Empirical evidence on real convergence across Brazilian states

Hilton Hostalácio Notini* †

Luiz Renato de Oliveira Lima‡
Graduate School of Economics - Getúlio Vargas Foundation
Rio de Janeiro, Brazil

July 24, 2006

Abstract

This paper examines the real convergence hypothesis across Brazilian states. In order to test for the existence of income convergence the order of integration of real Gross State Product (GSP) per capita series is examined as well as their differences with respect to the São Paulo state which is used as a benchmark state. Both parametric and semiparametric methods are used and the results show that convergence is achieved in the cases of Alagoas, Amazonas, Bahia, Goiás, Mato Grosso, Minas Gerais, Pernambuco, Piauí, Rio Grande do Sul, Rio de Janeiro and Santa Catarina and convergence is weakly achieved in the cases of Ceará, Maranhao, Pará, Paraná and Sergipe .The states of Espírito Santo, Paraíba and Rio Grande do Norte show no convergence.

O artigo examina a hipótese de convergência real entre os estados brasileiros. Para testar a existência ou não da convergência da renda a ordem da integração da série do produto real bruto do estado per capita é examinada assim como suas diferenças com respeito ao estado de São Paulo que é usado como base. Foram utilizados métodos paramétricos e semiparametric e os resultados mostram que ocorre convergência nos estados: Alagoas, Amazonas, Baía, Goiás, Mato Grosso, Minas Gerais, Pernambuco, Piauí, Rio Grande do Sul, Rio de Janeiro e Santa Catarina e ocorre convergência fraca nos estados: Ceará, de Maranhão, Pará, Paraná e Sergipe. Nos estados Espírito Santo, Paraíba e Rio Grande do Norte não há convergência.

- JEL Classification:C22, O49, O54, R11;
- Keywords: Growth model; Stochastic Convergence; Long memory.

1 Introduction

A great number of studies have examined the real convergence hypothesis empirically with varying results. In some cases these differences is caused by the

*I wish to thank Luis Gil-Alana for very useful comments and help. I also wish to thank Fábio Gomes and Pedro Cavalcanti Ferreira for their comments.

†E-mail: hilton@fgvmail.br

‡E-mail: luizr@fgv.br

use of different methodologies and differences in the data sample.

As it is well known, the neoclassical model predicts under some assumptions that the per capita output in an economy will converge to each country's steady-state (conditional convergence) or to a common steady-state (unconditional convergence), regardless of its initial per capita output level. Empirical testing of convergence hypothesis provides several definitions of convergence and thus, different methodologies to test it.

In a cross-section approach a negative correlation between growth rates and initial income is interpreted as evidence of unconditional β-convergence. In this context, one of the most generally accepted results is that while there is no evidence of unconditional convergence among a broad sample of countries, the conditional convergence hypothesis hold when examining more homogenous group of countries or regions. Examples in this context are Barro (1991), Barro and Sala-i-Martin (1991, 1992) and Mankiw et al. (1992). Another solution is to condition for additional explanatory variables like Mankiw et al. (1992).

Other branch of convergence studies uses the time series approach, in this case a stochastic convergence asks whether permanent movements in one country's per capita output are associated with permanent movements in another countries' output, that is, it examines, whether common stochastic elements matter, and how much persistent the differences among countries are. These earlier tests were initially proposed by Bernard and Durlauf (1995, 1996), who used the ideas of unit root and cointegration to test stochastic convergence. Since the seminal work of Nelson and Plosser (1982), the literature has noted how standard unit root tests have failed to reject the null of a unit root in output per capita. The problem with the unit root approach to test convergence is that unit root test are plagued by the well known low power problem of ADF tests as shown by Campbell and Perron (1991). A unit root in output implies that shocks are permanent so that output does not exhibit mean reversion. But conditional β-convergence means that aggregate shocks are absorbed at an uniform exponential rate.

More recently, Jones (1995), has observed that in line with the standard exogenous growth Solow model, the trend of output per capita for OECD economies is fairly smooth over time and does not exhibit any persistent changes in the post-second World War era.

These three stylized facts listed above seem to be inconsistent. On the one hand, a unit root in output implies that shocks are permanent so that output does not exhibit mean reversion. On the other hand, β-convergence implies that output converges to its steady-state level at a rate that, even if very low, is positive and uniform across economies. The Jones invariance property implies that steady-state output could well be represented by a smooth time dependent linear trend. If this is true, unit root tests and β-convergence are testing for the same hypothesis.

This paper starts from the observation that the size of the unit root component in GDP is usually found to be very low as stated in Cochrane (1988), Campbell and Mankiw (1987), Michelacci and Zaffarone (2000) and follows Quah (1995) in noting that cross-section and time series analysis cannot arrive at different conclusions. One possible explanation is that the speed with which aggregate shocks are absorbed is so low that standard unit root tests fail to reveal it. This could actually be the case if GDP per capita exhibits long memory. So we define real convergence as mean reversion in the differences in per capita output among states in Brazil and we test this hypothesis using methodology based on fractional integration[1]. In this case, the parameter of integration can take on non-integer values. For certain values of the parameter of integration the income differencial process can be nonstationary but mean-reverting. That is, income shocks are persistent but eventually die out. This property of fractionally integrated process can capture the observed slow speed of income convergence.

Another point that justifies the use of long memory process to study convergence is that the long memory processes are theoretically justified in terms of aggregation of ARMA process with randomly varying coefficients, what can be the case of the standard Solow-Swan model allowed for cross sectional heterogeneity in the speed with which different units in the same country or state adjust.[2]

The fractional integration approach has already been applied to test real convergence in Michelacci and Zaffaroni (2000), henceforth MZ, Silverberg and Verspagen (2000) and Cunado, Gil-Alana and Perez de Gracia (2003). All of these papers apply the Auto-Regressive-Fractionally-Integrated-Moving-Average model (AFIRMA) in countries data. The results are mixed; MZ could not reject the hypothesis that all the OECD countries are nonstationary and mean reverting. Therefore, according to these authors, the convergence hypothesis cannot be reject, and thus, convergence takes place, although at an hyperbolic very slow rate. However, Silverberg and Verspagen (2000) re-examine the MZ results using Beran‘s nonparametric FGN estimator and Sowell‘s exact maximum likelihood ARFIMA estimator. They find that MZ‘s results no longer hold and thus, no evidence of convergence. Cunado et. al. (2003) using OECD data (Australia, Canada, Japan and UK) and parametric and semiparametric methods find that convergence is achieved in all countries, especially Australia and Canada.

There are a lot of papers which study convergence across Brazilian states. In the great majority of these papers the cross-sectional approach is used like in Ferreira and Diniz (1995), Schwartsman (1996), Ferreira and Ellery (1996) and more recently Ferreira (2000). The cross-section results in general present signs of absolute convergence depending on the period analyzed. All studies listed above indicate the presence of β-convergence. Some papers like Azzoni et all (2000) and Menezes and Azzoni (2000) using panel data find the presence of β-convergence too. It can be said that the cross-section approach does not consider useful information present in the data. An improvement is to deal with panel data, in which the different conditions to steady state situations in distinct regions are taken into account. But strong assumptions have to be made about parameter homogeneity. Given these, difficulties, an alternative is the analysis of individual countries or regions over time.To our knowledge only Azzoni and Barossi-Filho (2002) apply the time series approach to study convergence in Brazil. Using data on income per capita for 20 states, covering the period 1947-1998 and a methodology of unit root tests with 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,

[1] This indicator of convergence has been widely used in other empirical works based on convergence like the work of Silverberg and Verspagen (2000) and Cunado et al. (2003).

[2] For a better discussion about this see Michelacci and Zaffaroni (2000)

GO, PE) and 5 no sign of convergence (AM, PA, PI, SC, SP).[3] Recent criticism about the Azzoni and Barossi-Filho (2002) methodology (intervention-analysis) has already emerged. Indeed, as pointed out by Montañés et al. (2005) unit root tests based on intervention analysis are very sensitive to the specification of the alternative model. This technique also suffers from some interpretation problems. The presence of regime break is presumably suggestive of a absence of convergence since it implies that there is some component of the difference of the state series and the benchmark state series that will not disappear over a sufficiently long time horizon. The time series definition of convergence is violated by any long-term predictability in output differences. Hence, claims by authors that allowing for data breaks produces evidence of convergence begs the question of what is meant by convergence. So it is another motivation to the use of a different time series methods to study convergence across Brazilian states.

In this article we will use this new methodology of ARFIMA models to study convergence across Brazilian states. São Paulo (SP) state is used as benchmark state so we examine the order of integration of the real state per capita GSP series in the Brazilian states as well as their differences with respect to São Paulo.

The article is divided as follows. This introduction is followed by section 2, where we review the Solow-Swan model. Section 3 reviews the literature of convergence. Section 4 reviews the theory of long memory process and presents the methods used to estimate the fractional differencing parameter. Section 5 presents our data set and the results of the parametric and semiparametric tests. Section 6 concludes.

2 The Solow growth model

We begin this section by briefly reviewing the Solow growth model. In his classic 1956 article, Solow proposed that we begin to study of economic growth by assuming a standard neoclassical production function with decreasing returns to capital. Solow's model takes the rates of saving, population growth, and technological progress as exogenous. There are two inputs, capital and labour, which are paid their marginal products. Production at time t is given by:

$$Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha}, \tag{1}$$

where: Y is output, K capital, L labour, and A the level of technology. L and A are assumed to grow exogenously at rates n and g:

$$\begin{aligned} L_t &= L_0 e^{nt}. \\ A_t &= A_0 e^{gt}. \end{aligned} \tag{2}$$

The number of effective units of labour, $A_t L_t$, grows ar rate $n + g$.

The models assumes that a constant fraction of output, s, is invested. Defining K as the stock of capital per effective unit of labour, $\widehat{k} = K/AL$, and $\widehat{y}$ as the level of output per effective unit of labour, $\widehat{y} = Y/AL$, the evolution of $\widehat{k}$ is governed by:

[3] Table 4 in appendice contains a complete description of the states name and abbreviation

$$\frac{d\widehat{k}_t}{dt} = s\widehat{k}_t^{\alpha} - (g+\delta)\widehat{k}_t. \tag{3}$$

Where δ is the rate of depreciation. Equation 3 implies that k converges to a steady-state value k^* defined by:

$$k^* = (\frac{s}{g+\delta})^{1/(1-\alpha)}. \tag{4}$$

The central predictions of the Solow model concern the impact of saving and population growth on real income.

We can consider a log-linear approximation of equation (3) around the steady state, so that

$$\frac{d[\ln(\widehat{y}_t)]}{dt} = -\beta[\ln(\widehat{y}_t) - \ln(\widehat{y}_t^*)], \tag{5}$$

with

$$\beta = (1-\alpha)(g+\delta), \tag{6}$$

where $\widehat{y}^* = (\widehat{k}^*)^{\alpha}$. Discretizing Eq. (5) and indicating with y_t the log of output per capita, viz. $y_t = \ln(Y_t/L_t)$ and by y_t^* the log of the level of output per capita in steady-state we get:

$$y_t - y_{t-1} = g + \beta y_{t-1}^* - \beta y_{t-1}, \tag{7}$$

or equivalently:

$$y_t - y_t^* = (1-\beta)[y_{t-1} - y_{t-1}^*]. \tag{8}$$

We now analyze the time series properties of both equation (7) and (8). Equation (7) is the basic equation used to test beta-convergence. The β-convergence is an inverse relationship between the growth rates of the state per capita incomes and the initial per capita income (PCIs) levels. It applies if a poor economy tends to grow faster than a rich one. This happens if the estimate of the coefficient beta in Eq. (7) is positive and significantly different from zero. It is important to say that the literature differentiates two kinds of β-convergence: unconditional and conditional β-convergence. The second one in addition to the initial per capita income it includes extra explanatory variables like population growth, rate of capital depreciation and technological progress.

If we find a positive and significantly beta coefficient the aggregate shocks that have pushed the current level of output away from the steady-state level will be absorbed at the exponential rate beta so that the dynamics of output will exhibit mean reversion. In practice, international empirical studies repeatedly find a 2% coefficient, uniform across countries and significantly different from zero (Quah 1993).

3 Literature Review

In this section we review the literature about convergence. First we discuss the recent literature about the use of AFIRMA models in growth studies after that we discuss the main articles that study convergence across Brazilian states. This section is divided in two subsections. In each one of them the articles will be present in chronological order. In the first subsection we present the international studies and in the second the Brazilian experience.

3.1 International experience

To our knowledge, the first paper to use the theory of AFIRMA processes to study convergence is Michelacci and Zaffaroni (2000). They start the article showing the conditions under which long memory arises naturally. In their point of view long memory arises as the result of aggregating heterogeneous units in the same economy what can be find in the Solow-Swan model augmented by cross-sectional heterogeneity. They use the log-periodogram estimator of Robinson (1995a) to find out the fractional integration parameter of the GDP per capita of the OECD countries. Their results show that the GDP per capita of all the countries in the sample exhibit long memory and that time series are non-stationary but mean reverting. The final conclusion of the paper is that the OECD countries present a rate of convergence of 2%.

In a note of the MZ paper Silverberg and Verspagen (2000) argue that MZ used estimators that are questionable for the purpose of clarifying the time series properties of their data. The first point is that MZ filter out a deterministic linear-in-logs instead of first-differentiate in logs as is usual in the long memory literature. Second they rely on a semiparametric Geweke and Porter-Hudak (GPH) method as modified by Robinson (1995a), which is known to be highly biased in small samples.

To avoid these problems Silverberg and Verspagen (2000) re-examine the results of MZ using Beran's nonparametric FGN estimator and Sowell's exact maximum likelihood ARFIMA estimator. In the authors point of view these methods avoid the small sample bias and arbitrariness of the cut-off parameters of Robinson's method and allow controlling for the short memory effects although the parametric ARFIMA estimator introduces specification problems of its own. They also look for the influence of the choice of sub-periods on the results.

Silverberg and Verspagen (2000) find no evidence for fractional integration in the range $(0.5, 1)$, an essential point in MZ argument in the context of β-convergence. Allowing for short memory, the evidence for persistence is very weak and antipersistence is as common as persistence. Finally, they apply the Robinson's method to their treatment of the data what show that the MZ results no longer hold. In their opinion, MZ results are probably artifacts of their detrending and data filtering techniques.

More recently Cunado et al. (2003) test the real convergence in some OECD countries (Australia, Canada, Japan and United Kingtom). The authors define real convergence as mean reversion in the differences in per capita output among countries. Their methodology differed from that used by MZ in the following aspects: first, while MZ subtract an exponential (common) trend to the data before testing for long memory Cunado et al. (2003) test for long memory using first differences of the logged series. Second, they use a parametric and semi-parametric techniques that avoid the small sample bias brought by the use of log-periodogram regression techniques. The parametric method is due to Robinson (1994), and the semiparametric procedure is the Quasi Maximum

Likelihood Estimate (QMLE).

The data are the annual log real GDP per capita in 1990 Geary-Khamis PPP-adjusted dollars. The series goes from 1900 to 2001 for five OECD countries (Australia, Canada, Japan, United Kingdom and United States), which came from Maddison (1995) and have been updated using the GGDC (Groningen Growth and Development Centre).

Both the parametric and semiparametric methods show that convergence is achieved in all countries, especially for the cases of Australia and Canada.

3.2 The Brazilian experience

There are a lot of papers which study convergence across Brazilian states. In the great majority of these papers the cross-sectional approach is used like in Ferreira and Diniz (1995), Schwartsman (1996), Ferreira and Ellery (1996) and more recently Ferreira (2000).

The cross-section results in general present signs of absolute convergence depending on the period analyzed. All studies listed above indicate the presence of β-convergence.

Other methodologies like panel data analysis was already applied to study convergence across Brazilian states, some examples are Azzoni et. al. (2000) and Menezes and Azzoni (2000). Both papers find the presence of β-convergence what supports the cross-sectional conclusions.

To the best of our knowledge only Azzoni and Barossi-Filho (2002) apply 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. 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.

Following the seminal paper of Perron (1989) which stress the importance of structural breaks for testing the null hypothesis of unit root, the authors use a methodology of unit root tests with endogenously structural break points. They calculate two unit root tests with break points, based on Lee and Strazicich (1999a) and Lee and Strazicich (1999b): one-break and two-break minimum LM tests. For each of them, they admit two possibilities for the model set up: crash model and break trend models.

They work with three different geographic levels. Initially they deal with the five official macro-regions of the country and consider whether or not their relative income levels are converging. After that they move to the second geographical level, in which they compare the relative income level of each state to the region it belongs to. Finally, they compare the relative income level of each state to the country. Their results indicate that there are signs of stochastic convergence of income at the macro regional level, with the exception of the North region. Convergence within the regions, that is, states converging to the income level in the region they belong to, is not homogeneous in the country. Finally, in the third geographical level of analysis 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 no sign of convergence(AM, PA, PI, SC, SP).

4 Theory of long memory process

Fractional integrated models were introduced in the economics literature by Granger (1980), and Granger and Joyeux (1980) and they were theoretically justified in terms of aggregation of ARMA process with randomly varying coefficients by Robinson (1978) and Granger (1980). 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 from an empirical data oriented approach in terms of the persistence of observed autocorrelations. When viewed as the time series realization of a stochastic process, the autocorrelation function exhibits persistence that is neither consistent with a covariance stationary process nor an unit root process. The extent of the persistence is consistent with an essentially stationary process, but where 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 $\rho_y(\tau) \sim \tau^{2d-1}$, for large τ, while a stationary ARMA process have a geometrically decaying function given by $\rho_y(\tau) \sim r^\tau$ where $|r| < 1$.

So to resume 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 that 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:

$$(1-L)^d y_t = v_t, \quad \text{t=1,2,...} \tag{9}$$

where L is the lag operator, y_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 (9) is called ARFIMA(0,d,0) (Auto-Regressive-Fractionally-Integrated-Moving-Average) model. If $d = 0$, then y_t is a standard or better short memory stationary process whereas y_t is a random walk if $d = 1$. For values of $d > -1$, 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 $-1/2 < d < 1/2$. So the process in (9) is stationary for values of the parameter d lying in the interval $(-1/2, 1/2)$. For values of d in the interval $(1/2, 1)$ the process is non-stationary, but exhibits mean reversion. In summary, for values of the fractional differencing less than the unity income shocks die out, even when the y_t process is non-stationary.

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 Y_t is said to exhibit long memory in terms of the power spectrum (when it exists) with parameter $d > 0$ if:

$$f(\lambda) \sim K\lambda^{-2d} \quad \text{as } \lambda \to 0^+. \tag{10}$$

In a non-stationary case ($d \geq 1/2$) $f(\lambda)$ is not integrable and thus it is defined as a pseudo-spectrum.

When ν_t is assumed to be a white noise process, the process y_t defined in equation (10) is called an ARFIMA(0,d,0) process and when ν_t is an (inverted)

ARMA (p,q) we obtain an ARFIMA(p,d,q) process. The power spectrum of the y_t process is given by

$$f_y(\lambda) = \mid 1 - e^{i\lambda} \mid^{-2d} f_v(\lambda) = (2sen(\lambda/2))^{-2d} f_v(\lambda), \qquad -\pi \leq \lambda < \pi \quad (11)$$

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

$$f_y(\lambda) \sim 4^{-d} f_v(0) \lambda^{-2d}. \qquad (12)$$

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

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

A more complete discussion about fractional integration process can be found in Baillie(1996).

4.1 Estimation and testing

There a lot of techniques to estimate and test AFIRMA models. In this section we will explain the techniques more used in empirical studies and justify our choice for two of them.

We use two different estimators for the fractional differencing parameter. First, we use a parametric testing procedure due to Robinson (1994) which has been widely employed in macroeconomic time series. The second procedure is the Quasi Maximum Likelihood Estimate (QMLE) due to Robinson (1995b) and available in OX AFIRMA package.

The parametric testing procedure due to Robinson (1994) is a Lagrange Multiplier (LM) test of the null hypothesis:

$$H_0 : d = d_0, \qquad (13)$$

in a model given by

$$y_t = \beta' z_t + x_t, \qquad t = 1, 2, ..., \qquad (14)$$

and (9), for any real d_0, where y_t is the time series we observe; $\beta = (\beta_1, ..., \beta_k)'$ is a $(k \times 1)$ vector of unknown parameters; and z_t is a $(k \times 1)$ vector of deterministic regressors that may include, for example, an intercept, (eg. $z_t \equiv 1$), or an intercept and a linear time trend, (in case of $z_t = (1, t)'$). Specifically, the test statistic is given by:

$$\widehat{r} = \frac{T^{1/2}}{\widehat{\sigma}^2} \widehat{A}^{-1/2} \widehat{a}, \tag{15}$$

where T is the sample size and

$$\widehat{a} = \frac{-2\pi}{T} \sum_{j=1}^{T-1} \Psi(\lambda_j) g(\lambda_j; \widehat{\tau})^{-1} I(\lambda_j) \tag{16}$$

$$\widehat{\sigma^2} = \frac{2\pi}{T} \sum_{j=1}^{T-1} g(\lambda_j; \widehat{\tau})^{-1} I(\lambda_j) \tag{17}$$

$$\widehat{A} = \frac{2}{T} (\sum_{j=1}^{T-1} \Psi(\lambda_j)^2 - \sum_{j=1}^{T-1} \Psi(\lambda_j) \widehat{\in} (\lambda_j)' \times (\sum_{j=1}^{T-1} \widehat{\in} (\lambda_j) \widehat{\in} (\lambda_j)')^{-1} \times \sum_{j=1}^{T-1} \widehat{\in} (\lambda_j) \Psi(\lambda_j) \tag{18}$$

$$\Psi(\lambda_j) = \log \mid 2 \sin \frac{\lambda_j}{2} \mid ; \widehat{\in} (\lambda_j) = \frac{\partial}{\partial \tau} \log g(\lambda_j; \widehat{\tau}); \lambda_j = \frac{2\pi j}{T}; \widehat{\tau} = \arg\min \sigma^2(\tau) \tag{19}$$

$I(\lambda_j)$ is the periodogram of v_t evaluated under the null, i.e.

$$\widehat{v_t} = (1-L)^{d_0} y_t - \widehat{\beta}' w_{t;} \tag{20}$$

$$\beta = \left(\sum_{t=1}^{T} w_t w_t' \right)^{-1} \sum_{t=1}^{T} w_t (1-L)^{d_0} y_t; \quad w_t = (1-L)^{d_0} z_t, \tag{21}$$

and the function g bellow is a known function coming from the spectral density function of $v_{t,}$

$$f(\lambda, \sigma^2, \tau) = \frac{\sigma^2}{2\pi} g(\lambda, \tau), \qquad -\pi < \lambda < \pi \tag{22}$$

Because these tests are purely parametric, they require specific modelling assumptions to be made regarding the short memory specification of $v_t$thus, if v_t is white noise, $g \equiv 1$, and if v_t is an AR process of the form $\phi(L) v_t = \varepsilon_t$, $g = \mid \phi(e^{i\lambda}) \mid^{-2}$, with $\sigma^2 = V(\varepsilon_t)$, so that the AR coefficients are function of τ.

Based on the null hypothesis H_0 (13), Robinson (1994a) established that under certain regularity conditions:

$$\widehat{r} \to_d N(0,1) \quad \text{as } T \to \infty \ , \tag{23}$$

and it is also efficient (in the Pitman efficiency theory sense) since it is the most powerful test against local departures from the null hypothesis. Thus, we are in a classical large sample testing situation: an approximate one-sided $100\alpha\%$ level test of H_0 (13) against the alternative $H_1 : d > d_0$ ($d < d_0$) will be given by the rule: "Reject H_0 if $\widehat{r} > z_a$ ($\widehat{r} < -z_a$)", where the probability that a standard normal variate exceeds z_a is α.

A problem with the parametric procedures is that the model must be correctly specified. Otherwise, the estimates are liable to be inconsistent. In fact, misspecification of the short run components of the process may ivalidate the estimation of the long run parameter d . This is the main reason for using also in this article a semiparametric procedure.

There exist several methods for estimating the fractional differencing parameter in a semiparametric way. One of the most used in the literature is the log-periodogram regression estimate which was initially propose by Geweke and Porter-Hudak (1983) and modified later by Kunsch (1986) and Robinson (1995a). The problem with this procedure is that it is highly biased in small samples. So to avoid this small-sample-bias problem, we use the QMLE and the parametric procedure.

The QMLE is a semiparametric procedure which is basically a local "Whittle estimate" in the frequency domain, considering a band of frequencies that degenerates to zero. The main reason for using the QMLE is based on its computational simplicity along with the fact that it requires a single bandwidth parameter, unlike other procedures where a trimming number is also required. It is important to say that Gil-Alana (2002) using Monte Carlo simulations show that the QMLE of Robinson (1995b) outperforms the other semiparametric models in a number of cases.

The estimate is implicitly defined by:

$$d_1 = \arg\min_d (\log C(d) - 2d1/m \sum_{j=1}^{m} \log \lambda_j) \tag{24}$$

$$\begin{aligned} \text{for } d &\in (-1/2; 1/2) : C(d) = 1/m \sum_{j=1}^{m} I(\lambda_j)\lambda_{j,}^{2d} \\ \lambda_j &= 2\pi j/T, \quad m/T \to 0 \end{aligned} \tag{25}$$

Under finiteness of the fourth moment and other conditions, Robinson (1995b) proves the asymptotic normality of this estimate, while Lobato (1999) extended it to the multivariate case.

5 Data and Results

Our data set consist of annual log real Gross State Product (GSP) per capita. The series runs from 1947 to 1998 for twenty Brazilian states[4], namely, Alagoas, Amazonas, Bahia, Ceará, Espírito Santo, Goiás, Maranhão, Mato Grosso, Minas Gerais, Pará, Paraíba, Paraná, Pernambuco, Piauí, Rio Grande do Norte, Rio Grande do Sul, Rio de Janeiro, Santa Catarina, São Paulo and Sergipe. The GSP data have been obtained from Azzoni (1997) and the population data have been obtained from Instituto Brasileiro de Geografia e Estatistica (IBGE).

As mentioned early, as an indicator of real convergence, we used the differences of the per capita GSP of each of the nineteen states with respect to the São Paulo state, used as the benchmark state because it is the richest state in Brazil.

[4]Although nowadays Brazil has 27 states, we work with a set of only 20 because 7 states don´t exist in 1947 (initial time series data analyzed).

The first thing that we do here is to perform the tests of Robinson (1994) described in the last section to the individual series as well as to their differences with respect to the São Paulo (SP). Denoting each of the time series y_t, we employ throughout the model given by (13) and (14), with $z_t = (1, 0)$ Thus, under the null hypothesis H_0:

$$y_t = \beta_0 + \beta_1 t + x_t, \qquad t = 1, 2, ... \tag{26}$$

$$(1 - L)^{d_0} x_t = v_t, \qquad t = 1, 2, ... \tag{27}$$

and we treat the case β_0 unknown and $\beta_1 = 0$ a priori. In other words we consider the case of an intercept and no linear trend. We will model the $I(0)$ process v_t to be both white noise and to have parametric autocorrelation.

We start with the assumption that v_t is white noise. We report test statistics for other fractionally integrated possibilities like $d_0 = 0$ and $d_0 = 0.5$. In all states, we observe a monotonic decrease of $\widehat{r}$ as d_0 increases. Such monotonicity is a characteristic of any reasonable statistic, given correct specification and adequate sample size, because for example, we would wish that if H_0 is rejected with $d_0 = 1$ against alternatives of form $H_a : d > 1$, an even more significant result in this direction should be expected when $d_0 = 0.75$ or $d_0 = 0.5$ are tested. However, misspecification inflates both numerator and denominator of $\widehat{r}$ to varying degrees, and thus affects $\widehat{r}$ in a complicated way. So computing $\widehat{r}$ for a range of d_0 values is thus useful in revealing possible misspecifications.

The test statistic reported in Table 1 is one-sided. We report the values of d_0 that yields the lowest statistics $\widehat{r}$. We report such a result for individual series as well as the differenced series with respect to the São Paulo state. Finally we mark in bold those values of d_0 for the differenced series that are smaller than the ones of the individual series. This reduction in the order of integration (d_0) is an evidence of convergence towards the output of São Paulo state.

Table1

Values of d which produces the lowest statistics in absolute value		
	Individual Series	With respect to the SP state
Alagoas	1.02	**0.70**
Amazonas	0.95	**0.91**
Bahia	1.28	**1.09**
Ceará	0.65	**0.64**
Espírito Santo	0.74	0.74
Goías	0.64	**0.36**
Maranhão	0.73	**0.66**
Mato Grosso	0.70	**0.67**
Minas Gerais	0.93	**0.76**
Pará	0.86	0.89
Paraíba	0.79	0.81
Paraná	0.78	**0.71**
Pernambuco	1.09	**0.94**
Piauí	0.82	**0.73**
R. G. do Norte	0.75	0.77
R. G. do Sul	1.12	**0.71**
Rio de Janeiro	1.42	**1.39**
S. Catarina	1.00	**0.91**
Sergipe	1.04	**0.97**
São Paulo	1.23	

Using white noise v_t we observe smaller $\widehat{d}$ in case of Alagoas, Amazonas, Bahia, Ceará, Goias, Maranhão, Mato Grosso, Minas Gerais, Paraná, Pernanbuco, Piauí, R. G. do Sul, Rio de Janeiro, Santa Catarina and Sergipe. Based on the tests of Robinson (1994) we can say that the states listed above present's convergence.

Another important result from Table 1 is that in twelve out of the twenty states, the individual series presents the d parameter between $\frac{1}{2} < d < 1$ so these states have infinite variance but they still have mean reversion, with the effect of the shocks dying away in the long run as opposed to the unit root case with shocks persisting forever. For these states the autocorrelation function exhibits persistence that is neither consistent with a covariance stationary process nor an unit root process. In the case of the series with respect to the SP state we find that for seventeen of nineteen series the d parameter is between $\frac{1}{2} < d < 1$.So there is evidence of long memory in the time series of Brazilian real GSP per capita.

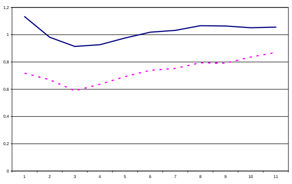
We also fitted AR models to v_t. The results are not reported in this article because we observe a lack of monotonicity in $\widehat{r}$ with respect to d_0 in all practically series. This could be explained in terms of model misspecification or the sample size.

Table 2 [5]reports the results based on the QMLE of Robinson (1995b), i.e., $\widehat{d_1}$ given by (24) and (25) for a range of values of m from 10 to 20. Since the time series are clearly nonstationary, the analysis will be carried out based on the first differenced data, adding then 1 to the estimated values of d to obtain the proper orders of integration of the series. The series in bold are the original states series and the other ones are the series in differences with respect to São

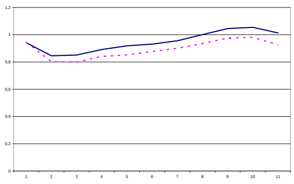
[5]In appendice the table 5 is the table 2 but in a larger scale

Paulo state.

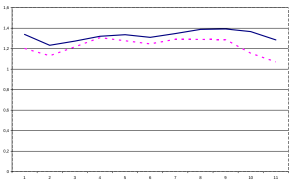
Alagoas



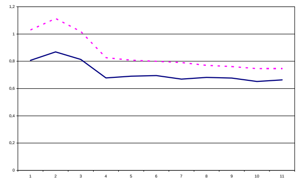
Amazonas



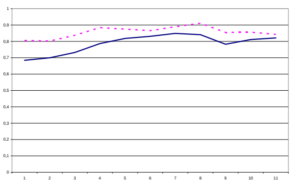
Bahia



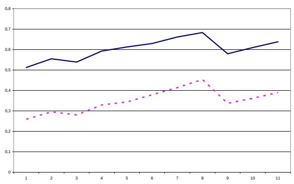
Ceará



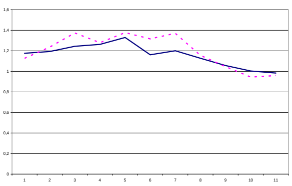
Espírito Santo



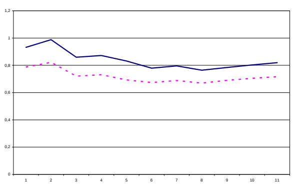
Goias



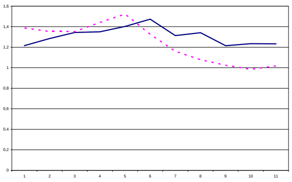
Maranhão



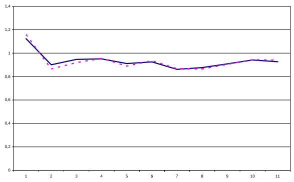
Mato Grosso



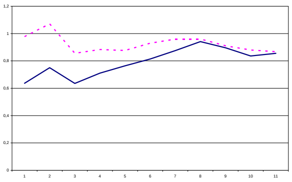
Minas Gerais



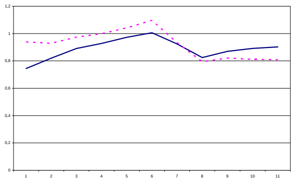
Pará



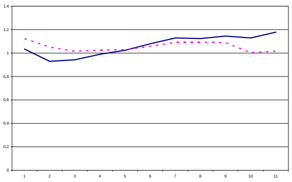
Paraíba



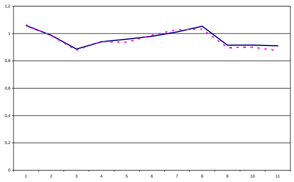
Paraná



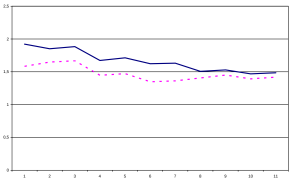
Pernanbuco



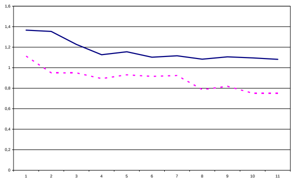
Piauí



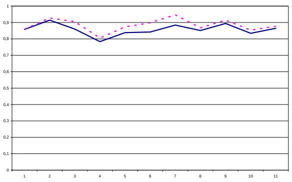
Rio de Janeiro



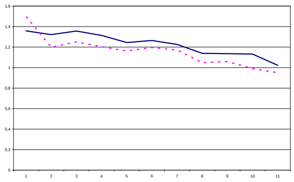
Rio Grande do Sul



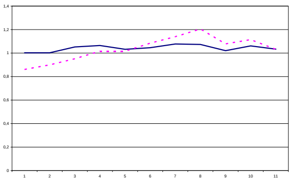
Rio Grande do Norte



Santa Catarina



Sergipe



We see that for Alagoas, Bahia, Goias, Minas Gerais, Pernanbuco and Rio Grande do Sul, the estimated values of d are strictly higher for the case of the individual series. So these states present a strong convergence. For the states: Amazonas, Mato Grosso, Piauí, Rio de Janeiro and Santa Catarina we have convergence. On the other hand, the values of d in Ceará, Espírito Santo, Paraíba and Rio Grande do Norte are higher for the differenced data so in these states we don't have convergence. Finally Maranhão, Pará, Paraná and Sergipe we have convergence for some values of m.

In Table 3 we combine the results from Table 1 (results from the parametric procedure) and Table 2 (results from QMLE) and presents our conclusions.

Table 3: States which presents convergence

	Parametric method	Semiparametric method (QMLE)	Conclusion
Alagoas	Y	Y	Y
Amazonas	Y	Y	Y
Bahia	Y	Y	Y
Ceará	Y	N	Y
Espírito Santo	N	N	N
Goías	Y	Y	Y
Maranhão	Y	For some values	Y
Mato Grosso	Y	Y	Y
Minas Gerais	Y	Y	Y
Pará	N	For some values	Y
Paraíba	N	N	N
Paraná	Y	For some values	Y
Pernambuco	Y	Y	Y
Piauí	Y	Y	Y
R. G. do Norte	N	N	N
R. G. do Sul	Y	Y	Y
Rio de Janeiro	Y	Y	Y
S. Catarina	Y	Y	Y
Sergipe	Y	For some values	Y

Y- represent that the method used showed that the state is converging

N- represent that the method used showed that the state is not converging

The results above are consistent with the ones obtained using the parametric method of Robinson (1994). For all states but Ceará, we have convergence in the parametric procedure and don´t have in the semiparametric one. The state of Pará which don't show convergence in the Robinson (1994) presents convergence for some values of m, so we can say that these ones are converging to the benchmark state. The states of Espírito Santo, Paraíba and Rio Grande do Norte show no convergence in both procedures.

6 Conclusion

In this paper we test the β-convergence hypothesis across Brazilian states. In order to do so we use ARFIMA models. In particular, we have examined the order of integration of the annual log real Gross State Product (GSP) per capita series for twenty states, namely, Alagoas, Amazonas, Bahia, Ceará, Espírito Santo, Goiás, Maranhão, Mato Grosso, Minas Gerais, Pará, Paraíba, Paraná, Pernambuco, Piauí, Rio Grande do Norte, Rio Grande do Sul, Rio de Janeiro, Santa Catarina, São Paulo and Sergipe as well as their differences with respect to the São Paulo state. We use two different estimators for the fractional differencing parameter: a parametric testing procedure due to Robinson (1994) and a semiparametric estimation method (QMLE) due to Robinson (1995b).

Combining the results from the two procedures we can say that the convergence hypothesis is strongly satisfied in the cases of Alagoas, Amazonas, Bahia, Goiás, Mato Grosso, Minas Gerais, Pernanbuco, Piauí, Rio Grande do Sul, Rio de Janeiro and Santa Catarina and weakly satisfied in the cases of Ceará, Maranhão, Pará, Paraná and Sergipe .The states of Espírito Santo, Paraíba and Rio Grande do Norte show no convergence. So the great majority of Brazilian real Gross State Product (GSP) per capita present's convergence to the GSP of São Paulo state. In others words the poor states are growing faster than the rich state (SP).

Another output of this work is to show that the order of integration of different states in Brazil are different, in line of this the time series tests of convergence based on cointegration are misspecified.

It is important to say that our time series results are in line with the works of Ferreira and Diniz (1995), Schwartsman (1996), Ferreira and Ellery (1996) and more recently Ferreira (2000) which uses a cross-sectional approach to study convergence across Brazilian states. So with this work we can say that the convergence process is occurring in Brazil and that cross-section and time series analysis don't arrive at different conclusions when convergence is analyzed with the use of Brazilian state data.

Our results are not directly compared with the ones from Azzoni and Barossi-Filho (2002) for two reasons: the first one is that the two works use 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; the second one is the difference in the data. We work with a set of twenty states maintaining its original data series while Azzoni and Barossi-Filho (2002) 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. Stressed these points let's compare our results with the results from Azzoni. We can say that while Espírito Santo, Paraíba and Rio Grande do Norte show no convergence in our paper, Azzoni and Barossi-Filho (2002) find that Amazonas, Pará, Piauí, Santa Catarina and São Paulo are not converging. All the other states are converging in both articles.

It is important to say 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. While we use long memory models Azzoni and Barossi-Filho (2002) use the unit root tests with endogenously structural break points, which have receive some criticism recently. Indeed, as pointed out by Montañés et al. (2005) unit root tests based on intervention analysis are very sensitive to the specification of the alternative model. 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 is more trustworthy than Azzoni and Barossi-Filho (2002).

Other issues such as potential presence of structural breaks on the data and the effect that this may have on the results can be study in future papers. Other interesting question is to study convergence across Brazilian states for specific regions like North, South, and Northeast etc each one with its own benchmark state that could be an aggregation of all participants' states of the region.

References

[1] Azonni, C., 1994. Crescimento Econômico e Convergência das Rendas Regionais: O Caso Brasileiro. Anais do XXII Encontro Nacional de Economia da ANPEC, 1, pp. 185-205.

[2] Azonni, C., 1997. Concentração regional e dispersão das rendas per capita estaduais: análise a partir de séries históricas estaduais de PIB. 1939-1995.

[3] Azonni, C., 2001. Economic Growth and Regional Income Inequalities in Brazil. Annals of Regional Science, Vol 35, No.1.

[4] 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.

[5] 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

[6] Baillie, R., T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5-59.

[7] Barro, R., 1991. Economic Growth in a cross-section of countries. Quarterly Journal of Economics, CVI, 407-55.

[8] Barro, R. and Sala-i-Martin, X., 1992. Convegence. Journal of Political Economy 100, pp.223-51.

[9] Barro, R. and Sala-i-Martin, X., 1995. Economic Growth. McGraw-Hill, Inc.

[10] 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.

[11] 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.

[12] Ferreira, A.H.B., 2000. Convergence in Brazil: Recent Trends and long Run Prospects. Applied Economics, 32,pp.479-489.

[13] 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.

[14] 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.

[15] Gil-Alana, L. A., 2000. Mean reversion in the real exchange rates, Economics Letters, 16, 285-8.

[16] Gil-Alana, L. A., 2002. Comparisons between semiparametric procedures for estimating the fractional differencing parameter, Preprint.

[17] 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.

[18] 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.

[19] Granger, C., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227-38.

[20] 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.

[21] Jones, C., 1995. Time series tests of endogenous growth models. Quarterly Journal of Economics 110, 495-525.

[22] Jones, C., 1997. Convergence Revisited. Journal of Economic Growth, 2, pp. 131-153.

[23] Kunsch, H., 1986. Discrimination between monotonic trends and long range dependence. Journal of applied Probability 23, 1025-30.

[24] Lee, J. and Strazicich, M., 1999a. Minimum LM unit root tests. Working Paper, University of Central Florida.

[25] Lee, J. and Strazicich, M., 1999b. Minimum LM unit root test with two structural breaks. Working Paper, University of Central Florida.

[26] Lobato, I., 1999. A semiparametric two-step estimator for a multivariate long memory process. Journal of Econometrics, 73, 303-24.

[27] Maia Gomes, 2002. Regional Development Strategies in Brazil. mimeo.

[28] Mankiw, G., Romer, P. and Weil, D. N., 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics, 107, 407-37.

[29] Mello, M. and Perreli R,.2003. Growth equations: a quantile regression exploration. The Quarterly Review of Economics and Finance 43, pp643-667.

[30] 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.

[31] Michelacci, C.; and Zaffaroni, P., 2000. Fractional beta convergence. Journal of Monetary Economics 45, pp. 129-153.

[32] Montañés, A., Olloqui, I., Calvo, E., 2005. Selection of the break in the Perron-type tests. Journal of Econometrics, 129, 41-64.

[33] 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.

[34] Perron, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. Econometrica, 57, 1361-1401.

[35] Quah, D. 1993., Empirical cross-section dynamics in economic growth. European Economic Review 37 (2/3), 426-34.

[36] Quah, D. 1995, Empirics for economic growth and convergence. European Economic Review 40, n6, 1353-75.

[37] Robinson, P. M., 1978. Statistical inference for a random coefficient autoregressive model. Scandanavian Journal of Statistics 5, 163-8.

[38] Robinson, P. M., 1994. Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association, 89, 1420-37.

[39] Robinson, P. M., 1995a. Log-periodogram regression of time series with long range dependence. Annals of Statistics, 23, 1048-72.

[40] Robinson, P. M., 1995b. Gaussian semiparametric estimation of long range dependence. Annals of Statistics, 23, 1630-61.

[41] Schwartsman, A., 1996. Convergence across Brazilian States. Discussion Paper, n 02/96. IPE, Universidade de São Paulo, 1996.

[42] 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.

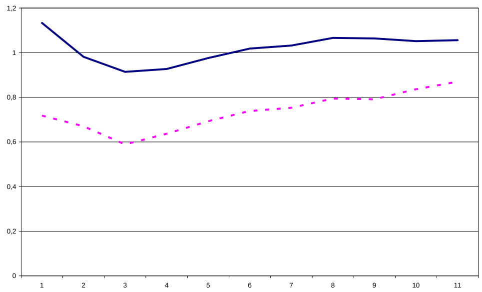
7 Apendice

Table 4

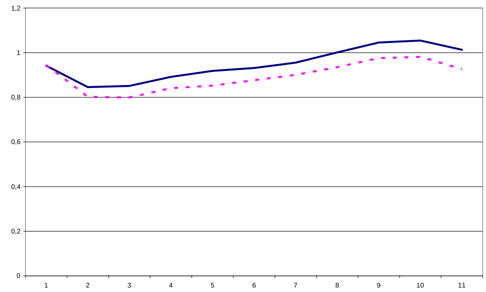
State	Abbreviation
Alagoas	AL
Amazonas	AM
Bahia	BA
Ceará	CE
Espírito Santo	ES
Goías	GO
Maranhão	MA
Mato Grosso	MT
Minas Gerais	MG
Pará	PA
Paraíba	PB
Paraná	PR
Pernambuco	PE
Piauí	PI
R. G. do Norte	RN
R. G. do Sul	RS
Rio de Janeiro	RJ
S. Catarina	SC
Sergipe	SE
São Paulo	SP

Table 5

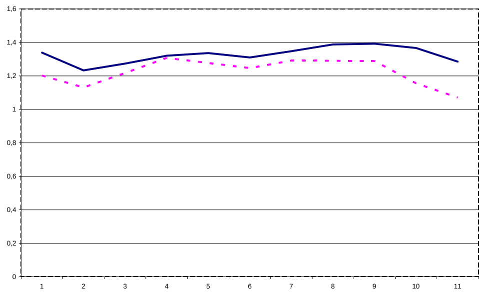
Alagoas



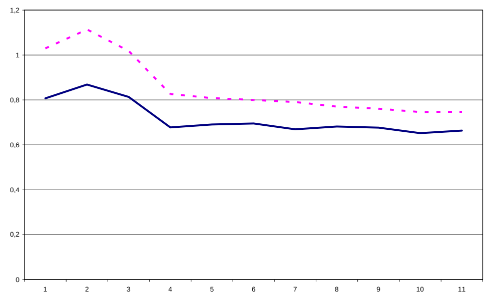
Amazonas



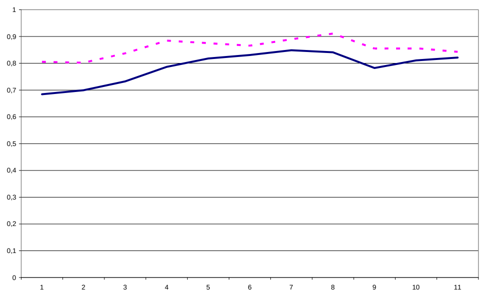
Bahia



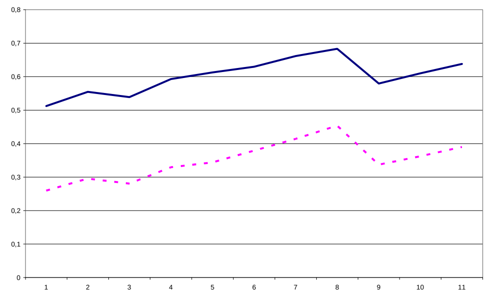
Ceará



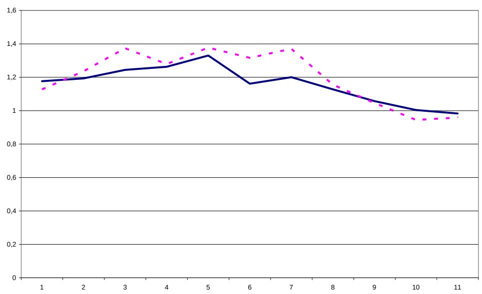
Espírito Santo



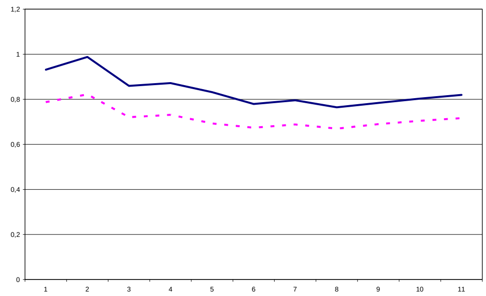
Goias



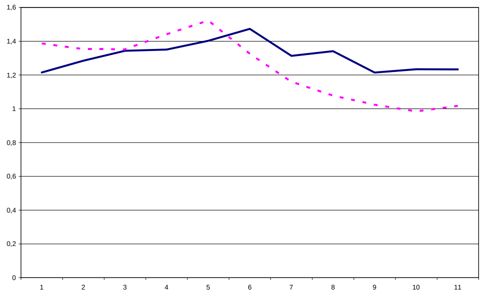
Maranhão



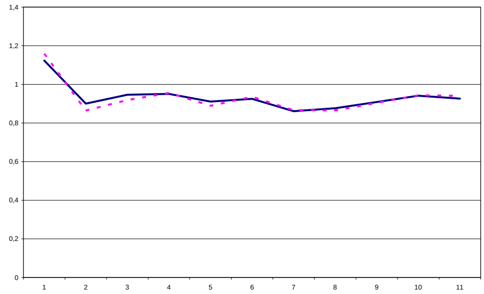
Mato Grosso



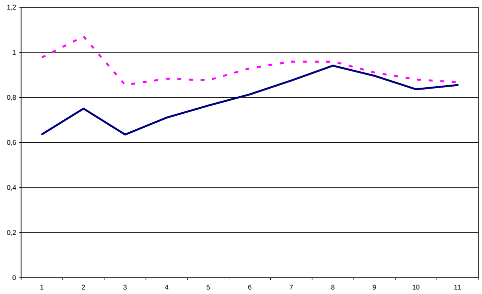
Minas Gerais



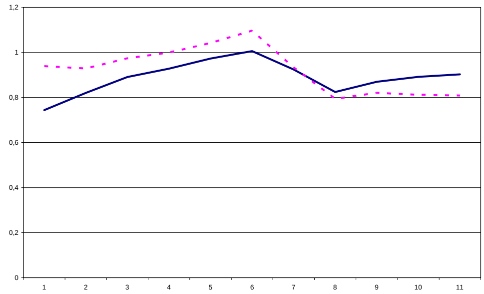
Pará



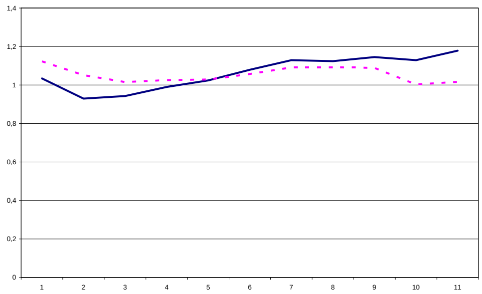
Paraíba



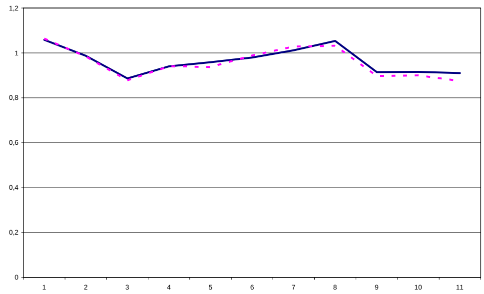
Paraná



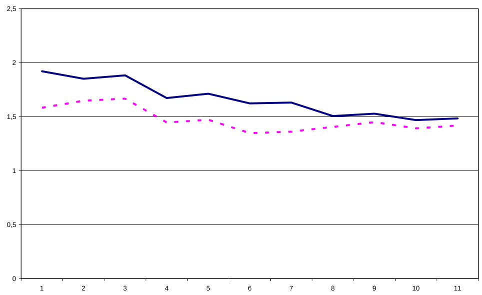
Pernanbuco



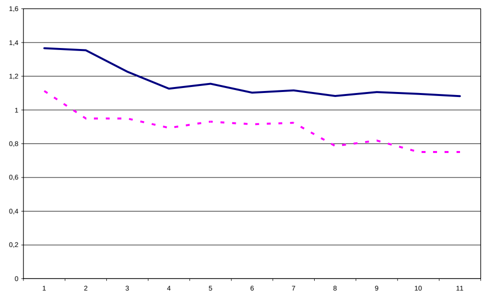
Piauí



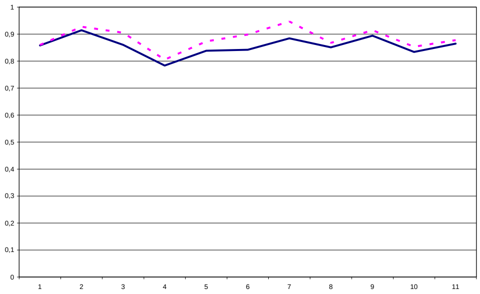
Rio de Janeiro



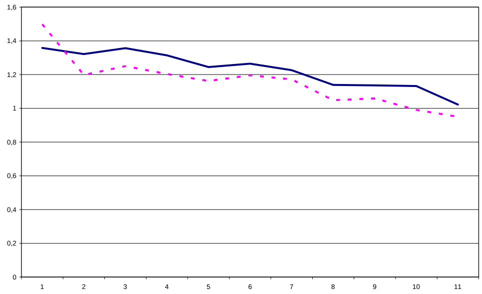
Rio Grande do Sul



Rio Grande do Norte



Santa Catarina



Sergipe