Insper | IBMEC São Paulo

Testing Convergence Across Municipalities in Brazil Using Quantile Regression

Eduardo Andrade

Márcio Laurini

Regina Madalozzo

Pedro L. Valls Pereira

Insper Working Paper

WPE: 034/2002

Testing Convergence Across Municipalities in Brazil Using Quantile Regression*

Eduardo Andrade	Márcio Laurini	Regina Madalozzo	Pedro Valls
Ibmec/SP	Ibmec/SP	Ibmec/SP	Ibmec/SP

First Version: June 27th, 2002

This Version: January 27th, 2003

Abstract

This paper compares the linear regression (OLS) and quantile regression approaches as ways to test the convergence hypothesis for Brazilian municipalities in the period from 1970 and 1996. The quantile approach not only circumvents some pitfalls of OLS, but is also a stronger test of convergence, and allows some quantiles to present convergence while others diverge. When controlling for regional differences, results from OLS and quantile regression are not significantly different, except for a few quantiles in the North and Northeast regions. The results suggest that the convergence hypothesis passes the stronger test imposed by the quantile regression. Depending on the region, GDP per capita is converging to the steady state at a rate from 0.39% to 3.64% per year.

Key words: convergence, quantile regression, stochastic kernel.

JEL Classification: C14, C23, R12, R13

*Address: Ibmec Business School, Rua Maestro Cardim 1170, 11o. andar, 01323-001, São Paulo, Brazil; Phone: (55)(11)3175-2311; Fax: (55)(11)3175-2315; E-mail: (eduardo.andrade@ibmec.br).

The authors would like to thank Patrícia Palomo and Samuel Kinoshita for excellent research assistance. The last author would like to acknowledge the partial financial support from CNPq and PRONEX.

1 Introduction

There are many empirical studies that try to test for income convergence across economies using different data sets.[1] One of their objectives is to evaluate whether Solow's prediction, that a poor economy grows at a faster rate than a rich one, can be somehow empirically confirmed.

In order to test Solow's prediction, many of these studies run an OLS regression in which the dependent variable is the average growth rate over some period of time and the explanatory variable is the initial income per capita. The idea is that a negative coefficient in this regression indicates that economies with lower initial levels of income per capita grow at a higher rate in subsequent years.

However, there are some important technical problems already discussed in the literature with this classical approach. First, it assumes that the estimated coefficient in the OLS regression is the same for all economies. This does not allow for the possibility that the impact of the initial income per capita on growth in the subsequent periods will differ across economies.[2] Second, it ignores the problem known in the literature as "Galton's Fallacy." As pointed out by Friedman (1992) and Quah (1993), a negative coefficient in the traditional OLS regression may not indicate that economies are converging to the same long-run steady state, but intead can signal regression to the mean. Finally, the presence of outliers and heterocedasticity can be a problem, which can bias the coefficient estimated from the OLS regression.

In this paper, we propose a different methodology to test convergence across economies, one which can solve the technical problems of the OLS regression mentioned above. The use of quantile regression is the main contribution of this study. This approach not only circumvents the pitfalls of the former one, but is also a stronger test of convergence, as it allows some quantiles to present convergence while others diverge.

Our objective here is to test whether there is convergence across the almost 4000 Brazilian municipalities in the period from 1970 to 1996. We compare the results obtained using the traditional OLS growth regression with those obtained employing quantile regression. Initially, we perform this comparison using the whole sample, which allows us to analyze if all municipalities in Brazil are converging to the same long-run steady state. Later, we break the sample into five groups. Each one is

[1] See Sala-i-Martin (1996) for a survey of these studies.

[2] See Bernard and Durlauf (1996).

formed of municipalities in each of the five regions in Brazil: south, southeast, center, northeast, and north. We compare the results obtained for each group using both methodologies.

To our knowledge, Mello and Novo (2002) is the only other study that uses quantile regression to test convergence across economies. They use, however, a different sample formed by 98 countries, known as Barro and Lee's data set.

The rest of this paper is organized as follows. The problems in using OLS to infer income convergence across economies are reviewed in section 2. Then, in section 3, we present the quantile regression methodology and explain how it can correct the problems with OLS growth regressions mentioned in the previous section. In section 4, we begin by briefly explaining the data set used in this empirical work, and test whether there is income convergence across Brazilian municipalities using quantile regression. We also compare the rate of convergence obtained in this study with other results where the traditional OLS approach is used. Conclusions are drawn in the last section.

2 Problems with OLS Growth Regressions

The traditional approach to growth, based on Solow's model, suggests that a poor economy tends to grow faster than a rich one. The explanation behind this result is that, given the same technological and behavioral parameters, the marginal productivity of capital is greater in the poor locations, as capital there is relatively scarcer. As a consequence, income or output per capita convergence across countries would occur in the long run.

The empirical growth literature has tested Solow's implication in the following way. It considers that the average growth rate over the interval from 0 to T is given by:

$$\left(\frac{1}{T}\right)\log(\frac{y_{iT}}{y_{i0}}) = \alpha + \beta\log(y_{i0}) + u_{i0,T}, \tag{1}$$

where α and β are constants, y_{iT} and y_{i0} are the income per capita in country or region i at time T and 0, respectively, and $u_{i0,T}$ represents the average of the error terms, u_{it}, between dates 0 and T. If $\beta < 0$, then the lower the initial income per capita, the greater the average growth rate in the years ahead. It is said that the data set exhibits absolute β-convergence if the β estimated through OLS is negative.

There are many empirical studies that follow this methodology to test the β-convergence hypothesis using different data sets. For regional data sets, such as Japanese prefectures, regions in Brazil, Italy, Spain, the United States, and areas of other countries, there is strong empirical support for the idea of β-convergence.[3][4] However, when countries' data set are used, the estimated coefficient for β is not necessarily negative, indicating nonexistence of β-convergence across countries.[5]

These different results for regional and non-regional data sets are to some degree expected. Regions inside a country are more likely to have the same technological and behavioral parameters, and thus have the same long-run steady state. The prediction in Solow's model that poor economies will eventually catch up with rich ones holds true only if they have the same long-run steady state. An example can illustrate this point. Suppose that richer economies have a greater income per capita steady-state equilibrium in comparison with poorer ones. Under this assumption, a negative estimated coefficient for β can occur if the poorer economies start far below their steady-states, whereas the richer ones start at values near their steady-states. As the long-run steady states are different, the notion of convergence in which "the limit of the long-run forecasts of output differences tending to zero as the forecasting horizon increases" does not hold even with a negative $\widehat{\beta}$.[6]

In order to deal with the possibility of multiple long-run steady states, the empirical growth literature uses another concept of convergence, the conditional β-convergence.[7] The idea is to introduce variables that are proxies for the steady-state in the regression based in equation (1). Hence, instead of running a regression using (1), one estimates:

$$\left(\frac{1}{T}\right)\log\left(\frac{y_{iT}}{y_{i0}}\right) = \alpha + \beta\log(y_{i0}) + \Psi X_{i0} + u_{i0,T}, \tag{2}$$

where $X_{i,t}$ is a vector of variables that hold constant the steady state of economy i. If $\widehat{\beta} < 0$ once $X_{i,t}$ is held constant, it is said that the data set exhibits conditional β-convergence. Once the vector

[3] For a summary of many of these results using regional data sets, see Sala-i-Martin (1996). For the case of Brazil, see Ferreira (1996) and Schwarstman (1996).

[4] Most of the empirical work uses the following non-linear version of equation (1): $\left(\frac{1}{T}\right)\log\left(\frac{y_{iT}}{y_{i0}}\right) = a - \left(\frac{1-e^{-\beta T}}{T}\right)\log(y_{i0}) + u_{i0,T}$. This non-linear equation comes from the log-linearized version of the equilibrium conditions of the Solow-Swan model. Barro and Sala-i-Martin (1995) show how to derive this equation.

[5] See also Sala-i-Martin (1996).

[6] See Bernard and Durlauf (1996) for an extensive discussion on this topic.

[7] Another concept of convergence widely used in the traditional literature on empirical growth is σ-convergence. According to this concept, a group of economies are converging if the dispersion of their income per capita tends to decrease over time. As this approach is not the focus of our analysis, we do not discuss it.

$X_{i.t}$ is added, the estimated coefficient for β in the regression using the countries' data set becomes negative, suggesting the existence of conditional β-convergence.[8]

However, there are other important criticisms of this traditional approach to testing the convergence hypothesis empirically which cannot be easily repaired by OLS. The new literature on empirical growth has presented three major objections to this traditional methodology.

First, as discussed in Bernard and Durlauf (1996), the estimated coefficient for β using OLS is equal to a weighted average of the ratio of differences of growth rates to differences of initial incomes, both from the sample mean. In order to obtain convergence, it is necessary that a weighted average of countries or regions with above average initial incomes grow at a slower rate than the mean growth for the cross-section. As it is an average, it is theoretically possible that some countries or regions are not converging to a common long-run steady state and that $\widehat{\beta}$ is negative. As pointed out by Bernard and Durlauf (1996), "the cross-section tests cannot identify groupings of countries which are converging ...[and] is thus ill-designed to analyze data where some countries are converging and others are not." In other words, it cannot identify the existence or lack of convergence clubs.[9]

Traditional OLS growth regressions establish not only the magnitude of the effects of the initial income per capita to be the same for all economies, but also the coefficients of the variables X_{i0}. Using equation (2), the estimated coefficients for Ψ are valid for all countries or regions. This assumption, however, seems to be unawarranted. It is very unlikely, for example, that the impact of a policy variable such as human capital on economic growth would be the same for all economies irrespective of their level of development. Mello and Novo (2002) discuss this point at great length and, in particular, they test using quantile regression whether policy variables affect the mean and the dispersion of the conditional distribution of GDP growth rates for the Barro and Lee's data set.

Second, Friedman (1992) and Quah (1993) argue that estimated coefficients for β in the traditional OLS growth regressions do not shed any light on whether the poorer economies are converging to the richer ones, that is, on whether convergence occurs or not. To make this association is to commit the problem known as Galton's classical fallacy. According to Friedman and Quah, a negative relationship between the average growth rate and the initial income may reflect regression to the mean and not convergence. In other words, "a negative correlation does not, in fact, imply a collapsing of the

[8] See Sala-i-Martin (1996).

[9] See Quah (2000) for a theoretical and empirical discussion of the possibility of emerging convergence clubs.

cross-section distribution." For example, Quah (1993) takes a non-collapsing invariant cross-section distribution and regresses the growth rates on an initial condition. He shows that the coefficient on the initial condition is always no greater than zero. Thus, a negative sign on the initial condition coefficient does not indicate a collapsing cross-section distribution. More generally, Quah (1993) shows that a cross-section distribution can diverge even when the initial conditions regression shows a negative correlation between time-averaged growth rates and the initial levels.[10]

Finally, the presence of outliers and heteroscedasticity can be a problem. With regard to the former problem of outliers, there is a risk of eliminating part of the sample. The use of a non-random sample can produce biased estimation. In fact,to restrict the sample is widely used in the classical empirical literature on growth in order to check whether there is conditional β-convergence. The idea is to restrict the sample to economies with the same long-run steady state. With regard to the latter problem, heteroscedasticity can arise due to the different economic behaviors and characteristics across the economies that form a data set used in the OLS traditional regressions. The use of dummy variables for different regions is a traditional approach to correct this problem.

The above discussion pinpoints some technical problems with the classical approach in the literature that is used to test empirically the existence or lack of convergence across economies. This discussion shows the need to employ another approach to circumvent the pitfalls of the traditional methodology. The next section deals with a possible alternative approach.

3 Quantile Regression

The use of quantile regression methodology to test convergence across economies is the main distinction of this paper. The OLS approach is easily computable and, by this reason, very appealing. Its results represent a conditional mean estimation. However, ordinary least squares do not perform adequately when the sample does not follow a normal distribution and presents heteroscedastic behavior with notable outliers. Koenker and Bassett (1978) introduced quantile regression as the estimation of the conditional quantile function as one possible solution to these problems.[11] This innovative approach brings not only more explanatory power to the results when compared to the details captured

[10] Bernard and Durlauf (1996) present complementary results to those derived by Quah (1993).

[11] See also Koenker and Portnoy (1996), Buchinsky (1999), and Koenker and Hallock (2001).

by the least squares approach, but also decreases the influence of outliers in the estimations.

The parallel between least squares and quantile regression approaches can be seen by comparing their objective functions. The least squares approach solves the minimization problem:

$$\min_{\beta \in R^p} \sum_{i=1}^{n} (y_i - x_i'\beta)^2$$

This equation results in the conditional mean function $E(Y|x)$. Quantile regression uses a similar procedure, directing its attention to the analogous optimization problem:

$$\min_{\beta \in R^p} \sum_{i=1}^{n} \rho_\tau (y_i - x_i'\beta(\tau)) \tag{3}$$

where ρ represents a "loss function" that can be calculated conditioned to each selected quantile τ, where $\tau \in (0,1)$. This loss function is defined as $\rho_\tau(u) = u(\tau - I(u < 0))$, where $I(.)$ is an indicator function, and u represents the difference between the actual value and the estimated one for each observation. Clearly, these differences may take positive or negative values. The indicator function assumes value 1 when the u is negative, and zero otherwise. As a result, the loss function will assume a negative value for all these observations.[12] This characteristic of the function $\rho_\tau(u)$ allows both the weighting and the optimization of the differences between the estimated and actual values of observations for each quantile.[13]

Consequently, the minimization problem expressed on the equation (3) provides solutions, $\widehat{\beta}(\tau)$, to the expectations of $\rho_\tau(y_i - x_i'\beta(\tau))$ with respect to each $\beta(\tau)$. Quantile regression, using linear programming methods, estimates the conditional median function, as well as all the possible choices of other conditional quantile functions. Compared to that, an OLS model restricts itself to estimations of the conditional mean function.

Quantile regression's characteristics are especially interesting in the present study. In the previous section, three specific problems with OLS regressions were mentioned, and quantile regression features answers to each one of them.

The main criticism offered by Bernard and Durlauf (1996) concerns the estimation of an equal

[12] This occurs because τ is between 0 and 1, while the indicator function assumes value 1 when u is negative.

[13] When using OLS regression, the squared value of the difference between the actual and the estimated value for each observation makes its optimization problem possible.

convergence rate for all countries or regions. This is a main attribute of ordinary least squares regression. Using quantile regression, however, instead of having an unique estimated coefficient, the conditional mean, a family of curves will be available for interpretation. Each estimated quantile focuses its attention on a particular segment of the conditional distribution, resulting in a broader description of the relationship between the variables. In a growth regression, these different estimated coefficients for β represent distinct rates of convergence. For each quantile, a specific rate of convergence will be available.

A second failure of OLS regression in characterizing the growth equation is related to Galton's fallacy, which points out the estimation of an expected conditional parameter as one of the caveats of a linear regression approach.[14] As Friedman (1992) and Quah (1993) point out, the estimation of a negative coefficient for the initial GDP per capita in the growth regression does not necessarily imply that the convergence hypothesis is fulfilled. This may happen because the linear regression estimation corresponds to a mean of the observations' distribution. Unlike OLS, quantile regression estimation provides a wider view of the observations distribution. As a result, if the observations contained in any quantile do not converge, its estimated coefficient for the initial GDP per capita would be non-negative. This characteristic of quantile regression means that it not only allows different coefficient estimates for each chosen quantile, but also allows the estimates to indicate convergence or divergence for each one of them.

Another problem that usually affects growth regressions in a linear regression is the presence of outliers and the heteroscedastic distribution of the observations. The quantile regression approach is known for its low sensitivity to outlier observations. In the linear squares regression, the failure of the normality assumption, especially with outliers that result in a long-tail distribution, results in poor estimates of the parameters. Quantile regression estimations, imposing different weights on observations according to the quantile to be estimated, are robust even for cases with a distribution far from Gaussian.

Given these specific characteristics of quantile regression, the estimation of a growth regression using this approach is developed in the following section. The usual difficulties faced in least squares estimation are avoided with this different methodology.

[14]Koenker (2000) describes the Galton fallacy and the past attempts to solve it.

4 Empirical Results

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

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 was created after 1970 while in 1970 it was a part of municipality B. In 1970, the GDP per capita of B included also that of A, as A did not exist at that time. In 1996, the data shows the GDP and the population of both municipalities separately, A and B. However, we do not have this same separated data for 1970. There are 845 cases such as this one. In such cases, we consider the two municipalities as only one. Second, there are other cases in which a new municipality, created after 1970, was formed by aggregating parts from different municipalities that existed in 1970. For example, municipality X was created after 1970 formed by parts of municipalities Y and Z that did exist in 1970. The latter municipalities continued to exist with the same name afterwards but they became smaller. There are 163 cases of this nature. It is not possible to identify what part of the GDP of the new municipality is related to each municipality that it came from. As a result, we treat municipalities X, Y, and Z as forming one unique municipality, and called it W. Therefore, municipality W is equal to the sum of municipalities Y and Z in 1970 and to the sum of municipalities X, Y, and Z in 1996, and we calculate the income per capita accordingly.

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 $3,781$.[15]

We now turn to the empirical analysis. First, equation (1) was estimated by OLS using all the

[15] The table in the appendix provides the number of municipalities in each state and region. It also includes the summary of all these adjustments implemented in the raw data.

3781 observations and the following results were obtained:

$$\left(\frac{1}{T}\right)\log(\frac{\widehat{y_{iT}}}{y_{i0}}) = \underset{(0.0029)}{0.0819} - \underset{(0.00042)}{0.00756}\log(y_{i0}) \tag{4}$$
$$R^2 = 0.07728 \quad F(1,3918) = 328.1[0.000]$$

where the standard error for the estimates of the coefficients are in parentheses and for the F statistics the p-value is reported. Both coefficients are significant and the $\widehat{\beta}$ is negative, indicating that there is convergence across Brazilian municipalities over the period covered in the analysis. However, without controlling for regional differences, the above result can be biased.[16]

Therefore, the next step is to add dummy variables in the OLSregression (1) to check if the estimated parameter changes significantly. We allow for different intercept and slope for the following regions in Brazil: North, Northeast, South, and Center.[17] Figure 1 reports the estimate of $\widehat{\beta}$ with a 95% confidence interval.[18] The coefficient estimated for β in regression (4) is reproduced in the first column in figure 1. Note that the coefficients for all regions are negative and they are significantly different from those obtained in regression (4).[19] The smaller new parameter in absolute terms estimated for β is obtained for the Southeast region, the most developed one, and it is almost two times greater than the previous estimate for β. These results suggest the importance of controlling for regional differences in order to obtain unbiased estimates and a greater speed of convergence.[20]

The next step is to estimate (1) for Brazil using quantile regression. Figure 2 shows the estimate of $\widehat{\beta}$ for several quantiles[21] and also the OLS estimate. Some of these estimates are significantly different from OLS.

Table 1 presents the estimates of $\widehat{\beta}$, its standard error and 95% confidence interval for quantile $10\%, 25\%, 50\%, 75\%$ and 90% as well as the OLS estimate. Note that it is between the quantiles

[16]See discussion in chapter 11 in Barro and Sala-i-Martin (1995).

[17]This excludes the dummy for the Southeast region.

[18]Notation used in the figure is LB for lower bound and UB for upper bound

[19]The new estimates are also significantly different from zero. The results can be obtained from the authors upon request.

[20]It will be seen at the end of this section that the magnitude of these new estimates (and as a consequence the speed of convergence) are more similar to the ones obtained in other related papers.

[21]In all figures that show the quantile regression results, we include estimates for quantiles from 10% to 90% with intervals of 5%.

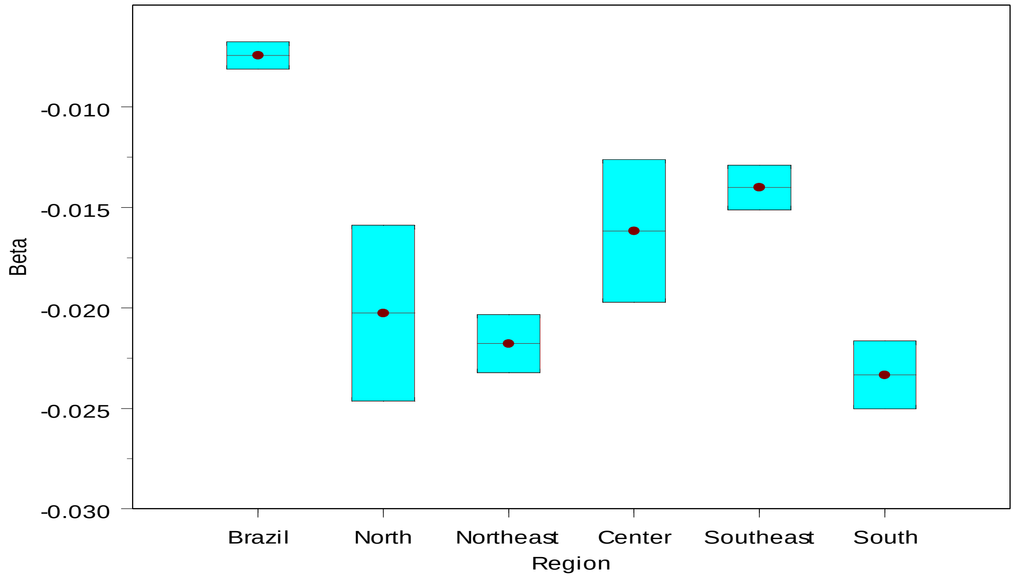

Figure 1: Beta Estimate with Confidence Interval for Brazil and Regions

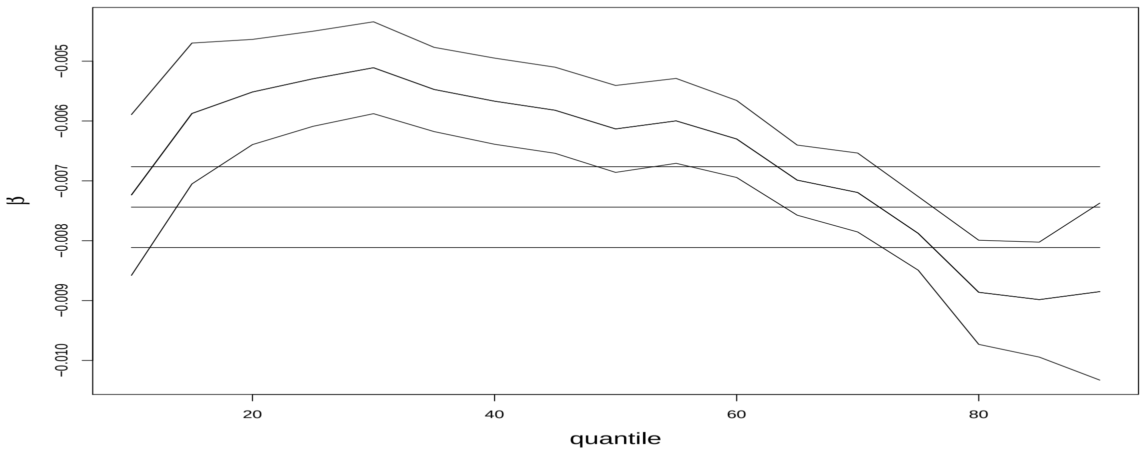

Figure 2: Quantile Regression - Brazil

25% and 50% that OLS estimates are significantly different from quantile estimates. For all the other quantiles, OLS is undistinguished from the quantile estimates. The results obtained using the quantile regression approach suggests two important implications. First, there is no indication that we can reject the hypothesis of convergence across Brazilian municipalities as the coefficients for all quantiles are significantly different from zero and negative. Second, the speed of convergence varies across municipalities. The quantile regression ordinates the municipalities by the dependent variable, the average growth rate in the period from 1970 to 1996, with those with higher growth rates belonging to the higher quantiles. Those municipalities in the quantiles 25% and 50% are converging at a slower rate, as the estimates for $\widehat{\beta}$ are lower in absolute terms.

As discussed before, the above estimates could be biased since we did not control for regional differences. To correct for this possible effect, we estimate (1) allowing for different intercept and slope for the regions in Brazil and using quantile regression. In fact, when we add regional dummy variables to control for this effect, the results change in two significant ways. First, as occurred when we used the OLS approach, the new estimates are very different from the ones obtained without the regional dummies, indicating a greater speed of convergence. Table 2 reports the new estimates and the old ones. The new $\widehat{\beta}$ are at least 2.63 times greater in absolute terms, suggesting a greater speed of convergence.[22] Second, the difference between the quantile estimate and OLS are now restricted to some quantiles for the North and Northeast regions. Figures 3 and 4 show the estimates for these regions. For all other regions there is no difference between OLS and quantile regression[23].

Table 3 presents the estimates of $\widehat{\beta}$ for North Region, its standard error and 95% confidence interval for quantile $10\%, 25\%, 50\%, 75\%$ and 90% as well as the OLS estimate. Note that for the 25% quantile the estimate is different from OLS.

Table 4 presents the estimates of β for Northeast Region, its standard error and 95% confidence interval for quantile $10\%, 25\%, 50\%, 75\%$ and 90% as well as the OLS estimate. The quantiles 10% and 90% present quantile estimates different from OLS.

One possible explanation for the difference between OLS estimates and the quantile estimates

[22] See in the end of this section a comparison of the magnitude of these new $\widehat{\beta}$ with the ones obtained in related papers.

[23] The results can be obtained from the authors upon request.

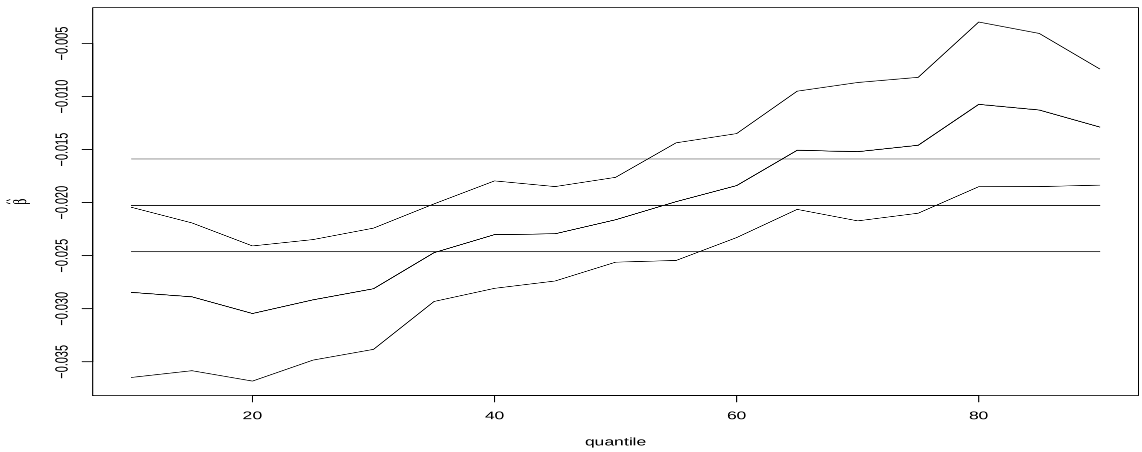

Figure 3: Quantile Regression for North Region

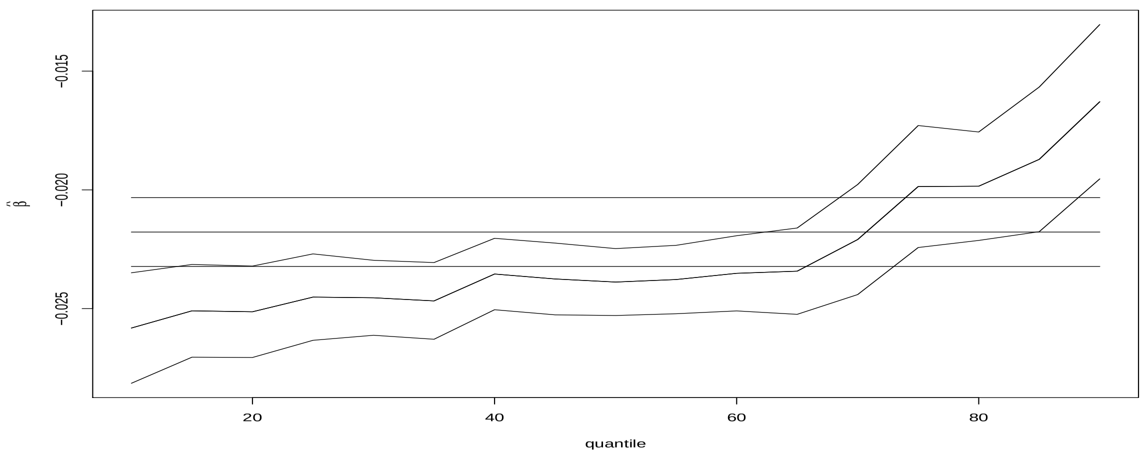

Figure 4: Quantile Regression for Northeast Region

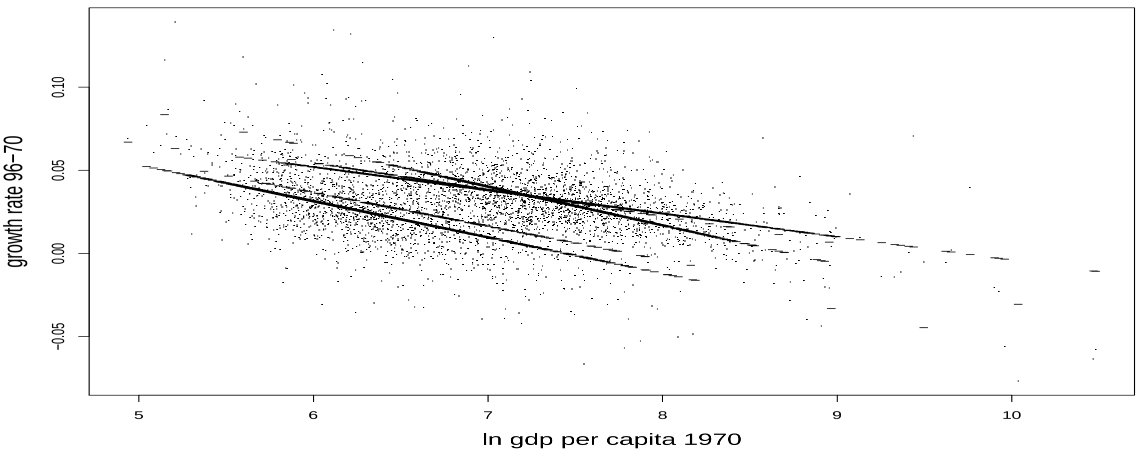

Figure 5: Fitted Values for OLS estimation allowing for regional effects

for Brazil and the quantile estimates when the regional effect are present could be the possibility of non-constant variance between quantiles. However, Bartlett's test for homogeneity of variances gives the result of 0.1376 with a p-value of 0.9999 which implies non-rejection of the null of homogeneity.

The empirical evidence, since all the estimates of β in all cases are negative, is in favor of convergence. Figures 5 through 7 illustrate the same result from a different perspective. Figure 5 gives a scatter diagram of the log of GDP per capita in 1970 versus the growth rate from 1970 to 1996. The lines represent the fitted values for OLS estimation allowing for region effects. Note that all slopes are negative which indicates those municipalities with lower initial levels of GDP per capital have grown faster. This is precisely the implication of Solow's model. The results obtained in the exercise above suggest that the hypothesis of convergence across Brazilian municipalities has passed the stronger test imposed by the quantile regression approach.When quantile regression is used we have the same picture. Figure 6 presents the 90% quantile allowing for regional effect and Figure 7 presents the 10% quantile also allowing for regional effect.

Figure 8 shows the half-life per region, that is, the number of years necessary to reduce by half the income differences across the municipalities. As the differences between the estimates for the quantile and OLS regressions are restricted to a few quantiles of the North and Northeast regions, we do not report the half-life per quantile. The half-lifes estimated range from 29 years for the South region to 49 years for the Southeast region. As the Southeast region is the more developed, it is expected that

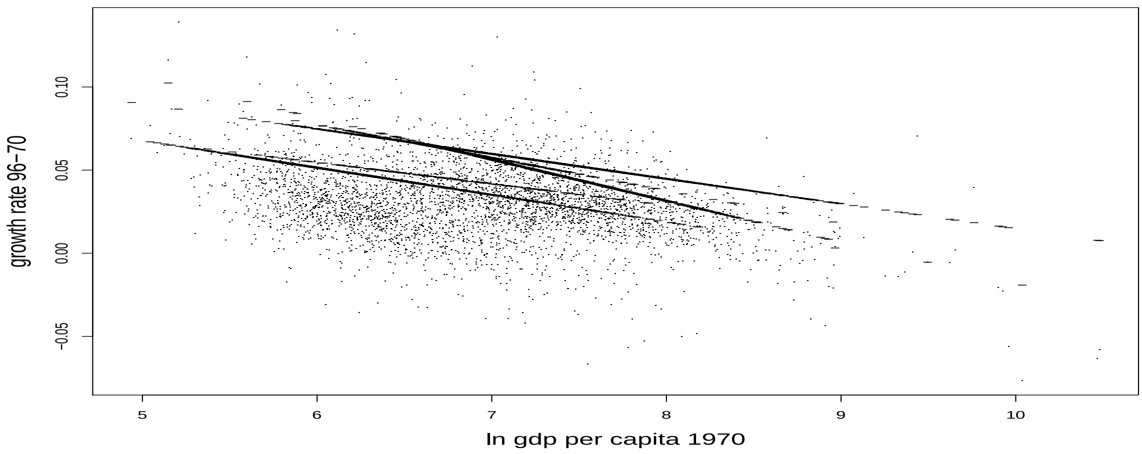

Figure 6: Fitted Values for 90 % Quantile

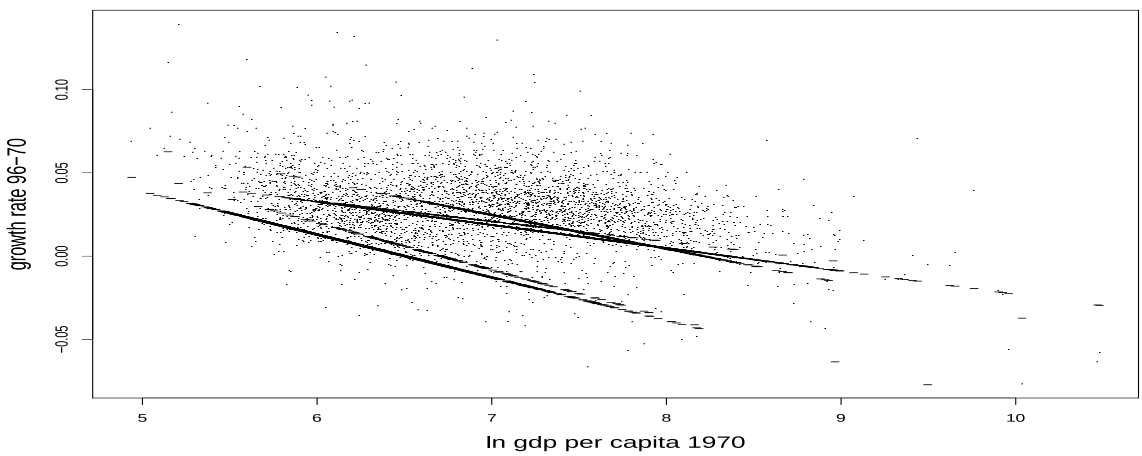

Figure 7: Fitted Values for 10 % Quantile

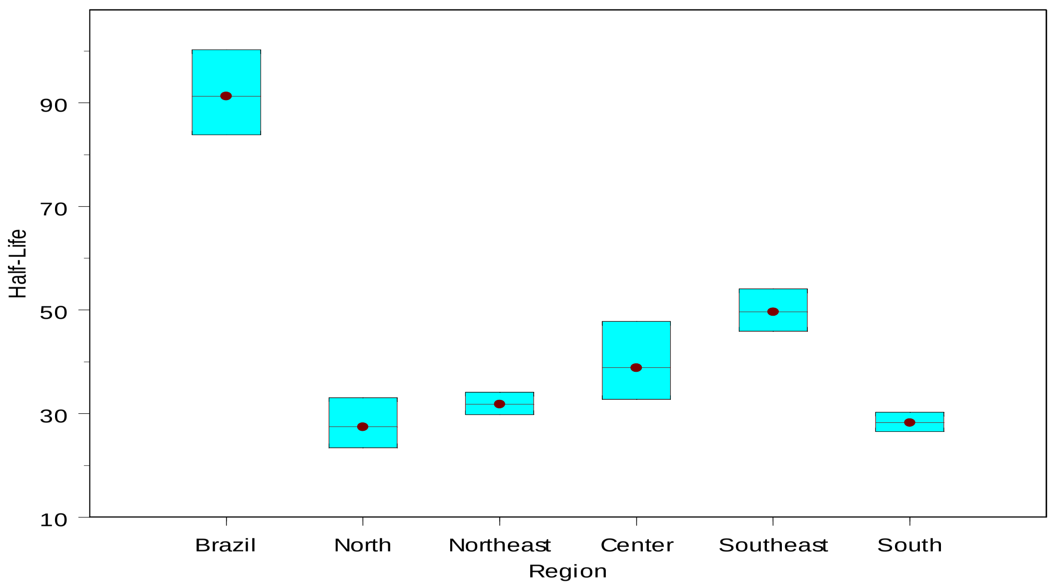

Figure 8: Half-Lifes

the reduction in income disparities across its municipalities will take a longer time. A surprising result is that high half-life for the South is lower than the ones observed for the Northeast and the Center regions, which are less developed than the South region.

Finally, it is interesting to compare the above results with others obtained in related papers. Ferreira and Ellery (1996) use data for the Brazilian states for the period from 1970 to 1990. They obtain a slower speed of convergence with $\widehat{\beta}$ around 0.014, which corresponds to a half-life of around 50 years. This difference may be related not only to the fact that we use a different time period as well as the fact that we use a finer division of the Brazilian space, municipalities versus states. Our overall result is close to the one obtained for the Japanese prefectures (β around 0.028 and half-life of 24.4 years) and are not very different from the ones found for U.S. states (β around 0.02 and half-life of 32 years).

5 Conclusion

This paper presents a new approach, the quantile regression method, to test the convergence hypothesis across regions in a country. We argue that the quantile approach represents an improvement with respect to the traditional approach used in the growth literature, the OLS regression. The reason is that it circumvents three important pitfalls of the OLS method. First, it solves the problem known in the literature as "Galton's fallacy." Second, it presents no biased estimates based on the presence of outliers and heteroscedasticity. Finally, it does not assume that the speed of convergence is the same for all economies. This attribute permits some quantiles to present convergence while others diverge. As a consequence of this feature, it presents a more rigorous test for the convergence hypothesis.

We use the quantile approach to test the convergence hypothesis for Brazilian municipalities in the period from 1970 to 1996. Our results are different from the OLS estimations, if the region indicators are not included. However, when the region dummies are added to the estimation, results from OLS and quantile regression are not significantly different, with a few exceptions for some quantiles in the North and Northeast regions. Therefore, we conclude that the convergence hypothesis across Brazilian municipalites passes the stronger test imposed by the quantile approach. That is, the conclusion in favor of convergence still remains.

Depending on the region, the GDP per capita is converging to the steady state at a rate from 0.39% to 3.64% per year. As a result, the half-lifes estimated range from 29 years for the South region to 49 years for the Southeast region. These results suggest a faster speed of convergence in comparison with earlier work on the Brazilian states. This difference may be due to the fact that we use a finer division of the Brazilian space, municipalities versus states, and a longer period of analysis. In terms of speed of convergence, our results are not very different from the ones obtained in other studies for other countries, such as Japan and the United States.

Besides the relevance of the present results' improvement over a simple linear regression approach, some critiques remain unanswered. Quah (1993a) points out that regressions of a mean growth rate in respect to initial income are not adequate to the complete understanding of income distribution for distinct economies. Aleatory shocks may modify the transition dynamic and the relative distribution would change from time to time. The key to this problem would be the dynamic analysis of the whole distribution, which will be a future development of this study.

6 References

Barro, R. and Sala-i-Martin, X. 1995. Economic Growth. McGraw-Hill, New York.

Buchinsky, M. 1999. Recent Advances in Quantile Regression Models: a Practical Guideline for Empirical Research. Journal of Human Resources, 3(1), 88-126.

Bernard, A. and Durlauf, S. 1996. Interpreting Tests of the Convergence Hypothesis. Journal of Econometrics 71, 161-173.

Ferreira, P. and Ellery Jr., R. 1996. Convergência entre a Renda Per-Capita dos Estados Brasileiros. Revista de Econometria, 16(1), 88-103.

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

Koenker, R. 2000. Galton, Edgeworth, Frisch, and Prospects for Quantile Regression in Econometrics. Journal of Econometrics 95, 347-374.

Koenker, R. and Bassett, G. 1978. Regression Quantiles. Econometrica, 46, 33-50.

Koenker, R. and Hallock, K. 2001. Quantile Regression: an Introduction. Department of Economics, University of Illinois at Urbana-Champain. (forthcoming in the Journal of Economic Perspectives "Symposium on Econometric Tools").

Koenker, R. and Portney, S. 1996. Quantile Regression. Working paper, University of Illinois at Urbana-Champaign, ORWP #97-0100.

Mello, M. and Novo, A. 2002. The New Empirics of Economic Growth: Quantile Regression Estimation and Inference of Growth Equations. Manuscript.

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

Quah, D. 1993a. Empirical Cross-section Dynamics in Economic Growth. LSE Working paper, December.

Quah, D. 2000. Cross-Country Growth Comparison: Theory to Empirics. LSE Working paper, January.

Sala-i-Martin, X. 1996. The Classical Approach to Convergence Analysis. The Economic Journal 106, 1019-1036.

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

7 Tables

Brazil	10%	25%	50%	75%	90%	*OLS*
Coeff	−0.00712	−0.00547	−0.00619	−0.00813	−0.00942	−0.00756
S.E.	0.000825	0.000493	0.000448	0.000413	0.000907	0.000417
LB	−0.00845	**−0.00626**	**−0.00691**	−0.00880	−0.01088	−0.00823
UB	−0.00579	**−0.00467**	**−0.00547**	−0.00746	−0.00796	−0.00689

Table 1 - Beta Estimates for Quantile and OLS

Brazil	10%	25%	50%	75%	90%	*OLS*
Coeff	−0,02042	−0,02143	−0,02189	−0,02291	−0,02486	−0,02333
SE	−0,018240	−0,019761	−0,020275	−0,021175	−0,02189	−0,021229
UB	−0,02261	−0,0231	−0,02351	−0,02465	−0,02784	−0,02543
LB	0,001358	0,001040	0,001006	0,001078	0,001846	0,001304

Table 2 - Quantile and OLS Estimates for Brazil with Regional Dummies

North	10%	25%	50%	75%	90%	*OLS*
Coeff	−0.0297	−0.0351	−0.0287	−0.0224	−0.0172	−0.0249
S.E.	0.00527	0.00340	0.00272	0.00391	0.00544	0.00261
LB	−0.0382	**−0.0406**	−0.0331	−0.0287	−0.0260	−0.0291
UB	−0.0212	**−0.0296**	−0.0244	−0.0161	−0.00849	−0.0207

Table 3 - Quantile Estimates for North Region

Northeast	10%	25%	50%	75%	90%	*OLS*
Coeff	−0.0257	−0.0245	−0.0237	−0.0199	−0.0161	−0.0215
S.E.	0.00141	0.00113	0.000850	0.00151	0.00221	0.000892
LB	**−0.0280**	−0.0263	−0.0250	−0.0223	**−0.0197**	−0.0230
UB	**−0.0234**	−0.0227	−0.0223	−0.0174	**−0.0126**	−0.0201

Table 4 - Quantile Estimates for Northeast Region

	Brazil	*North*	*Northeast*	*Center*	*Southeast*	*South*
half-life	91.31	27.44	31.83	38.89	49.67	28.26
UB	100.26	33.08	34.13	47.83	54.09	30.27
LB	83.82	23.43	29.82	32.74	45.91	26.50

Table 5 - Half-Life Estimates for Brazil and Regions

State/Region	Municipalities existing in 1996	Municipalities proceeding of more than one municipality	Municipalities proceeding of only one municipality	Municipalities without origin	Municipalities without GDP data[24]	Municipalities existing in 1970	Municipalities without population data in 1970	Total of grouped munipalities	Total municipalities used in 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

Table 6 : Database

[24]These municipalities were in spreadsheets divulged by IPEA, however did not contain data of the GDP