# Empirical Evidence on Convergence across Brazilian States

Luiz Renato Lima

Hilton Hostalácio Notini

Fábio Augusto Reis Gomes Insper Instituto de Ensino e Pesquisa

#### Abstract

This paper is aimed to analyze the convergence hypothesis across Brazilian States in a 60 year period (1947-2006). In order to test the existence of income convergence, the order of integration of the income differences between each State and São Paulo is examined. São Paulo is the richest State and for this reason is used as a benchmark. First of all, we employed the conventional unit root tests, finding evidence against the convergence hypothesis. However, given the lack of power of unit root tests, especially when the convergence is very low, we used ARFIMA models, which is also theoretically more appropriate [Michelacci and Zaffaroni (2000)]. Although the ARFIMA model present a point estimate lower than 1 for many States, we cannot reject the unit value, in general. Therefore, there is a lack of convergence.

JEL Codes: C22, O49, O54, R11;

Keywords: Growth Model; Stochastic Convergence; Long Memory; Brazil.

## 1 Introduction

According to the neoclassical growth model, economies converge to their own steady state, regardless of its initial per capita output. Furthermore, the speed of convergence is inversely related to the gap between current income and its steady state value (Solow, 1956). If the parameters that characterize the steady states of a group of countries are identical, then their difference lies in their initial level of capital and poor economies should grow faster because they are further away from the steady state. This proposition is named absolute convergence. On the other hand, one can allow for heterogeneity across economies by dropping the assumption that all economies have the same steady state. In this case, economies are convergent only after controlling for their steady states. This proposition is known as conditional convergence, and it gave rise to an immense empirical literature that analyzes its suitability. Initially, the literature focused on countries comparisons using cross-section data; however, some authors also explored time series dimension.

In the cross-section approach, a negative correlation between income growth rates and initial income is interpreted as evidence of unconditional  $\beta$ -convergence, where unconditional means that countries' characteristics are not taken into account. The conditional  $\beta$ -convergence analysis requires the inclusion of conditioning variables, like investment rate and population growth, in order to interpret the convergence rate as a measure of conditional convergence (Durlauf, 2001). In this context, one of the most generally accepted results is that, while there is evidence of unconditional convergence inside homogenous group of countries or regions, only the conditional convergence hypothesis holds when examining broad groups (Barro, 1991; Barro and Sala-i-Martin, 1991, 1992; Mankiw et al., 1992; Sala-i-Martin, 1996).

The time series approach investigates convergence looking for common long-run terms in countries' output. Durlauf and Bernard (1995) argued that two countries are convergent if the difference of their outputs is a mean-zero stationary process. Durlauf and Bernard (1996) relax this condition, stating that it is enough if countries' incomes have the same long-run forecast. In general, the works based on time series approach do not find evidence of convergence since they find that the output differences between pairs of countries contain a unit root (Durlauf, 1995, 1996; Campbell and Mankiw, 1989; Carlino and Mills, 1993).

While the conditional  $\beta$ -convergence means that aggregate shocks are absorbed at a uniform exponential rate (mean reversion), a unit root in output differences implies no mean

reversion. Are these results incompatible? As noticed by Mello and Guimaraes-Filho (2007), the speed of convergence found by cross section works is very low, which combined with the low power of unit root tests might explain the conflicting results.<sup>1</sup> Assuming a AR(1) model, the speed of convergence of 2% implies a coefficient around 0.98, which illustrates the issue. Probably, the unit root tests do not have enough power to reject the unit root null hypothesis when the coefficient is so close to one.

Mello and Guimaraes-Filho (2007) applied Auto-Regressive-Fractionally-Integrated-Moving-Average (AFIRMA) models to OECD economies, finding ample evidence of convergence.<sup>2</sup> While the unit root tests investigate if the order of integration, d, is equal to one or zero, the ARFIMA models estimated d allowing non-integer values and, this flexibility helps to increase the power of the test. Indeed, for certain values of d, income differential process can be nonstationary but mean-reverting which means that shocks are persistent but eventually die out. This property explains both: ARFIMA models are able to capture the observed slow speed of income convergence and unit root tests do not reject nonstationarity (unit root).

In addition, Mello and Guimaraes-Filho (2007) showed that if the difference of two economies' output is a ARFIMA(0,d,0) with d in the interval (-1/2,1), then countries' incomes have the same long-run forecast, being convergent. Therefore, the ARFIMA model offers a direct way to test convergence and reduces the lack of power of usual unit root tests. It is worth mentioning that Michelacci and Zaffaroni (2000) showed that the Solow growth model based on aggregation of cross-sectional heterogeneous unit can lead to a long memory process even if the aggregating elements are stationary with probability one.

If the lack of power of unit root test is solved, should we be indifferent between crosssection and time series approaches? Given the statement of Michelacci and Zaffaroni (2000), the answer is no. In any case, Bernard and Durlauf (1995) pointed out some stringent assumptions of the cross-section approach. First, the null hypothesis assumes that all countries are converging and at the same rate. Second, as mentioned, it is necessary to include conditioning variables to implement a conditional convergence test.

As argued by Durlauf (2001), the standard growth theory does not imply identical parameters across economies, and we should expect some kind of heterogeneity. On this matter, remember the results in favor of convergence groups.<sup>3</sup> In addition, Maddala and Wu (2000) used

<sup>&</sup>lt;sup>1</sup> See Campbell and Perron (1991) for a discussion about the lack of power of unit root tests.

<sup>&</sup>lt;sup>2</sup> Indeed, Quah (1995) noted that cross-section and time series analysis cannot arrive at different conclusions.

<sup>&</sup>lt;sup>3</sup> Durlauf and Johnson (1995) find evidence in favor of convergence clubs applying the time series approach.

a data set similar to that of Mankiw et al (1992), finding a faster convergence rate after allowing for distinct countries to have differing convergence rates.<sup>4</sup> Therefore, the analysis of cross-sectional data is not suited for cases where some countries are converging (at different rates) and others not. On the other hand, the time series analysis is not subject to these drawbacks because the convergence analysis is implemented for pairs of countries. Finally, about the conditioning variables, Durlauf (2001) stressed that there is no consensus about which variables should be included and some variables like saving rates and human capital bring together the ignored endogeneity problem.<sup>5</sup>

This paper is aimed to test the convergence hypothesis among states in Brazil and, for the reasons mentioned, the ARFIMA models are used. Many authors studied the convergence across Brazilian states using the cross-sectional approach and their results are in favor of  $\beta$ -convergence (Ferreira and Diniz (1995), Schwartsman (1996), Ferreira and Ellery (1996), Ferreira (2000) and Carvalho and Santos (2007)). The exception is Azzoni and Barossi-Filho (2002) which apply the time series approach to test the occurrence of stochastic convergence among Brazilian states. The authors used unit root tests, finding signs of convergence among the majority of states.

This paper is organized as follows. This introduction is followed by Section 2, where we briefly review the ARFIMA models, the convergence hypothesis concepts and the relevant literature. Section 3 shows our econometric approach while Section 4 presents the results. Finally, the last section summarize our main results.

## 2. ARFIMA Models and the Convergence Hypothesis

### 2.1 ARFIMA Models

This section presents a brief review of ARFIMA models. A more complete discussion about fractional integration process can be found in Baillie (1996).

Fractional integrated models were introduced in the economics literature by Granger (1980) and Granger and Joyeux (1980). They were theoretically justified, in terms of aggregation of ARMA process with randomly varying coefficients by Robinson (1978) and Granger (1980).

<sup>&</sup>lt;sup>4</sup> In fact, Maddala and Wu (2000) suggest that heterogeneity cause bias in the usual Barro-type regressions.

<sup>&</sup>lt;sup>5</sup> See Temple (1999) for a detailed discussion about this subject.

These models belong to a broad class of long range dependence process, also known as long memory process.

The presence of long memory can be defined in terms of the persistence of observed autocorrelations from an empirical data oriented approach. When viewed as time series realization of a stochastic process, the autocorrelation function exhibits persistence that is neither consistent with a covariance stationary process nor a unit root process. The extent of the persistence is consistent with an essentially stationary process, but the autocorrelations take far longer to decay than the exponential rate associated with the ARMA class. More specifically, an ARFIMA process has an autocorrelation function given by  $\rho_x(\tau) \sim \tau^{2d-1}$ , for large  $\tau$ , while a geometrically decaying function stationary а ARMA process has given by  $\rho_x(\tau) \sim r^{\tau}$  where |r| < 1. Thus, the importance of this class of process derives from smoothly bridging the gap between short memory I(0) process and I(1) process in an environment which maintains a greater degree of continuity.

In order to understand the idea of fractionally integrated processes, it is helpful to start with the stochastic process below:

$$(1-L)^{d}x_{t} = v_{t}, \quad t = 1, 2, 3, \dots$$
 (1)

where L is the lag operator,  $x_t$  is a discrete time scalar time series, t=1,2,...,  $v_t$  is a zero-mean constant variance and serially uncorrelated error term and *d* denotes the fractional differencing parameter which is allowed to assume non-integer values. The process in (1) is called ARFIMA(0,*d*,0) model. If *d*=0, then  $x_t$  is a standard or better short memory stationary process whereas  $x_t$  is a random walk if d=1. For values d>-1, the term  $(1-L)^d$  has a binomial expansion given by  $(1-L)^d = 1 - dL + d(d-1)L^2/2! - d(d-1)(d-2)L^3/3! + ....$  Invertibility is obtained whenever -1/2 < d < 1/2. Thus, the process in (1) is stationary if parameter *d* lies in the interval (-1/2, 1/2). In the interval (1/2, 1), the process is non-stationary, but exhibits mean reversion. In summary, if the fractional differencing parameter is less than the unity, income shocks die out, even when the  $x_t$  process is non-stationary. Therefore, the parameter *d* plays a crucial role in describing the persistence in the time series behavior: higher the *d*, higher will be the level of association between the observations.

Alternatively, we can define long range dependence in terms of the spectral density function. Assuming that K denotes any positive constant and  $\sim$  denotes asymptotic equivalence we have:

**Definition 1**: A real valued scalar discrete time process  $X_t$  is said to exhibit long memory in terms of the power spectrum (when it exists) with parameter d > 0 if:

$$f(\lambda) \sim K\lambda^{-2d}$$
 as  $\lambda \to 0^+$  (2)

In a non-stationary case  $(d \ge 1/2)$ ,  $f(\lambda)$  is not integrable and thus it is defined as a pseudospectrum. When  $v_t$  is assumed to be a white noise process, the process  $x_t$  defined in equation (2) is called an ARFIMA(0,d,0) process and when  $v_t$  is an (inverted) ARMA (p,q) we obtain an ARFIMA(p,d,q) process. The power spectrum of the  $x_t$  process is given by

$$f_{x}(\lambda) = |1 - e^{i\lambda}|^{-2d} f_{y}(\lambda) = (2\sin(\lambda/2))^{-2d} f_{y}(\lambda), \qquad -\pi \le \pi \le \pi$$
(3)

where  $f_v(\cdot)$  denotes the power spectrum of the  $v_t$  process. Thus from  $\sin(\varpi)/\varpi \sim 1$  as  $\varpi \to 0$ , when d = 0 as  $\lambda \to 0^+$  we have:

$$f_x(\lambda) \sim 4^{-d} f_v(0) \lambda^{-2d}.$$
(4)

For an ARFIMA process, the parameter *d* controls the low-frequency series behavior. In particular, the spectral density function of long range dependence process behaves like  $\lambda^{-2d}$  as  $\lambda \to 0$ , while in the traditional ARIMA model it is constrained to behave like  $\lambda^{-2}$  as  $\lambda \to 0$ . Whenever d > 0, the power spectrum is unbounded at the zero frequency, which implies that the series  $x_t$  exhibits long memory. When 0 < d < 1/2,  $x_t$  has both finite variance and mean reversion. When  $\frac{1}{2} < d < 1$ , it has infinite variance but it still shows mean reversion. When  $d \ge 1$ , the process has infinite variance and stops exhibiting mean reversion. When d = 1, there is a unit root process.

Last, it is important to mention that usually only the second moment properties are considered in order to characterize such behavior in terms of either the autocorrelation function at long lags, or the power spectrum near the zero frequency.

#### 2.2. Convergence Hypotyhesis

Based on Bernard and Durlauf (1991), we present the stochastic convergence definition.

**Definition 1 [Stochastic Convergence in per capita income]**: The logarithm of income per capita for economies *i* and *j*, denoted by  $y_{i,t}$  and  $y_{j,t}$ , respectively, is said to converge in a time series sense if their difference is a stationary stochastic process with zero mean and constant variance. That is, if  $y_{i,t} - y_{j,t} = \varepsilon_{ij,t} \approx I(0)$ , where  $\varepsilon_{ij,t} \sim (0, \sigma_{\varepsilon}^2)$ , then economies *i* and *j* converge in a time series sense.

Given this definition, we can test for pairwise convergence hypothesis by means of cointegration tests. Indeed, as the cointegration vector is [1,-1], we can apply unit root tests for income differences. However, Michelacci and Zaffaroni (2000) showed theoretically that the Solow growth model based on aggregation of cross-sectional heterogeneous unit can lead to a long memory process even if the aggregating elements are stationary with probability one. In this sense, conventional cointegration and unit root tests are misspecified.

For instance, Gonzalo and Lee (2000) showed that the Johansen's (1988) cointegration test tends to produce too much spurious cointegration relationships if the individual series are fractionally integrated. In addition, a speed of convergence about 2% is compatible with a AR(1) model with coefficient around 0,98 and a ARFIMA(0,d,0) model with *d* in the interval (0,5;1); in both cases the lower power of unit root tests explains the non-convergence result from the time series approach.

If the appropriated model is ARFIMA(0,*d*,0) with *d* in the interval (0,5;1) rather than a stationary ARMA, as claimed by Michelacci and Zaffaroni (2000), the absorption of shocks is hyperbolic (slow) rather than exponential (fast). Nevertheless, the  $\beta$ -convergence would apply in the sense that poorer economies would grow faster and converge toward their long-run steady state, which explain the results from the cross-section approach. Using a weaker convergence

criterion proposed by Bernard and Durlauf (1996) - reproduced below as definition 2 - Mello and Guimaraes-Filho (2007) formalize this explanation.

**Definition 2**: Convergence as equality of long-run forecasts at a fixed time. The logarithm of income per capita for countries *i* and *j*, denoted by  $y_{i,t}$  and  $y_{j,t}$ , respectively, is said to converge in time series sense if their long-run forecast of the log of income per capita for both countries are equal at a fixed time *t*. This condition can be written as  $\lim E(y_{i,t+k} - y_{j,t+k} | \mathfrak{T}_t) = 0$ , where  $\mathfrak{T}_t$  is the information set at time *t*.

Mello and Guimaraes-Filho (2007) assumed that  $y_{i,t} - y_{j,t}$  can be described by an ARFIMA(0,*d*,0) and then they showed that the time series convergence criterion 2 is satisfied whenever estimates of the parameter *d* lie on the interval (-1/2 ,1). Therefore, a direct test of convergence can be made using the ARFIMA models, taking into account any heterogeneity in the Solow-Swan model and overcoming unit root lack of power.

### 2.3 Literature Review

The same techniques applied to study convergence across countries can be used to study convergence across states or regions in a country. In this case the benchmark should be the richest state or region into the country. There are a lot of papers which study convergence across Brazilian states. Most of them used the cross-sectional approach like Ferreira and Diniz (1995), Schwartsman (1996), Ferreira and Ellery (1996) and Ferreira (2000). In general, they verified the existence of absolute convergence among Brazilian states for the period 1970 to 1985. In addition, in all cases there is presence of conditional  $\beta$ -convergence. As stressed by Azzoni (1997), one problem with these studies is the short time period of analysis which could avoid the right conclusion about the convergence hypothesis.

Indeed, Carvalho and Santos (2007) tested the absolute convergence hypothesis for the period 1980-2002. They found a very weak occurrence of convergence based on a -0.0062 significant beta coefficient and a speed of convergence of 0.7% which in turn generated a half-life of 103 years. According to these estimates, the Brazilian States would take 103 years to reduce the disparities between them. Thus, one contribution of this paper is to shed light to this discussion by increasing the time span of analysis to 60 years (four times bigger than the great

majority of Brazilian convergence studies) what would allow a better description of the convergence hypothesis given its long-term nature.

Azzoni *et al* (2000) and Menezes and Azzoni (2000) used panel data, finding support to conditional β-convergence. Only Azzoni and Barossi-Filho (2002) applied the time series approach to study convergence in Brazil. In order to test for the existence of stochastic convergence among Brazilian states they use data on per capita income for 20 states covering the period 1947-1998<sup>6</sup>. The authors used unit root tests which endogenously determine structural break points. They find that 14 out of 20 states analyzed present signs of convergence (AL, BA, CE, MA, MT, MG, PB, PR, RN, RS, RJ, SE), 3 states show weak convergence (ES, GO, PE) and 5 have no sign of convergence (AM, PA, PI, SC, SP)<sup>7</sup>.

However, Azzoni and Barossi-Filho's (2002) methodology has some caveats. First, the authors used as benchmark the average income rather than the richest State. Second, if determinist components like a broken trend are allowed, then the output difference is neither a zero-mean process nor its long-run forecast is zero.<sup>8</sup> In other words, the Definitions 1 and 2 are not attended. Last, Montañés *et al.* (2005) pointed out that unit root tests based on intervention analysis are very sensitive to the specification of the alternative model.

This paper is aimed to analyze the convergence hypothesis among the Brazilian States. To overcome the problems extensively discussed, we use the ARFIMA models as suggested by Michelacci and Zaffaroni (2000) and Mello and Guimaraes-Filho (2007). Especial attention is given to the presence of deterministic terms, once the time series definition of convergence is violated by any predictable long-term in output differences.

## 3. Econometric Methodoloy

### 3.1. Estimation and Inference in ARFIMA models

<sup>&</sup>lt;sup>6</sup> They maintain the original administrative organization of the country as in 1947 so the states that were created during the period considered have been added to the states that were originated from.

<sup>&</sup>lt;sup>7</sup> Table 4 in appendix contains a complete description of the states name and abbreviation.

<sup>&</sup>lt;sup>8</sup> For instance, Bernard and Durlauf (1995), Li and Papell (1999) and Attfield (2003) argue that the rejection of a unit root null hypothesis is a necessary but not sufficient condition to guarantee conditional convergence, in their point of view it is also necessary that the log of relative outputs are zero mean.

There are a lot of techniques available to estimate AFIRMA models. We estimate the model using the Nonlinear Least Squares Method (NLS) - sometimes referred to as the Approximate Maximum Likelihood Method. The NLS estimator is based on the maximization of the following likelihood function:

$$\ell_N(d,\Phi,\Theta) = -\frac{1}{2}\log\left(\frac{1}{T}\sum_{i=1}^N \tilde{e}_i\right) \tag{1}$$

where the residuals  $\tilde{e}_t$  are obtained by applying the ARFIMA(p,d,q) to  $u_t$  and the vectors  $\Phi$  and  $\Theta$  represent the p autoregressive and the q moving-average parameters, respectively.<sup>9</sup> In our case, p and q are equal to zero.

#### 3.2 Data

Our data set consists of annual log real Gross State Product per capita (GSP). The data range from 1947 to 2006 for twenty Brazilian states, namely, Alagoas (AL), Amazonas (AM), Bahia (BA), Ceará (CE), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rio de Janeiro (RJ), Santa Catarina (SC), São Paulo (SP) and Sergipe (SE). Although actually Brazil has 27 states, we work with a set of only 20 because 7 states did not exist in 1947. Thus, in order to maintain the original administrative organization the states created during the period analyzed have been added to the original states.

The GSP data have been obtained from Azzoni (1997), the population data and the annual price index (Índice Geral de Preços – Disponibilidade Interna, IGP-DI) have been obtained from Instituto Brasileiro de Geografía e Estatística (IBGE). We take SP State as our benchmark because it is the richest state in the country. Then, we analyze the evolution of  $\ln(Y_{sp,t}/Y_{i,t})$ , where *i* means the other 19 States.

Graph 1 displays the box plot of the ratio between SP and the other 10 richest States, ie,  $Y_{sp,t}/Y_{i,t}$ . Values greater than 1 (horizontal line), means that SP is richer than the other State under analysis. Even among the richest States there is a large difference in relation to SP. Looking at average values, apart from PR, RJ RS and SC, SP is at least two times the other States. In fact, only RJ was in some moment richer than SP. The dispersion inside each State is

<sup>&</sup>lt;sup>9</sup> The econometric package used for the estimations is Doornik & Ooms' (2001) OxMetrics and the numerical method used to maximize the likelihood function is BFGS.

not constant. For instance, while PB has a high dispersion, the RJ is very concentrated. Graph 2 displays the box plot of the ratio between SP and the other 9 poorest States. The difference between SP and these States is huge. On average, SP is around 8 times richer than PI, for instance. Among the poorest States, the dispersion is also far from constant.

[Graph 1]

#### [Graph 2]

## 5. Results

To have a kind of benchmark of the cross-section approach, we estimated the usual regression to test unconditional convergence:

$$\ln Y_{i,T} - \ln Y_{i,0} = a + b \ln Y_{i,0} + \varepsilon_{i,T}$$

where  $Y_{i,t}$  is the income level of State *i* at period *t*. This regression represents the core of crosssection approach and a negative value for *b* implies convergence. Graph 3 displays the result when the OLS method is employed. The value of *b* is estimated around -0,04, being significant at 5% level. However, the shaded area in the Graph 3 highlighted a sub-group of States with similar initial income and very distinct growth rate, which seems to be inconsistent with  $\beta$ -convergence hypothesis. Two of them present an impressive growth rate, GO (2.56%) and PB (2.23%). Looking at the initial level of income, there are also two remarkable States, SP (18.51) and RJ (20.90). We run the regression again, eliminating these four States. The results are displayed in Graph 4. The value of *b* approaches zero and, in fact, becomes not significant. We re-estimate the model using all states by means of LAD estimator, which is robust to outliers, and the coefficient *b* becomes insignificant at 5% level.

[Graph 3]

[Graph 4]

Although these analyses should be view with caution, given the low number of crosssection units and the absence of conditioning variables, at least we can safely say that (unconditional) convergence is not an obvious result.

Now we turn to time series approach, beginning our analysis with the unit root tests. We applied the augmented Dickey-Fuller (ADF) and Phillips-Perron (PP). Two specifications were considered, the first contains a constant, while the second does not have any deterministic component. If the unit root null hypothesis is rejected in specification 1, it does not mean that States are convergent, once in the long-run they deviate by a constant. The specification 2 does not have a constant, constituting a properly test of convergence. The results are displayed in Table 1. Using the 5% level of significance, the ADF and PP tests do not reject the null hypothesis for any State and specification. At 10% level of significance, the ADF test is able to reject the null hypothesis only for AL in specification 1 while the PP test reject unit root only for GO in specification 2. Therefore, there is strong evidence against convergence, even when we allow the States to deviate from SP by a constant term. Given the result of the cross-section approach, even if the convergence took place in this period, the convergence rate would be very low, impeding the unit root tests to reject a false unit root null hypothesis.

#### [Table 1]

Finally, we employ the ARFIMA models to estimate de decay rate d. Following the strategy of the unit root tests, we consider two specifications. The first contain a constant and the second does not have any deterministic term. The results are displayed in Table 2. First of all, it is worth mentioning that the procedure to estimate d for RJ did not converge when the specification 1 was employed. Allowing the constant term, in the fashion of Azzoni and Barossi-Filho (2002), 16 States present a point estimated of d in the interval (-1/2,1). For 5 (11) States we are able to reject the null hypothesis d=1 at 5% (10%) level. Thus, as the constant term is frequently significant, it seems that most States deviates from SP by a constant value. Looking at the results of specification 2, the number of States with point estimate of d lowers than 1, reduced to 10. From this subgroup only MT and PR rejected the null hypothesis d=1, still for a 10% significance level. Then, after all, there is a lack of convergence across Brazilian States.

## Conclusion

In this paper we test the real convergence hypothesis across Brazilian States using unit root tests and ARFIMA models. In particular, we have examined the order of integration of the annual log real per capita Gross State Product (GSP) series for twenty States, taking SP as a benchmark. Our results suggested a lack of convergence.

It is important to say that our time series results are very different from the works of Ferreira and Diniz (1995), Schwartsman (1996), Ellery and Ferreira (1996) and Ferreira (2000) which uses a cross-sectional approach to study convergence across Brazilian states. While these studies show that the States are convergent, we found the opposite result. One possible explanation is the broader time period used in here than the cross-section studies (which usually cover the period 1970-1995). Besides, it is relevant to stress that our analysis brings much more information about income convergence as we test the convergence hypothesis for each state separately. In a cross-sectional approach instead it is not possible to conclude which state is converging to the benchmark state and which are not, so we cannot point out which States deserve more government police to reduce interstates income disparities.

Our results are not directly compared with the ones from Azzoni and Barossi-Filho (2002) because the two works used different benchmarks. While we adopt São Paulo as our benchmark for the convergence analysis, Azzoni and Barossi-Filho (2002) adopt the whole country as their benchmark. It is important to mention that the use of long memory models and unit root tests with endogenously structural break points share the same purpose, which is to avoid the ADF tests low power problem. However, as mentioned the unit root tests with endogenously structural break points are very sensitive to the specification of the alternative model (Montañés et al., 2005). Besides, this model does not impose that the income differences will have a mean zero in the long-run. Considering these drawbacks in the intervention analysis, and the fact that long memory processes are theoretically justified in terms of aggregation of units with different speed of adjustment in the Solow-Swan model, we can say that our methodology and consequently our results are more trustworthy than Azzoni and Barossi-Filho (2002).

Other issues such as potential presence of structural breaks on the data and the effect that it may has on the results can be studied in future papers. Other consideration is studying convergence across Brazilian states in specific regions like North, South, and so on, each one with its own benchmark state that could be an aggregation of all constituents of the region.

## References

- Attfield, C.L.F. 2003. Structural Breaks and Convergence in Output Growth in the EU. Discussion Paper No 03/544 Department of Economics, University of Bristol.
- Azonni, C., 1997. Concentração regional e dispersão das rendas per capita estaduais: análise a partir de séries históricas estaduais de PIB. 1939-1995.
- Azonni, C., 2001. Economic Growth and Regional Income Inequalities in Brazil. Annals of Regional Science, Vol 35, No.1.
- Azzoni and Barossi-Filho, 2002. A Time Series Analysis of Regional Income Convergence in Brazil. In: XXX Encontro Nacional de Economia da ANPEC, Nova Friburgo. Anais do XXX Encontro Nacional de Economia da ANPEC, 2002.
- Azonni, C., Menezes-Filho, N., Menezes, T. and Silveira-Neto, R., 2000. Geography and regional income inequality in Brazil. Inter American Development Bank, Working paper
- Baillie, R., T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5-59.
- Barro, R., 1991. Economic Growth in a cross-section of countries. Quarterly Journal of Economics, CVI, 407-55.
- Barro, R. and Sala-i-Martin, X., 1992. Convegence. Journal of Political Economy 100, pp.223-51.
- Barro, R. and Sala-i-Martin, X., 1995. Economic Growth. McGraw-Hill, Inc.
- Bernard, A. B. and Durlauf, S. N. 1995. Convergence in International Output, Journal of Applied Econometrics, 10, pp.97-108.
- Bernard, A. B. and Durlauf, S. N. 1996. Interpreting tests of the convergence hypothesis. Journal of Econometrics 71, 161-173.
- Campbell, J. Y., and Mankiw, N. G., 1987. Are output fluctuations transitory?. Quarterly Journal of Economics 4, 857-880.
- Campbell, J. Y., and Mankiw, N. G., 1989. International evidence on the persistence of economic fluctuations. Journal of Monetary Economics, 23, 319-333.

- Campbell, J. Y. and Perron, P., 1991. Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots. NBER Macroeconomic Annual, p. 141-201.
- Caporale, G.M. and Gil-Alana, L.A., 2002. Fractional cointegration and real exchange rates, The Review of Financial Economics 13, 327-340.
- Carlino, G. A. and Mills, L. O., 1993. Are U.S regional incomes converging? A time series Analysis. Journal of Monetary Economics, 32, 335-346.
- Cochrane, J., 1988. How big is the random walk in GDP. Journal of Political Economy 96, 893-920.
- Cunado, J., Gil-Alana, L. A., Pérez de Gracia, F., 2003. Empirical evidence on real convergence in some OECD countries Applied Economics. 10, 173-76.
- Durlauf, S. N. (2001) Manifesto for a growth econometrics. Journal of Econometrics, 100:1, 65-69.
- Durlauf, S. and Johnson, P. (1995) Multiple regimes and cross country growth behaviour, Journal of Applied Econometrics, 10, 365-84.
- Ellery Jr., R. and Ferreira, P. C., 1996. Crescimento Econômico e Convergência entre a Renda dos Estados Brasileiros. Revista de Econometria, V.16 (1), pp.83-104.
- Engle, R.F. and Granger, C.W.J. 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica 55, 2, 251-276.
- Ferreira, A.H.B., 2000. Convergence in Brazil: Recent Trends and long Run Prospects. Applied Economics, 32,pp.479-489.
- Ferreira, A. H. B. and Diniz, C. C., 1994. Convergência Entre as Rendas Per Capita Estaduais no Brasil. Texto para Discussão n 79, CEDEPLAR/UFMG, Belo Horizonte.
- Geweke, P. and Porter-Hudak, S., 1983. The estimation and application of long memory time series models. Journal of Time Series Analysis, 4, 221-38.
- Gil-Alana, L. A., 2000. Mean reversion in the real exchange rates, Economics Letters, 16, 285-8.
- Gil-Alana, L. A., 2002. Comparisons between semiparametric procedures for estimating the fractional differencing parameter, Preprint.
- Gil-Alana, L. A., 2003. A Univariate Analysis of Unemployment and Inflation in Italy: A Fractionally Integrated Approach, Brazilian Review of Econometrics, 23, n 2, 227-54.
- Gil-Alana, L. A and Robinson, P. M., 1997. Testing of unit roots and other nonstationary hypotheses in macroeconomic time series, Journal of Econometrics, 80, 241-68.
- Granger, C., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227-38.

- Granger, C. and Joyeux, R., 1980. An introduction to long memory time series models and fractional differencing Journal of Time Series Analysis 1, 15-39.
- Jones, C., 1995. Time series tests of endogenous growth models. Quarterly Journal of Economics 110, 495-525.
- Jones, C., 1997. Convergence Revisited. Journal of Economic Growth, 2, pp. 131-153.
- Kunsch, H., 1986. Discrimination between monotonic trends and long range dependence. Journal of applied Probability 23, 1025-30.
- Li, Q. and Papell, D. 1999. Convergence of international output. Time series evidence for 16 OECD countries. International Review of Economics and Finance, 8, pp.267-280.
- Lobato, I., 1999. A semiparametric two-step estimator for a multivariate long memory process. Journal of Econometrics, 73, 303-24.
- Maddala, G. and Wu, S. (2000) Cross-country growth regressions: problems of heterogeneity, stability and interpretation, Applied Economics, 32, 635-642.
- Maia Gomes, 2002. Regional Development Strategies in Brazil. mimeo.
- Mankiw, G., Romer, P. and Weil, D. N., 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics, 107, 407-37.
- Mello, M. and Perreli R, 2003. Growth equations: a quantile regression exploration. The Quarterly Review of Economics and Finance 43, pp643-667.
- Menezes, T. and Azzoni, C., 2000. Convergência de renda real e nominal entre regiões metropolitanas brasileiras: uma análise de dados de painel", XXVIII Encontro da Anpec, Campinas, 2000.
- Michelacci, C.; and Zaffaroni, P., 2000. Fractional beta convergence. Journal of Monetary Economics 45, pp. 129-153.
- Montañés, A., Olloqui, I., Calvo, E., 2005. Selection of the break in the Perron-type tests. Journal of Econometrics, 129, 41-64.
- Nelson, C., Plosser, C., 1992. Trends and random walks in a macroeconomic time series: some evidence and implications. Journal of Monetary Economics 10, 139-62.
- Perron, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. Econometrica, 57, 1361-1401.
- Quah, D. 1993., Empirical cross-section dynamics in economic growth. European Economic Review 37 (2/3), 426-34.
- Quah, D. 1995, Empirics for economic growth and convergence. European Economic Review 40, n6, 1353-75.

- Robinson, P. M., 1978. Statistical inference for a random coefficient autoregressive model. Scandanavian Journal of Statistics 5, 163-8.
- Robinson, P. M., 1994. Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association, 89, 1420-37.
- Robinson, P. M., 1995a. Log-periodogram regression of time series with long range dependence. Annals of Statistics, 23, 1048-72.
- Robinson, P. M., 1995b. Gaussian semiparametric estimation of long range dependence. Annals of Statistics, 23, 1630-61.
- Robinson, P.M. and Yajima, Y., 2002. Determination of cointegrating rank in fractional systems, Journal of Econometrics 106, 217-241.
- Sala-i-Martin (1996) The Classical Approach to Convergence Analysis, The Economic Journal, 106, 1019-1036.
- Santos, C. and Carvalho, F. (2007) Dinâmica das disparidades regionais da renda per capita nos estados brasileiros: uma análise de convergência. Revista Economia e Desenvolvimento, 19, 78-91.
- Schwartsman, A., 1996. Convergence across Brazilian States. Discussion Paper, n 02/96. IPE, Universidade de São Paulo, 1996.
- Silverberg,G. and Verspagen,B., 2000. A Note on Michelacci and Zaffaroni, Long Memory, and Time Series of Economic Growth. Research Memoranda 031, Maastricht : MERIT, Maastricht Economic Research Institute on Innovation and Technology.



Graph 1 - Box Plot of the Ratio between SP and other States (richest)

Graph 2 - Box Plot of the Ratio between SP and other States (poorest)





Graph 3 - Unconditional Beta Convergence Analysis

Graph 4 - Unconditional Beta Convergence Analysis (sub-sample)



	Specification 1		Specification 2	
	ADF	PP	ADF	PP
State	(p-value)	(p-value)	(p-value)	(p-value)
AL	0.097	0.100	0.552	0.525
AM	0.597	0.601	0.426	0.448
BA	0.777	0.709	0.350	0.373
CE	0.613	0.380	0.367	0.414
ES	0.793	0.889	0.226	0.177
GO	0.716	0.872	0.136	0.067
MA	0.153	0.455	0.246	0.350
MG	0.853	0.766	0.234	0.274
MT	0.125	0.132	0.289	0.318
PA	0.236	0.208	0.554	0.551
PB	0.842	0.842	0.156	0.158
PE	0.285	0.232	0.639	0.639
PI	0.839	0.502	0.372	0.547
PR	0.491	0.645	0.418	0.403
RJ	0.143	0.132	0.270	0.247
RN	0.202	0.212	0.457	0.462
RS	0.509	0.615	0.378	0.402
SC	0.899	0.923	0.319	0.311
SE	0.589	0.589	0.339	0.342

 Table 1 - Results of Unit Root Tests - Sample: 1947-2006

Note: The ADF test used the Schwarz criterion to decide the number of lags of the dependent variable in the right side of the test equation. The PP test used the nucleus of Bartlett and the window of Newey-West. Specification 1 has a constant while the specification 2 does not have any deterministic component.

Table 2 - Results of ARFINIA(0,0,0) Model - Sample: 1947-2000							
	Specification 1			Specification 2			
State	Constant	d	H <sub>0</sub> : $d=1$	d	H <sub>0</sub> : $d=1$		
	(Std.Error)	(Std.Error)	(P-value)	(Std.Error)	(P-value)		
AL	1.627	0.727	-2.211	1.009	0.215		
	(0.208)	(0.123)	(0.031)	(0.043)	(0.830)		
AM	2.407	0.891	-0.986	1.056	0.639		
	(1.837)	(0.111)	(0.328)	(0.087)	(0.526)		
BA	0.557	1.092	0.810	1.059	1.639		
	(1.159)	(0.114)	(0.421)	(0.036)	(0.106)		
CE	1.707	0.713	-2.944	0.954	-0.879		
	(0.272)	(0.097)	(0.005)	(0.052)	(0.383)		
ES	0.825	0.833	-1.827	0.917	-1.281		
	(0.460)	(0.091)	(0.073)	(0.064)	(0.205)		
GO	1.018	0.833	-1.763	0.911	-1.340		
	(0.645)	(0.095)	(0.083)	(0.066)	(0.186)		
MA	2.024	0.822	-1.715	0.986	-0.358		
	(0.395)	(0.104)	(0.092)	(0.038)	(0.722)		
MG	1.085	0.835	-1.906	0.974	-0.497		
	(0.328)	(0.087)	(0.062)	(0.052)	(0.621)		
MT	0.836	0.713	-2.373	0.840	-1.860		
	(0.398)	(0.121)	(0.021)	(0.086)	(0.068)		
PA	2.615	0.914	-0.686	1.069	1.083		
	(2.372)	(0.126)	(0.495)	(0.064)	(0.283)		
PB	-2.421	1.025	0.252	1.057	1.029		
	(15.450)	(0.100)	(0.802)	(0.055)	(0.308)		
PE	1.721	0.919	-0.678	1.032	0.848		
	(1.024)	(0.119)	(0.500)	(0.038)	(0.400)		
PI	3.036	0.813	-1.984	1.098	2.040		
	(0.795)	(0.094)	(0.052)	(0.048)	(0.046)		
PR	0.640	0.742	-2.942	0.877	-1.887		
	(0.212)	(0.088)	(0.005)	(0.065)	(0.064)		
RJ	-	-	-	0.945	-0.440		
				(0.124)	(0.661)		
RN	1.608	0.797	-1.704	0.978	-0.340		
	(0.469)	(0.119)	(0.094)	(0.065)	(0.735)		
RS	0.671	0.785	-2.325	0.979	-0.272		
	(0.224)	(0.093)	(0.024)	(0.078)	(0.787)		
SC	2.222	0.943	-0.694	1.057	0.900		
	(2.496)	(0.083)	(0.491)	(0.063)	(0.372)		
SE	2.116	0.957	-0.359	1.011	0.199		
	(2.822)	(0.119)	(0.721)	(0.054)	(0.843)		

 Table 2 - Results of ARFIMA(0,d,0) Model - Sample: 1947-2006

Note: Specification 1 has a constant and specification 2 does not have deterministic components.