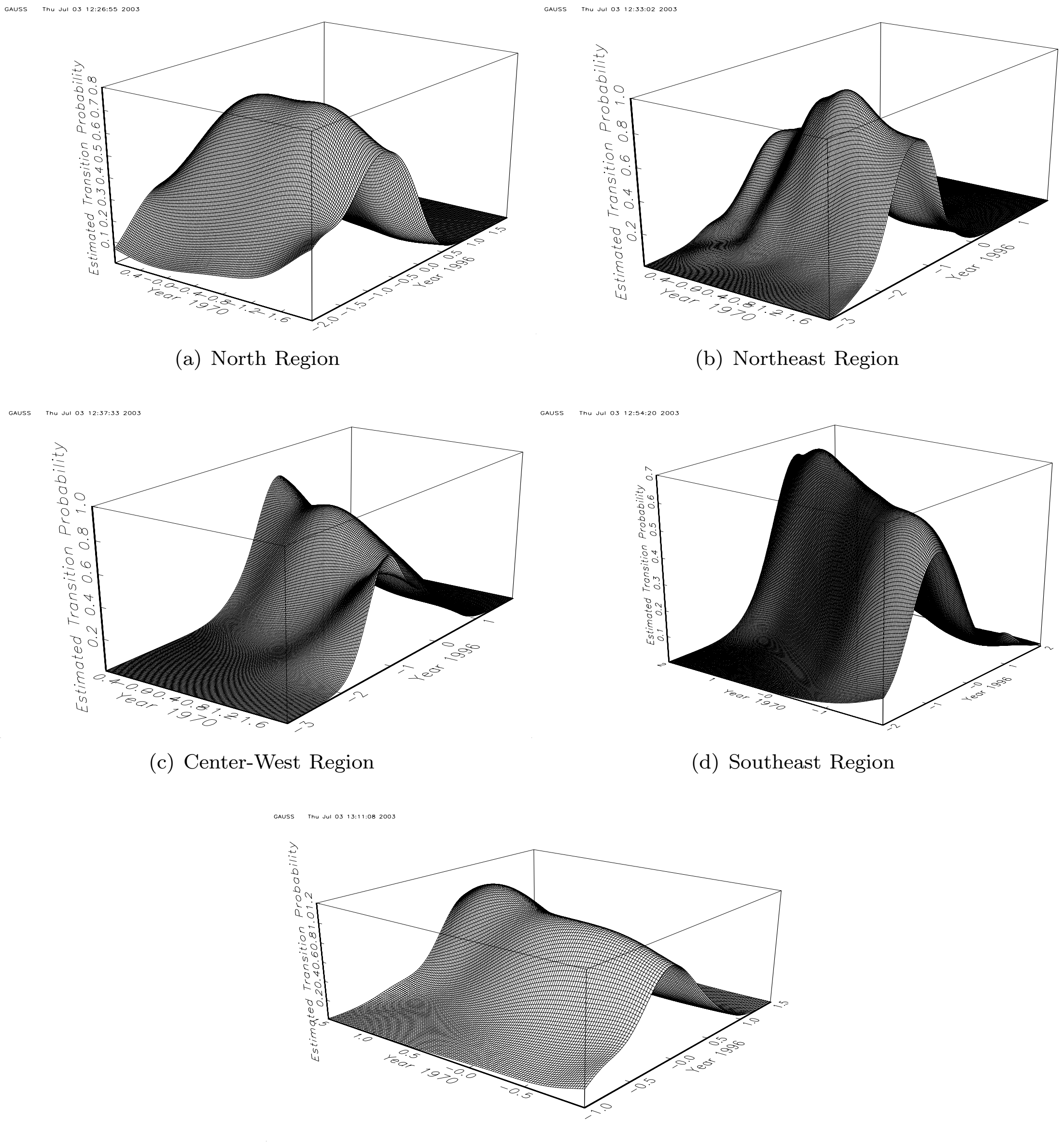


Figure 11: Stochastic Kernels - Regions (1970-1996)



(a) North Region

(b) Northeast Region

(c) Center-West Region

(d) Southeast Region

(e) South Region

Database

Data Construction

The data used in this paper come from two sources. Information about the population of each municipality is from IBGE, the Brazilian Bureau of Statistics. Information about GDP comes from IPEA, a Brazilian government research institute, which calculates a proxy for the GDP of each municipality.

The proxy for GDP is obtained in the following way. First, IPEA calculates a proxy for the value added in each of the three main sectors in the economy (agriculture, industry and services), for each municipality. For example, in order to construct a proxy for GDP in agriculture, IPEA uses the Municipal Agricultural Census which provides information on gross total production and total expenditures in the local agricultural sector. Subtracting expenditures from the value of production, one obtains a proxy for the value added by agriculture in each municipality. Similar procedures are employed to obtain a proxy for the valued added in industry and services.

Second, IPEA obtains the value added by every sector for each of the 27 states in Brazil, by adding up the proxies for the value added by every sector in all the municipalities of each state. Third, IPEA calculates, for every sector, the share of each municipality in its own state's value added in each sector. Fourth, IPEA multiplies this share by the state's sector GDP. Sector GDP's for each state are calculated by IBGE. This step produces an estimate of sector GDP for each municipality. Finally, the proxy for total GDP of each municipality is obtained by adding up the proxies for GDP's of all sectors (agriculture, industry and services).

It was necessary to make some adjustments in the raw data because the number of municipalities in Brazil increased substantially from 1970 to 1996. In 1970, Brazil was divided into $3,946$ municipalities. This number jumped to $4,988$ in 1996. The approach used in this paper is to work as if no new municipalities were created after 1970. In order to follow this strategy, it was necessary to make adjustments in the raw data in two ways. First, there are cases in which a new municipality was created after 1970, which would have been part of another municipality in 1970. For example, municipality A

Following the procedures proposed above and excluding 37 municipalities from the sample due either to a lack of information about the population, or GDP, or the origin of some municipalities that existed in 1996, the total number of municipalities used in the estimation is $3,781$.

Therefore, the database used in this empirical analysis is composed of the GDP per capita of the $3,781$ municipalities in Brazil. We evaluate whether convergence occurs across Brazilian municipalities by directly examining the cross-section distribution of income per capita over the period from 1970 to 1996, as suggested by Quah (1993, 1997).

Table 6: Database - Municipalities

State/ Region	Existing municip. 1996	Municip. arising from more than one municipality	Municip. arising from only one municipality	Municip. without origin	Municip. without GDP data	Municip. existing in 1970	Municip. without pop. data in 1970	Total agreggated municip.	Total municip. used in the estimations
PR	371	7	76	0	0	288	1	8	279
SC	259	2	59	0	0	198	0	5	193
RS	427	37	158	0	0	232	0	33	193
South	1057	47	292	0	0	718	1	46	671
ES	72	1	17	0	1	53	0	0	53
MG	755	1	31	0	0	723	0	10	713
RJ	81	0	18	0	0	63	0	0	63
SP	635	2	53	0	9	571	0	7	564
Southeast	1543	4	199	0	10	1410	0	17	1393
GO	232	5	51	8	0	168	0	5	163
MT	115	32	47	2	0	34	1	10	23
MS	77	9	18	0	0	50	0	11	39
DF	1	0	0	0	0	1	0	0	1
Center	425	46	116	10	0	253	1	26	226
MA	136	2	4	0	0	130	0	4	126
PI	148	4	30	1	0	113	0	7	106
CE	184	1	42	0	0	141	0	1	140
RN	152	1	1	0	0	150	1	3	146
PB	168	0	0	0	0	168	0	0	168
PE	177	2	9	1	0	165	0	7	158
AL	100	4	1	0	0	95	0	8	87
SE	75	0	1	0	0	74	0	1	73
BA	415	7	72	2	0	334	0	7	327
Northeast	1555	21	160	4	0	1370	1	38	1331
AC	22	4	11	0	0	7	0	2	5
AP	15	4	6	0	0	5	0	1	4
AM	62	10	8	0	0	44	0	12	32
PA	128	10	35	0	0	83	0	12	71
RO	50	8	32	8	0	2	0	1	1
RR	8	2	4	0	0	2	0	1	1
TO	123	7	62	2	0	52	0	6	46
North	408	45	158	10	0	195	0	35	160
TOTAL	4998	163	845	24	10	3946	3	162	3781