

Uma Análise da Hipótese da Convergência para os Municípios Brasileiros

Rodrigo Lara Pinto Coelho*, Lízia de Figueiredo†

Sumário: 1. Introdução; 2. Revisão da Literatura; 3. Estratégia Empírica; 4. Dados; 5. Resultados; 6. Absoluta vs. Clube; 7. Conclusão; A. Tabelas e Figuras.

Palavras-chave: convergência clube; métodos não-paramétricos; testes de robustez.

Códigos JEL: C14; R11.

Este artigo busca determinar qual das hipóteses da convergência – absoluta, condicional ou clube – melhor descreve a dinâmica da renda dos municípios brasileiros no período 1970-2000. Os resultados obtidos com base no método *regression tree* e em testes de robustez demonstram a importância das condições iniciais na determinação da taxa de crescimento da renda dos municípios, o que sugere adominância da hipótese da convergência clube sobre as demais.

This paper attempts to ascertain which of the convergence hypotheses – absolute, conditional or club – best describes the dynamics of the Brazilian municipalities income over the period 1970-2000. The results based on the regression tree method and on robustness tests demonstrate the importance of initial conditions on the determination of municipalities income growth rate, which suggests the dominance of the club convergence hypothesis over the other two.

1. INTRODUÇÃO

A persistência da extrema pobreza, em algumas economias, em contraste com altos níveis de renda e crescimento sustentado, por longos períodos, em outras é uma das questões mais intrigantes e controversas da literatura de crescimento econômico. No que se refere à literatura empírica, uma parte significativa dos esforços busca validar uma das três versões da chamada hipótese da convergência, usualmente com o objetivo de legitimar algum modelo teórico em detrimento de outro¹:

- convergência β absoluta ou incondicional – as rendas *per capita* de duas economias quaisquer tenderão para o mesmo nível no longo prazo (estado estacionário) independentemente de suas condições iniciais;

*Banco Central do Brasil. Endereço: Av. Álvares Cabral, 1605, 2º andar - Belo Horizonte(MG). Telefone: (31)3253-7254. E-mail: rodrigo.coelho@bcb.gov.br

†Centro de Desenvolvimento e Planejamento Regional – Cedeplar UFMG. Endereço: Rua Curitiba, 832, 8o andar, Sala 808 - Belo Horizonte (MG). Telefone: (31) 3279-9085. E-mail: lizia@cedeplar.ufmg.br.

¹Veja Galor (1996) para uma discussão teórica acerca da hipótese da convergência.



- convergência β condicional – duas economias com características estruturais comuns (mesmas preferências, tecnologias, taxas de crescimento populacional, políticas públicas, etc.) terão o mesmo nível de renda *per capita* no longo prazo independentemente de suas condições iniciais;
- convergência clube – as rendas *per capita* de duas economias quaisquer tenderão para um mesmo nível de longo prazo somente se essas partilharem das mesmas características estruturais e ao mesmo tempo possuírem condições iniciais semelhantes.

Além de auxiliar na escolha do melhor modelo teórico, entender a natureza da dinâmica da renda permite ao pesquisador realizar inferências quanto ao efeito de choques na distribuição da renda: se houver convergência condicional, choques não-estruturais podem afetar somente temporariamente o *ranking* das rendas, ao contrário do que ocorre no caso da convergência clube.²

Na prática, a melhor forma de abordar o problema da determinação da dinâmica das rendas de uma determinada economia permanece em debate. Essa questão é especialmente controversa quando diferenças permanentes nos níveis de renda *per capita* são observadas, de sorte que apenas a hipótese da convergência absoluta pode ser refutada. A dúvida, nessa situação, seria a que atribuir a causa dessas diferenças: às diferentes características estruturais ou às condições iniciais distintas de cada economia? A princípio, nenhuma hipótese deve ser descartada de plano e, mesmo na situação descrita anteriormente na qual se observa a persistência da desigualdade, a hipótese da convergência absoluta deve ser testada antes de ser rejeitada.

Nesse sentido, a abordagem econométrica natural consiste em testar, inicialmente, a hipótese da convergência β incondicional através da regressão das taxas de crescimento das economias em questão sobre os respectivos níveis iniciais de renda. Um coeficiente negativo e significativo, neste contexto, indicaria que as economias mais pobres estariam crescendo mais rapidamente que as mais ricas, evidenciando a tendência a um nível comum de renda de longo prazo. A hipótese da convergência β condicional, por outro lado, implica que as economias com níveis atuais de renda mais distantes dos seus respectivos estados estacionários terão taxas de crescimento maiores que aquelas localizadas mais próximas a eles. Assim, o procedimento econométrico é uma extensão do anterior, no sentido em que também se baseia na observação do sinal do coeficiente da renda inicial, com a diferença que agora variáveis de controle são incluídas para representar o estado estacionário de cada economia. Já a convergência clube é caracterizada por grupos formados por economias com condições iniciais semelhantes, ou de outra forma, por economias pertencentes a uma mesma zona de atração. Essa zona de atração, por sua vez, é caracterizada por um único estado estacionário para o qual todos os membros tendem no longo prazo. Assim, dado que todas as economias de um clube possuem condições iniciais semelhantes e tendem a um mesmo nível de longo prazo, as distâncias dessas economias a esse estado estacionário comum também serão similares e, conseqüentemente, as taxas de crescimento serão igualmente semelhantes. Desta forma, o método utilizado deverá ser capaz de agrupar aquelas economias que possuam, ao mesmo tempo, condições iniciais e taxas de crescimento semelhantes.

Neste artigo investigamos a hipótese da convergência da renda *per capita* para os municípios brasileiros, procurando determinar qual hipótese de convergência – absoluta, condicional ou clube – caracteriza de forma mais adequada o crescimento desse grupo. Para tanto, testamos separadamente cada uma das hipóteses contra as demais. O procedimento pode, então, ser sintetizado em três etapas. Na primeira, assumimos que as condições iniciais não são relevantes e testamos a hipótese da convergência absoluta contra a hipótese da convergência condicional por meio da análise da significância das variáveis de controle. Em seguida, admitimos que as condições iniciais são de fato relevantes, mas supomos

²No caso da convergência condicional, choques afetam o *ranking* das rendas apenas no curto prazo, mas não tem efeito duradouro, visto que o estado estacionário só é alterado por mudanças em características estruturais. No segundo caso, entretanto, a dinâmica da renda de cada economia é caracterizada possivelmente por múltiplos estados estacionários, sendo que apenas um deles está localizado na chamada zona de atração definida pelas condições iniciais. Assim, um choque capaz de deslocar uma economia para uma zona de atração distinta, levará essa economia a um novo estado estacionário e, portanto, terá um efeito permanente sobre a posição dessa economia no *ranking*.

que as economias compartilham das mesmas características estruturais, testando, assim, a hipótese da convergência absoluta contra a hipótese da convergência clube. Para tanto, determinamos os possíveis clubes através do método denominado *regression tree* e permitimos que cada deles seja descrito uma dinâmica específica, para, então, verificar se as dinâmicas são, de fato, distintas. Por fim, relaxamos a hipótese de características estruturais comuns e testamos se as condições iniciais são relevantes, mesmo quando incluímos variáveis de controle, testando, desta forma, a hipótese da convergência condicional contra a hipótese da convergência clube.

Os resultados obtidos sugerem que a hipótese da convergência clube domina as demais hipóteses, evidenciando, assim, a importância das condições iniciais na determinação da renda de longo prazo dos municípios. Ademais, dentre as possíveis candidatas a determinantes dos clubes, as variáveis que se mostraram mais relevantes foram primeiramente aquelas relacionadas ao estoque de capital físico e às características demográficas dos municípios e, em segundo lugar, as variáveis associadas ao nível de estoque de capital humano.

O restante deste artigo está organizado da seguinte forma. A seção 2 revê alguns trabalhos anteriores sobre a hipótese da convergência. A seção 3 descreve a estrutura empírica utilizada, com destaque para o método de *regression tree*, utilizado aqui para a obtenção dos clubes e seleção das condições iniciais relevantes. A seção 4 detalha os dados utilizados, enquanto a seção 5 apresenta os resultados. As conclusões finais são apresentadas na seção 6.

2. REVISÃO DA LITERATURA

Uma revisão da literatura empírica que trata da hipótese da convergência revela uma grande variedade de objetos de estudo e metodologias com resultados igualmente variados.³ A esse respeito, a abordagem mais comumente encontrada consiste na regressão da taxa de crescimento sobre o nível inicial da renda e, freqüentemente, algumas variáveis de controle. A lista de trabalhos nessa linha é interminável, cabendo destacar os de Mankiw et alii (1992) e Sala-i Martin (1996).

Recentemente, algumas abordagens não-paramétricas vêm ganhando espaço devido à maior robustez dos resultados, e, principalmente, pela insatisfação com os métodos tradicionais.⁴ Seguindo a metodologia proposta originalmente por Quah (1993) e, posteriormente, aprimorada pelo mesmo autor em Quah (1996, 1997), diversos trabalhos investigaram a hipótese da convergência através da análise da dinâmica da distribuição da renda *per capita* dos países, utilizando a matriz de transição de Markov e/ou o núcleo estocástico. Registre-se aqui os artigos de Jones (1997) e Bianchi (1997) que analisam a hipótese da convergência da renda e confirmam o surgimento de picos gêmeos (tradução da expressão cunhada por Quah *emerging twin peaks*).

Prosseguindo na trilha dos métodos não-paramétricos para a análise da hipótese da convergência, cabe destacar a técnica denominada de classificação e clustering. Um dos trabalhos pioneiros nessa linha é o de Durlauf e Johnson (1995). Utilizando o procedimento de *regression tree* e as variáveis renda *per capita* e taxa de alfabetização como determinantes dos grupos, os autores encontram quatro clubes de países com dinâmicas de crescimento fundamentalmente distintas. Empregando essa mesma técnica, Johnson e Takeyama (2003) buscam determinar qual forma de convergência melhor descreve a dinâmica da renda dos estados norte-americanos, concluindo que a hipótese da convergência clube domina as demais.

No que se refere ao estudo do caso brasileiro, os trabalhos acerca da convergência têm focado principalmente na verificação da convergência $\hat{\alpha}$ (condicional e absoluta) e $\hat{\alpha}_0$, utilizando a abordagem *cross*

³Durlauf e Quah (1999) apresentam uma revisão de algumas abordagens à hipótese da convergência

⁴Os métodos paramétricos comumente utilizados na literatura somente captam o que está ocorrendo com a média da distribuição, em oposição ao que ocorre com a distribuição como um todo. Os métodos paramétricos exigem ainda a formulação, a priori, de hipóteses sobre a distribuição real da variável e/ou sobre o formato das funções utilizadas, ao contrário dos métodos não-paramétricos.



section tradicional para dados estaduais. Dentre esses estudos, podemos destacar os trabalhos de Ferreira e Diniz (1995), Ferreira (1996), Ferreira e Ellery Jr. (1996) e Azzoni (2001). Estudos empregando metodologias alternativas também são abundantes na análise da convergência para os estados brasileiros: Barossi-Filho e Azzoni (2003) utilizando séries temporais; Azzoni et alii (2000), dados em painel; Ferreira (2000), Mossi et alii (2003) e Gondim e Barreto (2004), analisando a dinâmica da distribuição da renda a partir da utilização de matrizes de transição de Markov e/ou núcleos estocásticos; e Azzoni e Silveira-Neto (2005) utilizando técnicas de econometria espacial.

No que tange a análise da convergência para os municípios brasileiros, objeto de estudo deste artigo, o número de trabalhos é significativamente inferior. Os primeiros trabalhos que abordam essa questão são relativamente recentes e se concentram na análise da convergência para uma determinada região do país. Este é o caso, por exemplo, dos artigos de Porto Junior e Ribeiro (2000), para os municípios da região Sul e Porto Junior e Souza (2002), para os municípios nordestinos. Ademais, dentre os trabalhos que estudam o conjunto dos municípios brasileiros em sua totalidade, cabe citar os artigos de Laurini et alii (2003), Andrade et alii (2004) e Gondim e Barreto (2004) que, empregando a análise da dinâmica da distribuição da renda e baseando-se no surgimento de duas modas nessa distribuição, concluem pela tendência ao surgimento de dois clubes na distribuição de longo prazo da renda: um de baixa renda formado essencialmente pelos municípios das regiões Norte e Nordeste, e outro de alta renda composto por municípios das regiões Centro-Oeste, Sudeste e Sul.

Primeiramente, cabe alertar que as conclusões apresentadas acima no que se refere à tendência ao surgimento de uma distribuição bimodal de longo prazo se basearam em evidências fornecidas pela densidade condicional estimada. Conforme, Johnson (2000), contudo, uma densidade condicional bimodal não implica necessariamente em bimodalidade na distribuição de longo prazo. O autor, então, após encontrar uma distribuição condicional multimodal para os estados norte-americanos, estima formalmente a distribuição ergódica que, por sua vez, se mostra fortemente unimodal. Em segundo lugar, deve-se ressaltar que mesmo uma distribuição ergódica bimodal não necessariamente caracteriza a existência de clubes conforme estes são definidos. Uma distribuição bimodal de longo prazo, a princípio, indica apenas a presença de diferenças persistentes no perfil das rendas dos municípios brasileiros, não evidenciando, porém, sua causa. Essas diferenças poderiam, inicialmente, ser atribuídas tanto a heterogeneidades nas características estruturais dos municípios como a condições iniciais distintas e, portanto, são compatíveis tanto com a hipótese da convergência condicional como com a hipótese da convergência clube. Assim, a análise da dinâmica da distribuição não parece ser a metodologia mais adequada no contexto da determinação da hipótese da convergência que melhor descreve a dinâmica da renda dos municípios brasileiros. Nesse sentido, este artigo propõe uma abordagem alternativa que nos habilita verificar a existência de diferenças permanentes nos níveis municipais de renda, e também, identificar se essas diferenças se devem somente a heterogeneidades de características estruturais ou também das condições iniciais.

3. ESTRATÉGIA EMPÍRICA

Conforme Galor (1996), as diferenças entre as três hipóteses da convergência, quais sejam, absoluta, condicional e clube, se resumem à dependência na homogeneidade das características estruturais e das condições iniciais para a existência de convergência das rendas. A convergência absoluta, primeiramente, partindo do pressuposto da homogeneidade das características estruturais, prevê a existência de um nível de renda de longo prazo comum. As hipóteses da convergência condicional e clube, por outro lado, não fazem essa suposição, mas requerem que as economias tenham características estruturais comuns para que atinjam o mesmo nível de renda de longo prazo. Para essa última, contudo, essa condição é apenas necessária, uma vez que exige ainda que as condições iniciais também sejam similares. Sendo assim, uma forma de se testar a hipótese da convergência clube contra as duas demais é verificar a relevância das condições iniciais na determinação dos níveis de renda de longo prazo dos

municípios.

A abordagem utilizada neste artigo para determinar a hipótese da convergência mais adequada para descrever a dinâmica da renda *per capita* dos municípios brasileiros consiste em testar isoladamente cada uma das hipóteses contra as demais. Utilizaremos, para tanto, a estrutura empírica proposta por Johnson e Takeyama (2003) que, partindo de resultados conhecidos do modelo neoclássico de crescimento, obtém a seguinte equação:

$$\Gamma_{it} = \alpha - \beta_t(\log(y_{i0}) - \log(y_{BR0})) + \Pi X_i + \epsilon_i \quad (1)$$

onde $\Gamma_{it} = \frac{\log(y_{it}) - \log(y_{i0})}{t} - \frac{\log(y_{BRt}) - \log(y_{BR0})}{t}$; $\beta_t = \frac{1 - e^{-\lambda t}}{t}$; y_{it} é a renda *per capita* real do município i no instante t e y_{BRt} é a renda *per capita* brasileira real no instante t . λ é a taxa de convergência para o estado estacionário, α é uma constante, ϵ_i é o termo do erro, X_i é o vetor das variáveis de controle e Π é o vetor de parâmetros associado ao vetor X_i . Por fim, Γ_{it} é a taxa de crescimento do município i entre os instantes 0 e t relativa a taxa de crescimento brasileira para o mesmo período.

Segundo os autores, cada uma das três hipóteses acerca da convergência pode ser testada a partir de alguma versão restrita da equação acima. Um sinal negativo para o coeficiente da renda inicial, ou de outra forma, um sinal positivo para β , quando fazemos α e β serem idênticos para todas as observações e, ao mesmo tempo, forçamos Π para zero, seria uma evidência a favor da convergência absoluta. Por outro lado, na situação em que relaxamos essa última restrição, permitido que Π seja diferente de zero, um sinal positivo de β e a significância de alguma variável de controle corroborariam a hipótese da convergência condicional em detrimento da hipótese da convergência absoluta. Por fim, ao permitirmos que α e β assumam valores distintos para cada um dos possíveis clubes, estamos admitindo uma dinâmica específica para cada um deles, o que torna possível testar a hipótese da convergência clube contra as demais. Se supomos que os municípios compartilham das mesmas características estruturais, ou seja, se fazemos Π igual a zero na equação acima, podemos testar a hipótese da convergência clube contra a hipótese da convergência absoluta. Em contraste, ao abandonarmos essa suposição, estamos testando a hipótese da convergência clube contra a hipótese da convergência condicional.

A questão que fica, todavia, é como determinar os clubes de convergência, ou em outras palavras, qual o melhor critério a ser utilizado para agrupar os municípios. Mora (2005), por exemplo, utiliza uma medida de desigualdade obtida a partir da decomposição do índice de Gini para agrupar as economias. No nosso caso, como já mencionado, necessitamos de um método capaz de agrupar municípios com taxas de crescimento semelhantes e que pertençam a uma mesma zona de atração, a qual, por sua vez, é definida por um conjunto de condições iniciais. De forma a atender a estes requisitos, empregamos o método proposto originalmente no trabalho de Durlauf e Johnson (1995) para investigar a questão da convergência para os países do mundo denominado *regression tree*. A descrição dessa técnica se fundamenta naquela apresentada por Breiman (1993) e Durlauf e Johnson (1995). Nesse último, o algoritmo de *regression tree* é apresentado como um procedimento composto por uma sequência de quatro etapas a seguir descritas:

1. Seja $x_i, i = 1, \dots, I$ as variáveis representando as condições iniciais e y a taxa de crescimento de cada município. A observação relativa ao município será, assim, um vetor composto pela taxa de crescimento desse município e pelas respectivas variáveis representando as condições iniciais, $(y_j, x_{1j}, \dots, x_{Ij})$. Para cada uma das condições iniciais, classificamos os dados em ordem crescente e particionamos as observações em dois subgrupos de acordo com a regra: atribuímos a observação j ao conjunto $S_{(a,i)} \text{ se } x_{i,j} < a$, e, caso contrário, ao conjunto $S_{(a',i)}$. Variando o valor de “a” ao longo de toda faixa de valores que x_i pode assumir, obtemos todas as partições baseadas nessa variável.⁵ Fazendo isto para todas as variáveis explicativas, obtemos todas as partições possíveis. A melhor partição consistirá em um conjunto formado por uma variável x_i

⁵A rigor, “a” só assumirá valores iguais a média de duas observações subseqüentes, ou seja, $a = (x_{i,j} + x_{i,j+1})/2$, para $i = 1, \dots, I - 1$.



e algum valor de “a” que minimizem a seguinte soma dos quadrados dos resíduos (SQR) com relação a todas as partições possíveis:

$$\sum_{j \in S(a,i)} (y_j - \bar{y}_{(a,i)})^2 + \sum_{j \in S(a',i)} (y_j - \bar{y}_{(a',i)})^2 \quad (2)$$

onde $\bar{y}_{(a,i)}$ e $\bar{y}_{(a',i)}$ são os valores médios da variável para cada uma das novas partições. Iremos nos referir a essas partições a partir de agora como nós.

2. Repetindo o passo 1 para cada um desses nós, encontramos novamente alguma variável x_i e algum valor “a” que minimizam (1), obtendo 4 novos nós. Esse procedimento segue para os nós resultantes até que todas as observações de determinado nó possuam valores idênticos de y ou quando o número de observações em cada nó for inferior a determinado limite estipulado previamente. Neste caso, os nós serão denominados nós terminais e o número total de nós será representado por T . Se definirmos, então, T_L como o conjunto de todos os nós terminais, a soma dos quadrados dos resíduos da árvore, SQR_{Tree} será dada por:

$$SQR_{Tree} = \sum_{S_m \in T_L} \sum_{j \in S(a,i)} (y_j - \bar{y}_{S_m})^2 \quad (3)$$

A grandeza acima não possui uma interpretação natural e ainda depende da escala da variável dependente. Para contornar esse problema, utilizamos a seguinte normalização de (2), denominada erro médio quadrático relativo, EMQ :

$$EMQ = SQR_{Tree} / SQR_{Total} \quad (4)$$

onde $SQR_{Total} = \sum_j (y_j - \bar{y})^2$.

EMQ varia de 1, para a situação em que não foi realizada nenhuma partição, ou seja, quando a árvore se resume a apenas um nó contendo todos municípios e, tende assintoticamente para zero à medida que o número de partições aumenta.

3. O modelo obtido a partir dos passos 1 e 2 possui um número excessivo de nós, uma vez que não atribuímos um custo ao número de separações e expandimos a árvore até que algum critério que determina um nó terminal fosse atingido. Devemos, portanto, podar essa árvore, levando em consideração um custo associado ao número de partições. Seja este custo igual a $\alpha \cdot T$, onde α é uma constante associada ao custo de um número excessivo de nós. Utilizando, então, $\alpha \cdot T + SQR_{Tree}$ como critério para seleção da árvore ótima, podemos fazer o caminho de baixo para cima a partir da árvore original, removendo todas as partições que reduzam a soma acima, produzindo uma nova árvore. Este procedimento produzirá, então, uma árvore com $T * (\alpha)$ nós terminais associado a determinado valor de α . Fazendo α variar no intervalo $0 \leq \alpha \leq \infty$, obtemos uma sequência de k árvores, T_k , sendo cada uma delas uma versão reduzida da árvore original, que é um caso particular quando α é igual a 0. No outro extremo, por outro lado, temos a “árvore” obtida quando α tende para o infinito, que nada mais é que um nó contendo todas as observações.
4. Para escolhermos a melhor árvore dentre o conjunto obtido acima, precisamos de um estimador honesto de SQR_{Tree} . Uma forma consiste em dividir as observações aleatoriamente em dois grupos. A partir do primeiro grupo, denominado grupo de aprendizagem, utilizando os passos de 1 a 3, obtemos a sequência de árvores ótimas associadas aos distintos valores de α . Em seguida, utilizando o segundo conjunto de observações, chamado de conjunto teste, estimamos SQR_{Tree} para cada uma das árvores da seguinte forma:

$$SQR_{Tree} = \sum_j (y_j - d_k(\vec{x}_j))^2 \quad (5)$$

onde $d_k(\vec{x}_j)$ é o valor estimado de y_j pela k -ésima árvore.

A melhor árvore será aquela que estiver associada ao menor valor de SQR_{Tree} calculado pelas observações do segundo grupo.

A opção empregada neste artigo, por outro lado, consiste na utilização do método de validação cruzada. A escolha desse método se deve a maior robustez de seus resultados relativamente ao critério anterior.⁶ Esse método parte da divisão das observações aleatoriamente em subgrupos de tamanhos semelhantes. Chamaremos cada um desses conjuntos de $L_v, v = 1, \dots, V$, e o conjunto com todas as observações de L .⁷ Definiremos ainda $L^v = L - L_v$. Utilizando os passos 1 a 3, e cada um dos subconjuntos L^v , obtemos as árvores ótimas associadas a todos os possíveis valores de α . Denominaremos as árvores ótimas obtidas a partir das observações L^v associada a determinado valor de α de $T^v(\alpha)$. Seja α_k o valor de α associado à k -ésima árvore, T_k , obtida a partir de L e seja α'_k a média geométrica de α_k e α_{k+1} . O valor do erro médio quadrático relativo estimado pelo método de validação cruzada da árvore T_k , $EMQ^{cv}(T_k)$, será dado por:

$$EMQ^{cv}(T_k) = \frac{\sum_v \sum_{L^v} (y_j - d_k^{(v)}(\vec{x}))^2}{\sum_j (y_j - \bar{y})^2} \quad (6)$$

onde $d_k^{(v)}(\vec{x})$ é o estimador baseado na árvore $T^{(v)}(\alpha'_k)$.

Neste caso, a melhor árvore será aquela que possuir o menor valor de EMQ^{cv} .

Neste ponto, cabe registrar que a metodologia apresentada é apenas uma das possíveis opções para a seleção dos clubes. Uma alternativa interessante para trabalhos futuros e que se mostra superior ao *regression tree*, especialmente no que tange o poder de previsão e o esforço computacional empregado, é proposta por da Rosa et alii (2003). Este método, denominado *Smooth Transition Regression Tree (STR Tree)*, se assemelha ao *regression tree* em diversos aspectos, porém apresenta uma particularidade que o torna bastante atrativo. Em contraste com o método empregado neste artigo, o STR Tree não separa necessariamente os elementos em grupos mutuamente excludentes, de sorte que cada elemento pode pertencer a mais de um conjunto, sendo seu grau de associação com cada um desses variado. Assim, dependendo do valor do parâmetro de suavização utilizado, teremos situações que variam desde o caso onde todos os elementos pertencem a um mesmo grupo, passando pela situação onde os elementos pertencem aos diversos grupos com graus de associação diferentes, chegando ao outro extremo, onde cada elemento é alocado em apenas um conjunto, que é exatamente a separação que seria obtida caso o método de *regression tree* fosse utilizado.

4. DADOS

Um estudo envolvendo variáveis no nível de agregação municipal deve ter o cuidado de considerar grandezas que sejam comparáveis intertemporalmente. Neste artigo, estamos interessados no período de 1970 a 2000, quando o número de municípios brasileiros aumentou de 3.951 para 5.507. Nesse sentido, uma análise que pretenda fazer alguma inferência acerca da evolução da renda municipal deve

⁶ Os resultados obtidos a partir da utilização do critério baseado na separação das observações em um grupo de aprendizagem e outro de teste se mostraram muito sensíveis quando o tamanho do grupo de teste era pequeno. À medida que esse tamanho é aumentado, contudo, os resultados se tornam mais robustos e tendem para o resultado obtido quando utilizamos o critério de validação cruzada.

⁷ No caso específico deste artigo, dividimos o conjunto de observações em 10 subgrupos, conforme proposto por Breiman (1993).



levar essa questão em consideração, pois, caso contrário, seus resultados serão, no mínimo, distorcidos. O problema, contudo, não se restringe a agregar os novos municípios a outro, já que, em alguns casos, um município teve sua origem em dois ou mais municípios. Devemos, portanto, trabalhar com o conceito de áreas mínimas comparáveis⁸ (AMC) que são, por definição, as áreas geográficas mais desagregadas possíveis que podem ser comparadas entre dois pontos no tempo. Para o período em questão, o número de AMC é de 3.659, inferior, portanto, ao número de municípios existentes em 1970, demonstrando claramente que alguns municípios tiveram sua origem em mais de um município.⁹

As variáveis utilizadas neste artigo para representar tanto as candidatas a condições iniciais bem como as possíveis variáveis de controle abrangem diversas características socioeconômicas e demográficas dos municípios para o ano de 1970,¹⁰ focando, principalmente, naquelas relacionadas aos estoques de capital físico e humano. Essa escolha, contudo, foi orientada, em grande parte, pela disponibilidade dos dados para o ano em questão no formato descrito acima, de sorte que algumas variáveis refletindo características possivelmente relevantes não foram levadas em consideração nesta análise. Por outro lado, tendo em vista a infinidade de possibilidades no que se refere às candidatas a variável de controle e também a condição inicial encontradas na literatura teórica de crescimento,¹¹ não fazemos qualquer suposição quanto às melhores variáveis para representá-las e deixamos aos métodos a incumbência de selecioná-las. A tabela 1 descreve cada uma das variáveis selecionadas, apresentando a unidade de medida de cada uma delas, o ano da observação, assim como a média municipal e o coeficiente de variação.

5. RESULTADOS

Nesta seção, procuramos evidências a favor de uma das três hipóteses da convergência, testando separadamente cada uma delas contra as demais, o que sugere que a análise seja dividida em três partes. Na primeira delas, assumimos que as condições iniciais são irrelevantes na determinação dos níveis de renda de longo prazo dos municípios e testamos a hipótese da convergência absoluta contra a hipótese da convergência condicional. Em seguida, supomos que todos os municípios possuem características estruturais comuns, mas diferem quanto às condições iniciais, podendo, assim confrontar as hipóteses da convergência à absoluta e clube. Na terceira parte, por fim, relaxamos a suposição de características estruturais comuns e testamos a hipótese da convergência condicional contra a hipótese da convergência clube.

5.1. Absoluta vs. Condicional

Iniciamos nossa análise estimando (1) com a restrição de $\Pi = 0$ e assumindo que α e β sejam idênticos para todos municípios. A primeira coluna da tabela 2 apresenta o resultado dessa regressão. O valor negativo e significativo do coeficiente do nível inicial da renda indica que, na média, os municípios mais pobres cresceram mais rapidamente que os mais ricos, corroborando, assim, a hipótese da convergência absoluta.

O passo seguinte consiste em relaxar a restrição $\Pi = 0$ e estimar (1) incluindo separadamente cada uma das variáveis de controle em nível e na forma logarítmica. Estes resultados são apresentados na tabela 3, de onde se observa que, das 33 variáveis testadas, 29 apresentaram coeficientes significativos ao nível de 5% em pelo menos uma das formas. Ademais, o coeficiente da renda inicial se manteve

⁸Os dados utilizados neste artigo foram obtidos do IPEADATA (<http://www.ipeadata.gov.br>).

⁹Neste artigo, utilizaremos o termo município quando estaremos, de fato, nos referindo a AMC.

¹⁰As variáveis “capital humano *per capita*” e aquelas relacionadas à violência referem-se ao ano de 1980 devido à indisponibilidade de dados para 1970.

¹¹Ver Durlauf et alii (2004) para uma síntese das diversas alternativas de determinantes de crescimento já propostas.

negativo e significativo em todas as situações e os coeficientes estimados para as variáveis de controle apresentaram os sinais esperados, cabendo registrar algumas exceções, notadamente aquelas relacionadas ao nível de violência e a algumas características demográficas.

Ao estimarmos as regressões incluindo cada variável de controle separadamente, estamos sujeitos a problemas associados à omissão de variáveis, o que, em certo grau, compromete a análise acima. Neste caso, contudo, a omissão de variáveis é a situação mais provável, uma vez que a teoria econômica não explicita a “verdadeira” equação de crescimento e, portanto, não identifica quais variáveis explicativas compõe essa equação. Isto, por sua vez, credencia os pesquisadores a simplesmente experimentar todas aquelas variáveis que podem afetar a taxa de crescimento, gerando, conseqüentemente, uma série infundável de trabalhos, cada qual defendendo a utilização de determinadas variáveis. Mesmo que o pesquisador inclua um grande número de variáveis de forma a evitar a omissão de variáveis, ao realizar algumas regressões combinando as variáveis disponíveis, ele se deparará com situações onde uma variável que era significativa em determinada regressão, deixa de sê-la quando incluímos ou excluimos algumas variáveis.

Uma forma de contornar o problema descrito acima consiste na utilização dos chamados testes de robustez. Os testes que utilizaremos aqui e que descreveremos brevemente se baseiam nos trabalhos de Levine e Renelt (1992) e Sala-i Martin (1997). O procedimento comum a ambos se fundamenta na realização de regressões da seguinte forma:

$$\Gamma = \alpha + \beta_{yi}Y + \beta_{zi}Z + \beta_{xi}X + \epsilon \quad (7)$$

onde Γ é novamente a taxa de crescimento, Z é a variável cuja robustez queremos testar, Y é o conjunto de variáveis comum a todas as regressões e X é um grupo formado normalmente por três variáveis extraídas de um conjunto de tamanho N .¹²

Ambos os testes se baseiam nos valores dos coeficientes β_z e nos seus respectivos desvios-padrão obtidos nas diversas regressões realizadas, mas diferem quanto ao critério utilizado para caracterizar a robustez da variável. O primeiro, proposto por Levine e Renelt, se fundamenta na técnica denominada *extreme bounds analysis* – EBA. Segundo esse critério, uma variável é dita robusta se seus limites inferior e superior possuem o mesmo sinal, sendo o limite inferior dado pelo menor coeficiente significativo subtraído de duas vezes seu desvio-padrão, $\beta_{zi} - 2\sigma_{zi}$, e o limite superior, pela soma do maior coeficiente significativo a duas vezes seu desvio, $\beta_{zi} + 2\sigma_{zi}$. O teste proposto por Sala-i-Martin, por outro lado, se baseia na análise de toda distribuição do estimador β_z e, em uma de suas versões, se resume a verificação da significância da seguinte estatística:

$$t = \frac{\sum_i \beta_{zi}}{\frac{\sqrt{\sum_i \sigma_{zi}^2}}{M}} \quad (8)$$

onde M é o número de regressões realizadas o teste da robustez de cada variável.

Neste caso, uma variável será considerada robusta se a estatística t for superior a determinado valor crítico estabelecido com base na significância do teste.

A tabela 4 apresenta os resultados de ambos os testes para as 33 variáveis de controle em nível e log.¹³ As colunas “Valor Médio” e “Desvio Padrão” referem-se respectivamente ao numerador e ao denominador de (8), enquanto as colunas “Limite Inferior” e “Limite Superior” referem-se aos resultados do teste EBA. Como se pode observar, com base no teste EBA, apenas a variável associada à defasagem

¹² Repare que neste caso onde X é composto por três variáveis, o procedimento implicará a realização de $\frac{N!}{(N-3)!3!}$ regressões para cada variável testada.

¹³ Note que isso implica em 66 variáveis e, conseqüentemente, em um total de 2.882.880 regressões.



escolar pode ser considerada robusta. Os resultados do segundo teste, por outro lado, caracterizam 19 variáveis (em nível e/ou log) como robustas.¹⁴

Desta forma, com base nos resultados obtidos anteriormente quando testamos isoladamente cada uma das variáveis de controle, juntamente com os resultados obtidos a partir dos dois testes de robustez, podemos concluir pela relevância das variáveis de controle na estimação da equação 1 e, conseqüentemente, pela existência de diferenças significativas nas características estruturais dos municípios brasileiros o que, por sua vez, demonstra a superioridade da hipótese da convergência condicional sobre a hipótese da convergência absoluta.

6. ABSOLUTA VS. CLUBE

Nesta subseção, assumiremos que os municípios possuem um conjunto homogêneo de características estruturais, podendo, assim, fazer $\Pi = 0$ em (1). Permitiremos, por outro lado, que cada clube seja descrito por um modelo linear distinto, em outras palavras, relaxaremos a restrição de α e β idênticos para todo conjunto de municípios.

Dado que já estimamos na subseção anterior a regressão restrita relativa à hipótese da convergência absoluta, a etapa seguinte é determinar os clubes. Conforme já adiantado, o procedimento adotado neste artigo consiste em aplicar o método de *regression tree*. A figura 1 apresenta a árvore ótima obtida a partir dessa técnica. Sua interpretação não é direta, necessitando que algumas diretrizes sejam fornecidas. Acima de cada nó intermediário, os nomes e os níveis de corte das variáveis utilizadas para fazer a partição são apresentados. Desses nós partem duas ramificações, uma para esquerda, referente aos municípios com níveis iniciais da variável em questão abaixo do nível de corte, e outro para a direita, relativo aos municípios com nível superior a esse valor de corte. Abaixo dos nós terminais, apresentam-se os valores de taxa de crescimento médio de cada clube. Registre-se, por fim, que iremos nos referir aos nós terminais ou clubes por números, sendo o nó mais à esquerda o clube de número 1, o nó à direita desse, o número 2 e assim por diante.

Isto posto, podemos passar à descrição dos resultados obtidos a partir desse experimento (Figura 1). Em primeiro lugar, os resultados demonstram a importância das variáveis relacionadas aos estoques de capital físico e humano na determinação dos clubes, tendo cada uma delas uma relação distinta com a taxa de crescimento. Enquanto na maioria dos casos, os clubes com menores estoques de capital físico apresentaram as maiores taxas de crescimento, foram os clubes com maiores níveis iniciais de capital humano aqueles que mais cresceram. Além disso, a proporção da população urbana também se mostrou significativa na caracterização dos clubes com maiores estoques de capital residencial, sendo os municípios com a menor fração aqueles que apresentaram o maior crescimento. Isto poderia estar indicando que nos municípios maiores (maiores estoques de capital), as externalidades negativas originadas pela aglomeração estariam prevalecendo em relação as externalidades positivas. Destaque-se, por fim, a amplitude das taxas de crescimento dos 8 clubes para todo o período, variando de 0,5% abaixo da média nacional a 2,1% acima.

Uma vez determinados os clubes, podemos estimar o modelo linear seguido por cada um. Isto será feito a partir da estimação da seguinte versão da equação 1:

$$\Gamma_{it} = \alpha + \sum_{j=2}^8 \alpha_j \cdot D_{ij} - (\beta_t + \sum_{j=2}^8 \beta_{tj} \cdot D_{ij})(\log(y_{i0}) - \log(y_{BR0})) + \epsilon_i \quad (9)$$

onde $D_{ij} = 1$ se o município i pertencer ao clube j , para $j = 2, \dots, 8$.

¹⁴A discrepância entre os resultados dos dois testes se deve principalmente ao fato do primeiro ser, de fato, extremo, uma vez que, havendo uma única mudança de sinal no coeficiente da variável testada em uma das 43.680 regressões realizadas, ela será considerada não-robusta.

A equação 9 permite testar a existência de distintos interceptos e distintas sensibilidades em relação à renda inicial de cada grupo.

A tabela 2 apresenta a estimação dessa equação. Apenas a título de exercício, são exibidos nas colunas 2 e 3 os resultados quando restringimos que apenas o intercepto seja comum ou que apenas a inclinação seja comum para todos os clubes. A quarta coluna apresenta os resultados quando relaxamos as duas restrições e permitimos que cada clube possua intercepto e inclinação distintos dos demais. Inferências acerca da significância desses coeficientes obtidos não são diretas. Conforme Johnson e Takeyama (2003), o método utilizado para determinação dos clubes e, conseqüentemente, das *dummies*, afetou a eficiência dos estimadores, tornando essa análise comprometida, uma vez que a significância dos coeficientes dessas variáveis será artificialmente alta. Em outras palavras, isso quer dizer que o erro-padrão estimado é certamente inferior ao valor verdadeiro, de sorte que qualquer coeficiente que seja considerado não significativo com base nos erros-padrão estimados também será não significativo quando os erros-padrão apropriados forem utilizados. Assim, utilizando esse raciocínio, podemos afirmar que os coeficientes α_2 , β_2 e β_7 não são estatisticamente diferentes de zero.¹⁵ Quando estimamos a nova equação restringindo os coeficientes acima para zero, vemos que também os coeficientes α_5 , α_7 , e β_5 não são significativos.¹⁶ Neste sentido, podemos afirmar que, apesar de possuírem características iniciais e taxas de crescimento distintas, os clubes 1, 2, 5 e 7, compartilham um modelo linear comum. Necessitamos, agora, verificar se os clubes 3, 4, 6 e 8 também possuem interceptos ou inclinações comuns. Para tanto, inicialmente somamos as *dummies* relativas aos clubes 1, 2, 5 e 7, obtendo uma *dummy* que é igual a um para todos os municípios que foram originalmente classificados nesses clubes. Em seguida, estimamos (9), incluindo essa nova *dummy* e também as *dummies* dos clubes 3, 4, 6 e 8, deixando sempre uma delas de fora de forma a evitar a multicolinearidade perfeita. Fazendo isto, vemos que as inclinações destes quatro clubes não são significativamente diferentes uma das outras e, nem os interceptos dos grupos 4 e 6 entre si. A tabela 5 resume os resultados obtidos, de onde se observa em todos os casos um sinal negativo para o nível inicial de renda. Esse fato conjuntamente com a significância de algumas das variáveis *dummies* são considerados como evidências a favor da hipótese da convergência clube.

Antes de passar a subseção seguinte, é interessante notar a relação existente entre a localização geográfica de cada município e o modelo linear que descreve a dinâmica de sua renda. A tabela 6 demonstra esse fato. Dessa tabela, percebe-se que, enquanto 84,7% dos municípios da região Sul, 69,1% da região Sudeste e 85,2% da Centro-Oeste pertencem ao conjunto formado pelos clubes 4 e 6, apenas 12,7% dos municípios da região Nordeste e 39,2% da região Norte pertencem a esses mesmos clubes. Por outro lado, 81,6% dos membros dos clubes 1, 2, 5 e 7 são municípios nordestinos e 74,1% do clube 8 são municípios da região Sudeste. O mapa da figura 2 ilustra esses fatos. Note que se agruparmos os municípios membros do clube 3 àqueles dos clubes 1, 2, 5 e 7 e, da mesma forma, os municípios do clube 8 àqueles dos clubes 4 e 6, obteremos, um resultado muito semelhante àquele encontrado por Laurini et alii (2003) e Andrade et alii (2004), qual seja, da formação de dois clubes de convergência de renda para os municípios brasileiros, sendo o de baixa renda composto pelos municípios das regiões Nordeste e Norte, e o de renda mais elevada pelos municípios das regiões Centro-Oeste, Sudeste e Sul. Esse resultado adicionado ao fato dos coeficientes do nível inicial de renda serem negativos para todos os clubes sugere que a desigualdade existente se deve em grande parte às diferenças de renda entre as regiões e, em menor grau, às diferenças dentro de cada uma das regiões.

¹⁵ A estatística do teste de Wald da hipótese $\alpha_2 = \beta_2 = \beta_7 = 0$ é igual 3,21 e tem p-valor associado de 0,021.

¹⁶ A estatística do teste de Wald da hipótese $\alpha_2 = \alpha_5 = \alpha_7 = \beta_2 = \beta_5 = \beta_7 = 0$ é igual 3,08 e tem p-valor associado de 0,005.



6.1. Condicional vs. Clube

Nas subseções anteriores, encontramos evidências a favor da rejeição da hipótese da convergência absoluta quando a confrontamos com as hipóteses alternativas. No primeiro embate, os testes demonstraram a importância das variáveis de controle na modelagem das diferentes características estruturais de cada município, subsidiando a escolha da hipótese da convergência condicional em detrimento da hipótese de convergência absoluta. No segundo, mostramos a relevância das condições iniciais na determinação dos clubes e a convergência β presente em cada um dos clubes, resultado que corroborou a opção pela hipótese da convergência clube. Resta-nos, agora, colocar as hipóteses da convergência condicional e clube frente a frente e determinar qual delas melhor descreve a dinâmica da renda dos municípios brasileiros. Para isso, começaremos testando se os coeficientes das *dummies* determinadas na subseção anterior permanecem significativos quando incluímos as 29 variáveis de controle obtidas na subseção 5.1. Os resultados desse experimento são apresentados na tabela 7, onde a primeira coluna exibe os valores da estatística de Wald para o teste de que os coeficientes das *dummies* são todos iguais a zero, na presença de cada uma das variáveis de controle (em nível e log) e, na segunda coluna, os valores dos coeficientes das variáveis de controle. Uma observação dessa tabela deixa clara a relevância das condições iniciais, já que, em todas as situações em que foram testadas, as *dummies* se mostraram fortemente significativas tanto de forma conjunta como individualmente. Mais ainda, vemos que o coeficiente da renda inicial permanece negativo e significativo em todas as situações e que, dentre as 29 variáveis testadas, 5 deixaram de ser significativas após a inclusão das *dummies*.

Vimos anteriormente que ao incluímos cada uma das variáveis de controle isoladamente poderíamos estar incorrendo em omissão de variáveis, fato esse que justificou a utilização dos testes de robustez. O teste EBA indicou que apenas a variável associada à defasagem escolar era robusta e, sendo essa uma das variáveis já testadas isoladamente, não precisamos realizar mais esse teste. Os resultados obtidos para essa situação demonstram, portanto, a significância dos coeficientes das *dummies* assim como o dessa variável de controle. Se, por outro lado, nos basearmos no teste de robustez proposto por Sala-i-Martin, devemos testar a significância dos coeficientes das *dummies* na presença de todas as variáveis consideradas robustas ao mesmo tempo. O resultado desse teste demonstra que todas as *dummies* apresentam coeficientes significativos conjuntamente¹⁷ e individualmente, com exceção do coeficiente de intercepto para o clube 3. Além disso, das 19 variáveis que se mostraram robustas anteriormente com base nesse teste, 13 se tornaram não-significativas quando avaliadas individualmente.

Desta forma, consideramos que os resultados obtidos nessa subseção demonstram a importância das condições iniciais e são, portanto, mais consistentes com a hipótese da convergência clube, apesar de diversas variáveis de controle terem permanecido significativas mesmo após a inclusão das variáveis *dummy*. Esse fato demonstra a persistência de diferenças nas características estruturais dos municípios que não desaparecem mesmo quando controlamos para as condições iniciais. Ressalte-se, contudo, que essa dependência da taxa de crescimento com relação às variáveis de controle é totalmente compatível com a hipótese da convergência clube. A evidência a favor das condições iniciais, por outro lado, não é acolhida pela hipótese da convergência condicional, fato que nos leva a concluir pela dominância da hipótese da convergência clube sobre a hipótese da convergência condicional.

7. CONCLUSÃO

Utilizando a metodologia proposta por Johnson e Takeyama (2003), buscamos determinar qual das hipóteses da convergência – absoluta, condicional ou clube – descreve de maneira mais adequada a dinâmica da renda dos municípios brasileiros. Os resultados obtidos demonstram a importância das condições iniciais na determinação do nível de renda de longo prazo, o que sugere a dominância da

¹⁷A estatística do teste de Wald para a hipótese que todos os coeficientes das *dummies* são conjuntamente iguais a 0 é igual 46,2 com p-valor associado muito próximo a 0.

hipótese da convergência clube sobre as demais. Esse resultado, por sua vez, implica a existência de zonas de atração e múltiplos estados estacionários na dinâmica da renda desses municípios e, por conseguinte, na possibilidade de choques não-estruturais terem efeitos permanentes sobre o nível de renda de longo prazo dos municípios.

Ademais, os resultados obtidos atribuem um papel central na determinação dos clubes aos níveis iniciais de capital físico e humano, e são, portanto, consistentes com modelos teóricos que prevêem as chamadas armadilhas de pobreza. Essa relação existente entre o nível inicial de capital físico e a formação dos clubes é bem acolhida no modelo proposto por Feyrer (2003). Nesse modelo, partindo da suposição de que países mais ricos têm uma taxa de fecundidade menor, Feyrer obtém um equilíbrio estável composto por dois clubes: o primeiro constituído por economias com níveis iniciais de capital físico mais elevados e tendendo para um nível de renda de longo prazo superior; e o segundo, com níveis iniciais de capital físico abaixo de determinado valor crítico e tendendo ao estado estacionário inferior. A justificativa para a dependência com o nível inicial de capital humano, por outro lado, pode ser atribuída à existência de externalidades na acumulação desse fator. Azariadis e Drazen (1990), por exemplo, supõe que exista uma relação crescente entre o estoque de capital humano de uma economia e os efeitos das externalidades associadas a esse fator, de sorte que economias com níveis iniciais de capital humano mais elevados tendem a crescer a taxas maiores que aquelas com níveis mais baixos.

Os resultados demonstram ainda o caráter regional dos clubes visto que o clube de baixa renda é formado predominantemente pelos municípios das regiões Norte e Nordeste, o outro, com nível de renda mais elevado, é composto basicamente pelos municípios das regiões Centro-Oeste, Sudeste e Sul. Mais do que isso, o sinal negativo encontrado para o coeficiente do nível inicial de renda para cada um dos clubes parece demonstrar que o surgimento desses clubes se deve em grande parte a desigualdades existentes entre as regiões e não dentro das regiões. Esses resultados são consistentes com aqueles obtidos nos trabalhos de Laurini et alii (2003), Andrade et alii (2004) e Gondim e Barreto (2004). Além disso, deixam claras as diferenças existentes entre as dinâmicas das rendas desses clubes, uma vez que são descritas por modelos lineares essencialmente distintos. Esse fato, por sua vez, vai de encontro à hipótese comumente assumida de um modelo único para todos os municípios e, portanto, sugere a adoção de modelos compatíveis com múltiplos regimes ao se buscar descrever a dinâmica da renda dos municípios brasileiros.

Referências Bibliográficas

- Andrade, E., Laurini, M., Madalozzo, R., & Valls Pereira, P. L. (2004). Convergence clubs among brazilian municipalities. *Economics Letters*, 83(2):179–184. available at <http://ideas.repec.org/a/eee/econl/v83y2004i2p179-184.html>.
- Azariadis, C. & Drazen, A. (1990). Threshold externalities in economic development. *The Quarterly Journal of Economics*, 105(2):501–26. available at <http://ideas.repec.org/a/tpr/qjecon/v105y1990i2p501-26.html>.
- Azzoni, C. & Silveira-Neto, R. (2005). Decomposing regional growth: Labor force participation rates, structural changes, and sectoral factor reallocation. *The Annals of Regional Science*, 39(2):221–239. available at <http://ideas.repec.org/a/spr/anresc/v39y2005i2p221-239.html>.
- Azzoni, C. R. (2001). Economic growth and regional income inequality in brazil. *The Annals of Regional Science*, 35(1):133–152. available at <http://ideas.repec.org/a/spr/anresc/v35y2001i1p133-152.html>.
- Azzoni, C. R., Menezes-Filho, N., Menezes, T., & Silveira-Neto, R. (2000). Geography and income convergence among brazilian states washington. Disponível em <http://www.iadb.org/res/publications/pubfiles/pubR-395.pdf>. Acesso em 30 jun. 2006.



- Barossi-Filho, M. & Azzoni, C. R. (2003). A time series analysis of regional income convergence in Brazil São Paulo. Disponível em http://www.econ.fea.usp.br/nereus/td/Nereus_09_03.pdf. Acesso em 3 jul. 2006.
- Bianchi, M. (1997). Testing for convergence: Evidence from non-parametric multimodality tests. *Journal of Applied Econometrics*, 12(4):393–409. available at <http://ideas.repec.org/a/jae/japmet/v12y1997i4p393-409.html>.
- Breiman, L. (1993). *Classification and Regression Trees*. Chapman & Hall, New York.
- da Rosa, J. C., Álvaro Veiga, & Medeiros, M. C. (2003). Three-structured smooth transition regression models based on cart algorithm. Textos para discussão 469, Department of Economics PUC-Rio (Brazil). available at <http://ideas.repec.org/p/rio/texdis/469.html>.
- Durlauf, S. N. & Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics*, 10(4):365–84. available at <http://ideas.repec.org/a/jae/japmet/v10y1995i4p365-84.html>.
- Durlauf, S. N., Johnson, P. A., & Temple, J. R. W. (2004). Growth econometrics poughkeepsie. Disponível em <http://irving.vassar.edu/VCEWP/VCEWP61.pdf>. Acesso em 14 abr. 2006.
- Durlauf, S. N. & Quah, D. T. (1999). The new empirics of economic growth. In Taylor, J. B. & Woodford, M., editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 4, pages 235–308. Elsevier. available at <http://ideas.repec.org/h/eee/macchp/1-04.html>.
- Ferreira, A. (1996). Evolução recente das rendas per capita estaduais no Brasil: o que a nova evidência mostra. *Revista Econômica do Nordeste*, 27(3):363–374.
- Ferreira, A. (2000). Convergence in Brazil: recent trends and long-run prospects. *Applied Economics*, 32(4):479–489.
- Ferreira, A. & Diniz, C. (1995). Convergência entre as rendas per capita estaduais no Brasil. *Revista de Economia Política*, 15(4).
- Ferreira, A. & Ellery Jr., R. (1996). Convergência entre a renda per capita dos estados brasileiros. *Revista de Econometria*, 16(1):83–104.
- Feyrer, J. (2003). Convergence by parts. Disponível em <http://www.dartmouth.edu/~jfeyrer/parts.pdf>. Acesso em 29 mai. 2006.
- Galor, O. (1996). Convergence? inferences from theoretical models. *Economic Journal*, 106(437):1056–69. available at <http://ideas.repec.org/a/ecj/econjl/v106y1996i437p1056-69.html>.
- Gondim, J. L. B. & Barreto, F. A. (2004). O uso do núcleo estocástico para identificação de clubes de convergência entre estados e municípios brasileiros. Anais do XXXII Encontro Nacional de Economia [Proceedings of the 32th Brazilian Economics Meeting] 053, ANPEC - Associação Nacional dos Centros de Pósgraduação em Economia [Brazilian Association of Graduate Programs in Economics]. available at <http://ideas.repec.org/p/anp/en2004/053.html>.
- Johnson, P. A. (2000). A nonparametric analysis of income convergence across the US states. *Economics Letters*, 69(2):219–223. available at <http://ideas.repec.org/a/eee/ecolet/v69y2000i2p219-223.html>.
- Johnson, P. A. & Takeyama, L. N. (2003). Convergence among the U.S. states: Absolute, conditional, or club? Disponível em <http://irving.vassar.edu/VCEWP/VCEWP50.pdf>. Acesso em 8 mai. 2006.

- Jones, C. I. (1997). On the evolution of the world income distribution. *Journal of Economic Perspectives*, 11(3):19–36. available at <http://ideas.repec.org/a/aea/jecper/v11y1997i3p19-36.html>.
- Laurini, M., Andrade, E., & Pereira, P. L. V. (2003). Clubes de convergência de renda para os municípios brasileiros: Uma análise não-paramétrica. Encontro brasileiro de econometria. anais, SBE. Disponível em <http://www.sbe.org.br/ebe25/087.pdf>.
- Levine, R. & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review*, 82(4):942–63. available at <http://ideas.repec.org/a/aea/aecrev/v82y1992i4p942-63.html>.
- Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, 107(2):407–37. available at <http://ideas.repec.org/a/tp/qjecon/v107y1992i2p407-37.html>.
- Mora, T. (2005). Evidencing european regional convergence clubs with optimal grouping criteria. *Applied Economics Letters*, 12(15):937–940. available at <http://ideas.repec.org/a/taf/apeclt/v12y2005i15p937-940.html>.
- Mossi, M., Aroca, P., Fernandez, I., & Azzoni, C. R. (2003). Growth dynamics and space in brazil. *International Regional Science Review*, 26(3):393–418.
- Porto Junior, S. & Ribeiro, E. (2000). Dinâmica de crescimento regional: Uma análise empírica da região sul. *Revista Econômica do Nordeste*, 31(número especial):454–482.
- Porto Junior, S. & Souza, N. (2002). Crescimento regional e novos testes de convergência para os municípios da região nordeste do brasil. Disponível em http://www.ufrgs.br/ppge/pcientifica/2002_11.pdf. Acesso em 21 mar. 2006.
- Quah, D. (1993). Empirical cross-section dynamics in economic growth. *European Economic Review*, 37(2-3):426–434.
- Quah, D. (1996). Empirics for economic growth and convergence. *European Economic Review*, 40:1353–1375.
- Quah, D. (1997). Empirics for growth and distribution: stratification, polarization, and convergence clubs. *Journal of Economic Growth*, 2(1):27–59. available at <http://www.springerlink.com/content/g427052144633146/fulltext.pdf>.
- Sala-i Martin, X. X. (1996). The classical approach to convergence analysis. *Economic Journal*, 106(437):1019–36. available at <http://ideas.repec.org/a/ecj/econjl/v106y1996i437p1019-36.html>.
- Sala-i Martin, X. X. (1997). I just ran four million regressions. Disponível em <http://papers.nber.org/papers/w6252.pdf>. Acesso em 14 abr. 2006.

A. TABELAS E FIGURAS



Tabela 1 – Descrição das Variáveis

Descrição das Variáveis	Ano	Unidade	Mnemônico	Média	Coefficiente de variação
% da população de 10 a 14 anos	1970	%	POP10A14	13,5	0,08
% da população de 15 ou + anos de idade	1970	%	POP15	56,2	0,08
% da população de 25 ou + anos de idade	1970	%	POP25	36	0,1
% da população de 7 a 14 anos	1970	%	POP7A14	22,4	0,07
% da população que vive em domicílio com instalação adequada de esgoto	1970	%	POPESG	18,3	1,18
% da população que vive em domicílio com abastecimento adequado de água	1970	%	POPAGUA	27,9	0,9
% da população urbana	1970	%	POPURB	33,2	0,64
% de crianças de 10 a 14 anos > 1 ano atraso	1970	%	CRESTATR	81	0,19
% de crianças de 10 a 14 anos de idade que trabalham	1970	%	CKTRAB	15,8	0,6
% de crianças de 7 a 14 anos de idade fora da escola	1970	%	CRFESC	39,8	0,45
% de pessoas de 25 ou + anos de idade com mais de onze anos de estudo	1970	%	POPEST11	0,46	1,56
% de pessoas de 25 ou + anos de idade com menos de oito anos de estudo	1970	%	POPEST8	97,4	0,03
% de pessoas de 25 ou + anos de idade com menos de quatro anos de estudo	1970	%	POPEST4	85,5	0,14
% dos domicílios com água canalizada rede geral	1970	%	DOMAGCA	0,15	1,29
% dos domicílios com iluminação elétrica	1970	%	DOMILUEL	0,24	0,94
% dos domicílios com instalações sanitárias (rede geral)	1970	%	DOMINSA	0,05	2,31
Capital humano per capita	1980	RS de 2000/habit.	H	15	0,53
Defasagem escolar média criança 10 a 14	1970	Anos	DEFESC	3,29	0,27
Densidade demográfica	1970	Habit/Km2	DENDEM	61,3	5,02
Estoque per capita de capital governo – comércio	1970	RS de 1999/habit.	KGOVCOM	0,24	22,7
Estoque per capita de capital governo – indústria	1970	RS de 1999/habit.	KGOVIND	1,57	22,1
Estoque per capita de capital governo – serviços	1970	RS de 1999/habit.	KGOVSER	0,4	23,7
Estoque per capita de capital privado – agropecuária	1970	RS de 1999/habit.	KPRIAGP	1669,8	1,01
Estoque per capita de capital privado – comércio	1970	RS de 1999/habit.	KPRICOM	69,7	1,46
Estoque per capita de capital privado – indústria	1970	RS de 1999/habit.	KPRIIND	477	15,6
Estoque per capita de capital privado – residências	1970	RS de 1999/habit.	KPRIRES	1281,2	0,93
Estoque per capita de capital privado – serviços	1970	RS de 1999/habit.	KPRISER	35,1	2,84
Estoque per capita de capital privado total (residências + empresas)	1970	RS de 1999/habit.	KPRITOT	2289	1,4
Número de homicídios por 1000 habitantes	1980	Unidade/1000 habit./Ano	HOMICID	5,84	1,69
Número de suicídios por 1000 habitantes	1980	Unidade/1000 habit./Ano	SUICIDIO	2,88	2,37
Número de vítimas de acidentes de trânsito por 1000 habitantes	1980	Unidade/1000 habit./Ano	ACITRAN	9,67	1,71
Número médio de anos de estudo das pessoas de 25 ou + anos de idade	1970	Anos	ANOEST	1,37	0,59
Taxa analfabetismo (população 15 anos ou +)	1970	%	TAXANFB	44	0,41

Tabela 2 – Resultado das Regressões

	Modelo Base ($\alpha_i = 0, i = 2, \dots, 8$) ($\beta_i = 0, i = 2, \dots, 8$)	Intercepto Comum ($\alpha_i = 0, i = 2, \dots, 8$)	Inclinação Comum ($\beta_i = 0, i = 2, \dots, 8$)	Intercepto e Inclinação distintos	Modelo Restrito ($\alpha_2, \beta_2, \beta_7 = 0$)	Modelo Restrito ($\alpha_2, \alpha_5, \alpha_7 = 0$) ($\beta_2, \beta_5, \beta_7 = 0$)
α	0,142595* (0,044268)	0,035579 (0,04546)	-0,706111* (0,114098)	-2,112031* (0,272365)	-2,811668* (0,230227)	-3,208974* (0,161730)
β	-1,020730* (0,033154)	-1,122705* (0,03774)	-1,430867* (0,047637)	-2,292243* (0,131818)	-2,446639* (0,1127)	-2,640117* (0,084785)
α_2	-	-	-0,445341* (0,138791)	-0,918013* (0,568869)	-	-
β_2	-	0,220628* (0,073933)	-	-0,321069* (0,290988)	-	-
α_3	-	-	-0,839080* (0,095961)	1,036292* (0,338376)	1,405929* (0,305605)	1,803235* (0,258093)
β_3	-	0,695929* (0,056737)	-	0,959669* (0,187708)	1,114065* (0,174907)	1,307543* (0,158457)
α_4	-	-	1,020700* (0,087326)	3,155876* (0,285526)	3,525513* (0,245684)	3,922819* (0,183097)
β_4	-	-0,446791* (0,049672)	-	1,253306* (0,151704)	1,407702* (0,135462)	1,601180* (0,113341)
α_5	-	-	-0,586408* (0,116018)	-1,256335* (0,409770)	-0,886698* (0,383283)	-
β_5	-	0,322629* (0,068672)	-	-0,686019* (0,229222)	-0,531622* (0,218936)	-
α_6	-	-	1,034902* (0,113806)	2,993211* (0,283206)	3,362848* (0,242979)	3,760154* (0,179445)
β_6	-	-0,325539* (0,101847)	-	1,367977* (0,176973)	1,522373* (0,163311)	1,715851* (0,145535)
α_7	-	-	-1,421416* (0,199524)	-1,689944* (0,589559)	-0,461012* (0,211760)	-
β_7	-	1,170937* (0,163813)	-	-0,916607* (0,454867)	-	-
α_8	-	-	0,539839* (0,137192)	2,268932* (0,281749)	2,638569* (0,241275)	3,035875* (0,177127)
β_8	-	-0,163545* (0,152580)	-	1,035849* (0,190758)	1,190245* (0,178183)	1,383723* (0,162071)
R^2	0,21	0,3	0,35	0,39	0,38	0,38

* indica significância ao nível de 5%.

**Tabela 3** – Resultados dos Testes para Convergência Condicional

Variável	Nível			Log		
	β	Controle	R^2	β	Controle	R^2
ACITRAN	-1,092775* (0,034903)	0,010486* (0,001664)	0,22	-1,120701* (0,036620)	0,121674* (0,019376)	0,22
ANOEST	-1,641816* (0,049327)	0,790488* (0,047810)	0,26	-1,748397* (0,048665)	2,255571* (0,114932)	0,28
CRESTATR	-1,847765* (0,048261)	-0,056657* (0,00253)	0,30	-1,721479* (0,047758)	-3,616666* (0,185116)	0,28
CRFRESC	-1,593873* (0,041995)	-0,038168* (0,001859)	0,29	-1,578391* (0,043262)	-1,418579* (0,075076)	0,28
CRTRAB	-1,044375* (0,034642)	-0,006719* (0,002873)	0,21	-1,034286* (0,034743)	-0,060781* (0,046620)	0,21
DEFESC	-2,029961* (0,047944)	-1,147310* (0,042311)	0,34	-1,974906* (0,048345)	-4,410777* (0,13278)	0,33
DENDEM	-1,005094* (0,033634)	-0,000230* (0,000086)	0,21	-0,996930* (0,034567)	-0,060118* (0,024895)	0,21
DOMAGCA	-1,139446* (0,042548)	0,762737* (0,172010)	0,21	-1,164395* (0,042663)	1,162786* (0,218558)	0,21
DOMILUEL	-1,365837* (0,049543)	1,580349* (0,170156)	0,23	-1,431046* (0,050129)	2,493206* (0,231199)	0,23
DOMINSA	-1,034232* (0,038189)	0,174776* (0,245282)	0,21	-1,040053* (0,038319)	0,298002* (0,296317)	0,21
H	-1,074197* (0,040793)	0,009021* (0,004014)	0,21	-1,127485* (0,043754)	0,337473* (0,090477)	0,21
HOMICID	-1,044665* (0,033618)	0,010761* (0,002680)	0,21	-1,045788* (0,034383)	0,056655* (0,020837)	0,21
KGOVCOM	-1,021076* (0,033194)	0,001072* (0,004737)	0,21	-1,012832* (0,033524)	-0,157875* (0,100202)	0,21
KGOVIND	-1,021985* (0,033207)	0,0011209* (0,001767)	0,21	-1,013722* (0,034175)	-0,041820* (0,049447)	0,21
KGOVSER	-1,019047* (0,033203)	-0,002594* (0,002757)	0,21	-1,017679* (0,033240)	-0,112502* (0,089471)	0,21
KPRIAGP	-1,197785* (0,034136)	0,000244* (0,000016)	0,26	-1,187335* (0,0322615)	0,479242* (0,024289)	0,29
KPRICOM	-1,047074* (0,039973)	0,000362* (0,000307)	0,21	-1,118778* (0,045985)	0,106454* (0,034541)	0,21

* indica significância ao nível de 5%.

Tabela 4 – Testes de Robustez

Variável	Nível			Log		
	Valor Médio	Desvio Padrão	Limite Inferior	Limite Superior	Valor Médio	Desvio Padrão
ACITRAN	0,0083 [*]	0,0017	-0,0053	0,0153	0,0994 [*]	0,0198
ANOEST	0,6820 [*]	0,0643	-0,7384	1,9192	0,1798 [*]	0,0293
CRESTAIR	-0,0525 [*]	0,0046	-0,2720	0,0628	-2,1285 [*]	0,3100
CRFESC	-0,0320 [*]	0,0024	-0,0772	0,0138	-1,0368 [*]	0,0951
CRTRAB	-0,0008	0,0029	-0,0202	0,0297	0,0167	0,0165
DEFESC	-1,1969 ^{**}	0,0686	-3,1113	-0,3970	-2,9008 [*]	0,2255
DENDEM	-0,0002 [*]	0,0001	-0,0008	0,0009	-0,0605 [*]	0,0221
DOMAGCA	-0,4911	0,3951	-21,3820	2,7003	0,8508	0,5010
DOMILJEL	0,5610	0,3353	-20,6090	4,5985	2,6014 [*]	0,4560
DOMINSA	-1,0367	0,6722	-21,5420	14,6730	-0,4436	0,8144
H	0,0014	0,0044	-0,0499	0,0473	0,0835	0,0962
HOMICID	0,0116 [*]	0,0028	-0,0104	0,0422	0,0518 [*]	0,0218
KGOVCOM	0,0005	0,0046	-0,0090	0,0263	-0,1983 [*]	0,0981
KGOVIND	0,0012	0,0017	-0,0033	0,0095	-0,0846	0,0485
KGOVSR	-0,0035	0,0028	-0,0204	0,0212	-0,1320	0,0885
KPRIAGP	0,0002 [*]	0,0002	-0,0005	0,0006	0,4674 [*]	0,0265
KPRICOM	-0,0001	0,0003	-0,0028	0,0027	0,0144	0,0335
KPRIND	0,0001 [*]	0,0002	-0,0004	0,0004	0,1193 [*]	0,0192
KPRIRES	-8,64E+06	1,38E+08	-1,67E+10	0,0004	0,0616	0,0645
KPRISER	-0,0003	0,0003	-0,0029	0,0020	0,2010 [*]	0,0283
KPRITOT	5,51E+06	1,26E+08	-0,0005	0,0007	1,0612 [*]	0,0712
POP10A14	0,1066	0,0690	-0,1601	1,9984	0,5529	0,9186
POP15	0,0064	0,0126	-0,5221	0,2462	0,9878	0,7688
POP25	-0,0094	0,0171	-0,4399	0,5889	-0,3590	0,6321
POP7A14	0,0416	0,0534	-0,1948	1,4156	0,3227	1,1770
POPAGUA	0,0108 [*]	0,0015	-0,0152	0,0271	0,2261 [*]	0,0235
POPESG	0,0077 [*]	0,0018	-0,0252	0,0285	0,1784 [*]	0,0241
POPEST11	-0,0752	0,0531	-0,7378	0,4707	0,0010	0,0014
POPEST4	-0,0331 [*]	0,0052	-0,3044	0,0608	-0,7528 [*]	0,3505
POPEST8	-0,0339	0,0429	-2,3144	1,0246	5,0635	3,9121
POPURB	-0,0109 [*]	0,0017	-0,0594	0,0195	-0,1230 [*]	0,0505
SUICIDIO	0,0198 [*]	0,0040	-0,0146	0,0486	0,1189 [*]	0,0278
TAXANFB	-0,0476 [*]	0,0025	-0,0746	0,0120	-1,5770 [*]	0,0976

* indica robustez do coeficiente com base apenas no teste proposto por Sala-i Martin (1997).

** indica robustez do coeficiente com base em ambos os testes.

**Tabela 5** – Modelos Lineares dos Clubes

	Clubes 1, 2, 5 e 7	Clube 3	Clubes 4 e 6	Clube 8
α	-3,208974	-1,139334	0,559153	-0,177465
β	-2,640118	-1,144800	-1,144800	-1,144800

Tabela 6 – Quantidade de Municípios por Região/Clube

	Clubes 1, 2, 5 e 7	Clube 3	Clubes 4 e 6	Clube 8
Sul	13	28	503	50
Sudeste	116	39	969	277
Centro-Oeste	13	2	190	18
Nordeste	783	322	165	28
Norte	35	51	56	1

Tabela 7 – Resultados dos Testes para Convergência Clube

Variável	Nível			Log		
	Wald	β	Controle	Wald	β	Controle
ACITRAN	244,5*	-1,182236* (0,049114)	0,006639* (0,001486)	249,0*	-1,208720* (0,049525)	0,101443* (0,017376)
ANOEST	206,8*	-1,472090* (0,055952)	0,562143* (0,050219)	183,0*	-1,475829* (0,054681)	1,486204* (0,121250)
CRESTATR	167,5*	-1,605731* (0,057170)	-0,037129* (0,002598)	187,5*	-1,551510* (0,057184)	-2,375990* (0,186284)
CRFRESC	170,8*	-1,351677* (0,050673)	-0,023059* (0,001937)	182,9*	-1,366870* (0,051730)	-0,851238* (0,076583)
CRTRAB	249,8*	-1,148377* (0,048558)	-0,004265* (0,002602)	250,5*	-1,147784* (0,048601)	-0,045007* (0,042345)
DEFESC	130,5*	-1,671474* (0,056292)	-0,767922* (0,045751)	144,4*	-1,664255* (0,057083)	-2,915242* (0,182634)
DENDEM	248,7*	-1,142013* (0,048624)	-0,000069* (0,000078)	250,3*	-1,163864* (0,049321)	0,049836* (0,023473)
DOMAGCA	246,4*	-1,192031* (0,052127)	0,434332* (0,175992)	244,2*	-1,196447* (0,051965)	0,607972* (0,220354)
DOMILUEL	246,0*	-1,422516* (0,058276)	1,504284* (0,178556)	240,3*	-1,441632* (0,058031)	2,159247* (0,237444)
H	252,9*	-1,210226* (0,051996)	0,013748* (0,003967)	250,3*	-1,226054* (0,053559)	0,319913* (0,089948)
HOMICID	250,8*	-1,158537* (0,048534)	0,009669* (0,002379)	252,0*	-1,161320* (0,048700)	0,062627* (0,018626)
KPRIAGP	194,0*	-1,215355* (0,049321)	0,000109* (0,000016)	161,6*	-1,191693* (0,048285)	0,250637* (0,028056)
KPRICOM	251,1*	-1,178915* (0,053017)	0,000458* (0,000287)	250,3*	-1,214696* (0,054272)	0,092190* (0,032197)
KPRIIND	254,9*	-1,227631* (0,051940)	0,000069* (0,000016)	258,8*	-1,455384* (0,058467)	0,164149* (0,017676)
KPRIRES	251,4*	-1,099678* (0,051894)	-0,000063* (0,000026)	246,8*	-1,133921* (0,054998)	-0,024939* (0,059343)

* indica significância ao nível de 5%.



Figura 1 – Árvore de regressão

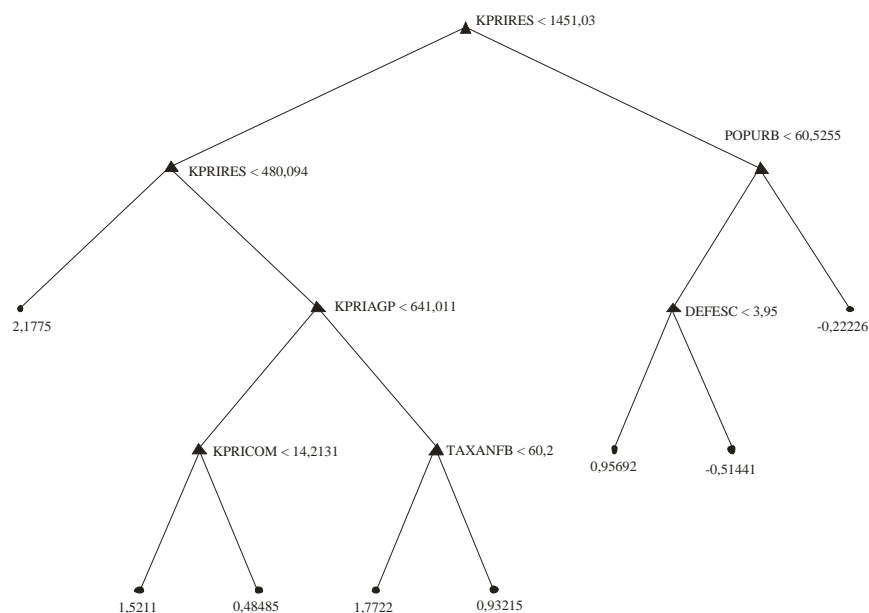


Figura 2 – Distribuição geográfica dos clubes de convergência

