



**Clubes de Convergência de
Renda para os Municípios
Brasileiros: Uma Análise
Não-Paramétrica**

Márcio Laurini

Eduardo Andrade

Pedro L. Valls Pereira



Copyright Insper. Todos os direitos reservados.

É proibida a reprodução parcial ou integral do conteúdo deste documento por qualquer meio de distribuição, digital ou impresso, sem a expressa autorização do Insper ou de seu autor.

A reprodução para fins didáticos é permitida observando-se a citação completa do documento

Clubes de Convergência de Renda para os Municípios Brasileiros: Uma Análise Não-Paramétrica¹

Márcio Laurini
Ibmec/SP

Eduardo Andrade
Ibmec/SP

Pedro L. Valls Pereira
Ibmec/SP

Resumo

Neste artigo analisamos a evolução da distribuição da renda relativa per capita para os municípios brasileiros no período de 1970-1996. As análises realizadas são baseadas em metodologias não paramétricas, sem assumir distribuições de probabilidade e formas funcionais para os dados. Realizamos dois testes de convergência - um teste para Sigma convergência baseado no princípio do *Bootstrap* e outro de Beta convergência usando *Smoothing Splines* para as regressões de crescimento. As evidências obtidas mostram a necessidade de se modelar a dinâmica da renda dos municípios brasileiros como um processo de clubes de convergência, utilizando a metodologia de matrizes de transição e núcleo estocástico. Os resultados mostram a formação de dois clubes de convergência, um clube de baixa renda formado pelos municípios das regiões Norte e Nordeste e outro clube de alta renda formado pelos municípios das regiões Centro-oeste, Sudeste e Sul. A formação dos clubes de convergência é confirmada por um teste de *bootstrap* para multimodalidade.

Palavras Chave: Clubes de Convergência, Métodos Não Paramétricos, Dinâmicas de Distribuição.

1 Introdução

A hipótese de convergência de renda per capita pode ser sintetizada como uma tendência de diminuição progressiva no tempo das diferenças entre as rendas relativas entre as economias ricas e pobres. A convergência é uma das principais previsões do modelo neoclássico de crescimento proposto por Solow [22] e Swan [23], sendo uma consequência do pressuposto da existência de retornos decrescentes para os fatores de produção. Isto implica que a produtividade do capital é maior em economias relativamente mais pobres, o que levaria a uma maior taxa de crescimento nas economias com um estoque de capital mais baixo, e com o passar do tempo a progressiva redução na diferença entre taxas de crescimento levaria a uma convergência para mesmos níveis de taxas de crescimento (Beta Convergência) e renda (Sigma Convergência) no estado estacionário. Devido a uma maior homogeneidade nos parâmetros tecnológicos e comportamentais, causada pela ausência de barreiras a mobilidade de capital dentro de um mesmo país, a convergência entre rendas para municípios dentro de um mesmo país seria ainda mais provável.

Tradicionalmente, os testes para convergência e a modelagem da distribuição de renda são baseados nos pressupostos de que a distribuição dos dados é conhecida, por exemplo assumir que a distribuição dos dados é Normal, e nos testes de Beta Convergência assumir que a relação entre o a taxa de crescimento e o logaritmo da renda inicial é linear. Nossa análise mostra que o pressuposto de linearidade na regressão de crescimento pode esconder relações de divergência para algumas faixas de renda relativa. Os testes de convergência baseados em regressões corte transversal, como o uso de regressões de crescimento relacionando a taxa de crescimento como função da renda inicial, são criticados por Quah pelo fato de que modelar uma média condicional pode ser inadequado para se analisar a hipótese de convergência. O primeiro problema existente nessa regressão é assumir que o coeficiente estimado é o mesmo para todas as economias. O segundo problema é conhecido como “Falácia de Galton”, apontado por Friedman[11] e Quah [15], que mostra que o coeficiente negativo encontrado em regressões de crescimento pode não significar convergência, e sim um sintoma de uma regressão a média.

¹Ibmec Business School, Rua Maestro Cardim 1170, 11o. andar, 01323-001, São Paulo, Brasil; Fone: (55)(11)3175-2311; Fax: (55)(11)3175-2315; E-mail: (marciopl@ibmec.br). Os autores gostariam de agradecer a Patrícia Palomo e Samuel Kinoshita pela excelente assistência acadêmica. O último autor reconhece o apoio financeiro parcial do CNPq e PRONEX.

Relaxando estes pressupostos de linearidade e distribuição, testamos convergência e modelamos a dinâmica da renda relativa dos municípios brasileiros usando métodos não-paramétricos. Realizamos testes de Sigma-Convergência usando as estatísticas tradicionais do Coeficiente de Variação e o Índice de Theil, que medem a dispersão entre rendas, onde obtemos as distribuições desses estimadores através do uso de *Bootstrap*. O teste de Beta-Convergência usando o estimador não paramétrico de *Smoothing Spline*, relaxa a linearidade imposta pela estimação por mínimos quadrados ordinários, e derivamos um teste de convergência baseado na primeira derivada deste estimador. Este teste mostra que a hipótese de convergência, representada por uma relação negativa entre taxa de crescimento e renda inicial não é válida para todos os níveis de renda inicial, o que mostra que existem indícios de divergência para as rendas relativas dos municípios brasileiros. Este resultado é consistente com a bimodalidade obtida na estimação não paramétrica via Função de Núcleo para as densidades das rendas nos anos de 1970 e 1996. Esta bimodalidade, que pode ser interpretada como a formação de clubes de convergência de renda conforme propostos por Quah [16], é testada estatisticamente através de um teste de multimodalidade utilizando-se *bootstrap*.

Modelamos a evolução da distribuição da renda relativa dos municípios brasileiros usando as metodologias de Dinâmicas de Distribuição proposta por Quah [16], que modelam a evolução da renda como um processo de Markov. A vantagem dessa metodologia é que ela formula uma lei de movimento para toda a distribuição das rendas entre os períodos em análise, permitindo capturar a existência dos clubes de convergência verificados nos dados. Este processo de Markov para as renda relativas é modelado utilizando-se uma formulação discreta, através do uso de matrizes de transição, e em uma formulação contínua, conhecida como Núcleo Estocástico, que evita os problemas associados a discretização do processo de transição na estimação de matrizes de transição.

Nossa análise mostra que existem evidências apontando para a formação de dois clubes de convergência, um formado pelas regiões mais ricas dos municípios do Sudeste, Sul e Centro-Oeste e outro formado pelas regiões relativamente mais pobres dos municípios do Nordeste e Norte, e que a hipótese de convergência para um mesmo nível de taxas de crescimento e renda relativa é rejeitada pelos dados.

A base de dados consiste nas rendas per capita de 3781 municípios brasileiros nos anos de 1970 e 1996, contruída a partir de dados de renda e população obtidos do IPEA e do IBGE. A base de dados é a mesma utilizada por Andrade et al. [1], que explica os ajustes necessários na base de dados para contornar o problema da alteração do número de municípios nesse período (Tabela 6). Este artigo está organizado da seguinte forma: na Seção 2 descrevemos alguns estudos anteriores sobre convergência para o Brasil; na Seção 3 realizamos um teste de sigma convergência através do uso de *bootstrap*; na Seção 4 testamos não parametricamente a hipótese de beta convergência através do uso de *Smoothing Splines*; na Seção 5 estimamos densidades usando funções Núcleo e testamos a presença de bimodalidade. A Seção 6 contém as dinâmicas de distribuição estimadas e a Seção 7 apresenta as conclusões finais.

2 Estudos Anteriores

Os estudos anteriores sobre convergência de renda no Brasil utilizam-se quase que exclusivamente de dados de renda em nível estadual, pela dificuldade em se obter os dados de renda para os municípios. Os artigos de Ferreira e Diniz [9] e Schwartzman [21] encontram convergência absoluta para as rendas per capita do estados brasileiros para o período entre 1970 e 1985. Azzoni [2] critica este resultado, apontando que o período de 1970-1985 utilizado nestes estudos foi um período de convergência e redução de desigualdades muito forte, mas esta dinâmica de convergência não se manteve posteriormente, além de mostrar alguns problemas com a construção dos dados utilizados no estudo.

Em resposta às críticas de Azzoni [2], Ferreira [8] estima matrizes de transição de Markov para os dados de pib per capita estaduais para os anos de 1970 até 1995. Os resultados das distribuições ergódicas (distribuição de longo prazo das rendas per capita) estimadas por Ferreira [8] mostram uma tendência de concentração nas categoriais médias de renda, e o desaparecimento

das categorias de renda acima de 120% da média nacional, e pouca alteração na distribuição de renda dos extratos pobres e muito pobres.

Utilizando medidas de associação espacial, Mossi et all [18] encontra o resultado de polarização de rendas com forte componente espacial. No cluster de baixas rendas se encontrariam principalmente os estados da região nordeste (estados de PI, CE, RN, PB e BA), e os estados da região sul e sudeste (RJ, SP, PR e MG) formariam o cluster de rendas elevadas. Mossi et all [18] utilizam matrizes de transição e núcleos estocásticos na análise da evolução das rendas per capita estaduais. Os resultados de sua estimação de matrizes de transição mostram uma elevada persistência nas classes extremas (a divisão é feita em 5 extratos de renda, analisando o período 1939-1998). A estimação de núcleos estocásticos realizados por Mossi et all [18] mostra as mesmas características de elevada persistência, tanto na análise incondicional quanto condicionada espacialmente. Os principais resultados confirmam que a dinâmica da distribuição de rendas é fortemente influenciada por fatores regionais, e que existem dois clusters de convergência de rendas, um de baixas rendas formado pelos estados da região nordeste e outro de alta renda pelo estados da região sudeste e sul.

Utilizando regressões tradicionais de crescimento estimadas por mínimos quadrados ordinários e regressão quantílica, Andrade et All [1] não consegue rejeitar a hipótese de Beta Convergência para o Brasil e para as regiões em separado usando a mesma base de dados de rendas municipais usadas em nosso estudo. Pontual e Porto Júnior [19] estudam convergência para os municípios da região sul no período de 1970-1991, encontrando indícios da formação de clubes de convergência dentro desta região, e também para os estados brasileiros para o período 1985-1998, mostrando uma tendência de estratificação da renda dos estados brasileiros em três grupos, um de estados pobres com 26,9% dos estados, um de renda média com 52% dos estados e um de estados ricos com 11,4%.

Em relação aos estudos anteriores, nosso artigo utiliza dados de renda municipais e substitui os testes paramétricos de sigma e beta convergência por metodologias não paramétricas, encontrando resultados mais robustos em favor da hipótese da formação de clubes de convergência para o Brasil. A modelagem utilizada de Dinâmicas de Distribuição consegue capturar a lei de movimento das rendas relativas per capita sem os problemas associados a discretização do processo Markoviano, e os resultados do processo da formação de dois clubes de convergência são confirmados estatisticamente por um teste de multimodalidade.

3 Sigma Convergência

Uma definição simples para o processo de sigma convergência é de convergência para um único ponto de renda, o que pode ser entendido como uma dinâmica contínua de diminuição entre as diferenças de rendas entre economias, implicando em uma menor dispersão e desigualdade entre rendas.

Para analisar a dispersão entre rendas relativas, duas medidas freqüentemente usadas na literatura para se testar sigma convergência são o Índice de Theil e o Coeficiente de Variação, que medem o grau de desigualdade existente na distribuição dos dados. Os métodos tradicionais de se verificar sigma convergência usando estes indicadores de desigualdade consiste em montar uma série temporal com os valores dos índices medidos a cada ano e através de uma regressão linear contra o tempo verificar se existe uma tendência significativa de redução da desigualdade, o que corresponderia a um parâmetro negativo nessa regressão.

Como estamos utilizando apenas os anos de 1970 e 1996 em nossa análise, a forma de se testar estatisticamente se houve uma redução na desigualdade de renda, através dos coeficientes de variação e índices de Theil estimados para os anos de 1970 e 1996 foi obter a distribuição destes estimadores usando o método de bootstrap e construir intervalos de confiança para os valores estimados.

O método do bootstrap consiste em tratar a amostra disponível como sendo a população, e a partir de seguidas reamostragens desta amostra, obter a distribuição dos estimadores ou estatísticas de teste. Dada a necessidade de condições de regularidade muito pouco restritivas, o

Tabela 1: Intervalos de Confiança por Bootstrap - Índice Theil 1970 e 1996

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

método de bootstrap permite aproximações acuradas das distribuições em amostras finitas. O bootstrap também é vantajoso por substituir a necessidade de derivações matemáticas por tempo de computação quando estas derivações são excessivamente complicadas. Aplicando o método de bootstrap aos Índice de Theil e ao coeficiente de variação podemos testar se a redução destes estimadores é estatisticamente significativa, sem a necessidade de se assumir a priori que os dados provêm de alguma distribuição.

A tabela 1 mostra os intervalos de confiança obtidos via bootstrap para os Índices de Theil e o Coeficiente de Variação para os dados de renda per capita municipal para todos os municípios do Brasil nos anos de 1970 e 1996. Os intervalos de confiança foram obtidos pelo método de bootstrap não paramétrico de BC_a percentile (*Bias Corrected and Accelerated*). Este método necessita de um menor número de replicações de bootstrap para aproximar corretamente e com mais acurácia às distribuições dos estimadores, e também é invariante a transformações nos estimadores, conforme Efron e Tibshirani [7]. A Tabela 1 contém os valores correspondentes aos percentis 0.01, 0.025, 0.05, 0.95, 0.975 e 0.99 das distribuições obtidas usando o *bootstrap*, permitindo a construção de intervalos de confiança.

Os testes de sigma convergência mostram que embora exista uma redução na desigualdade entre rendas per capita municipais para todas as regiões exceto a região Norte, onde houve um aumento nos Índices de Theil e no coeficiente de variação, a redução na desigualdade correspondente a hipótese de sigma convergência só é estatisticamente válida para a região Sul, onde ao nível de significância de 1% rejeitamos a hipótese nula de que tanto o índice de Theil quando o coeficiente de variação sejam os mesmos para os períodos de 1970 e 1996, já que os intervalos de confiança em 1996 para a região Sul não se sobrepõem aos intervalos de confiança dos dois indicadores para esta região em 1970. Para as demais regiões onde houve redução na desigualdade, estatisticamente não podemos rejeitar que os a hipótese nula de que os indicadores são iguais.

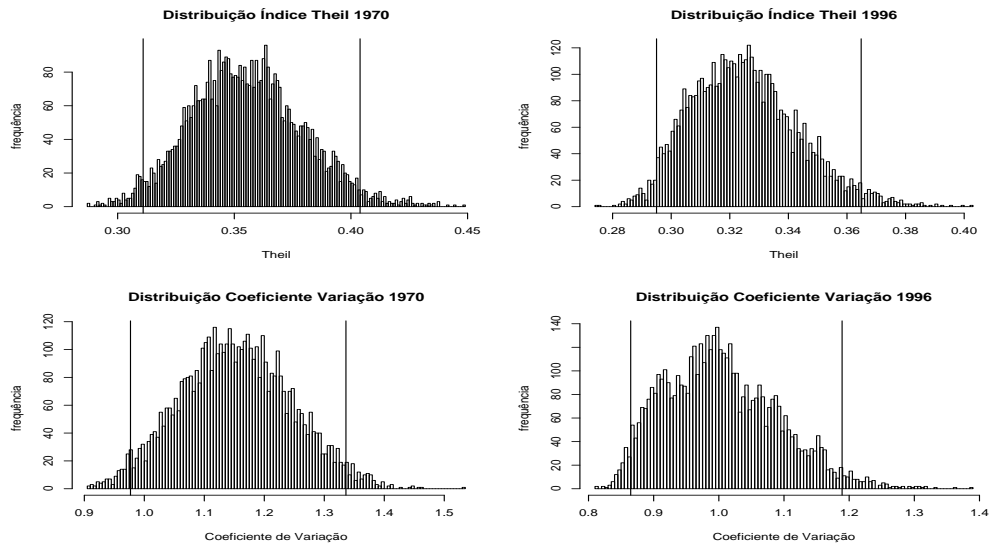
Uma das condições necessárias para a validade dos resultados do procedimento de bootstrap é que as amostras sejam provenientes de um processo independente, mas as análises nas seções posteriores mostram que pode existir um fator regional para a distribuição de rendas. Para controlar este efeito, que representaria uma violação do requisito de independência necessário ao bootstrap, realizamos um procedimento conhecido como *Bootstrap* Estratificado. Nesse procedimento, realizamos a reamostragem para todos os municípios do Brasil forçando a que o número de municípios dentro de cada região incluídos em cada reamostragem seja o mesmo, o que equivale a que fazer uma reamostragem dentro de cada região e calcular o resultado para o país todo.

As distribuições obtidas para o coeficiente de variação e para o índice de Theil via bootstrap

Tabela 2: *Bootstrap* Estratificado - Theil e Coeficiente de Variação 1970-1996

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

Figura 1: *Bootstrap* Estratificado - Índices de Theil e Coeficiente de Variação



estratificado para o Brasil se encontram na figura 1. As linhas verticais marcam os valores para um intervalo de confiança de 5% de significância. A Tabela 2 contém os valores superiores e inferiores dos intervalos de confiança obtidos por este método, e mostram que o resultado obtido com o bootstrap sem estratificação se mantém: apesar de uma redução nos valores calculados para o índice de Theil e o coeficiente de Variação, não podemos rejeitar que eles sejam estatisticamente iguais entre 1970 e 1996.

4 Beta Convergência

A hipótese de Beta Convergência pode ser vista como a existência de uma relação negativa entre a taxa de crescimento e o valor da renda inicial, sendo causada pela presença de retornos decrescentes na função de produção² utilizada nos modelos de crescimento de Solow [22] e Swan [23]. Beta Convergência é um fator necessário mas não suficiente para a existência de sigma convergência, já que choques exógenos nas taxas de crescimento poderiam aumentar a dispersão entre rendas.

A regressão de crescimento consiste na estimação da seguinte equação:

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

onde Y_{it} e Y_{i0} são as rendas nos períodos T e no período inicial, T é o número de períodos, α e β são constantes e μ_{it} é o erro médio entre as datas 0 e T para a taxa de crescimento. A hipótese de Beta Convergência é dada pela existência de um β negativo nesta equação.

O problema com esta abordagem é que a formação de clubes de convergência não pode ser capturada por uma estimação paramétrica usando mínimos quadrados, já que esta impõe uma mesma taxa de convergência para todos os níveis de renda. O conceito de clubes de convergência, segundo Quah [16], equivale ao desaparecimento das faixas intermediárias de renda, e o surgimento de dois atratores para rendas, um de alta e outro de baixa renda. Este comportamento seria visualizado pela existência de dois picos distintos na densidade empírica dos dados.

Para capturar a formação de clubes de convergência, é necessário o uso de métodos não paramétricos que possibilitem estimar o parâmetro β equivalente a cada nível de renda inicial. Para modelar a relação entre taxa de crescimento e renda inicial sem adotar uma forma funcional definida utilizamos a técnica de regressão não paramétrica conhecida como *Smoothing Spline*. Um *Smoothing Spline* pode ser definido com a solução do problema de se minimizar o seguinte funcional:

$$S_\lambda(g) = \sum_{i=1}^n (Y_i - g(X_i))^2 + \lambda \int (g''(x))^2 dx \quad (2)$$

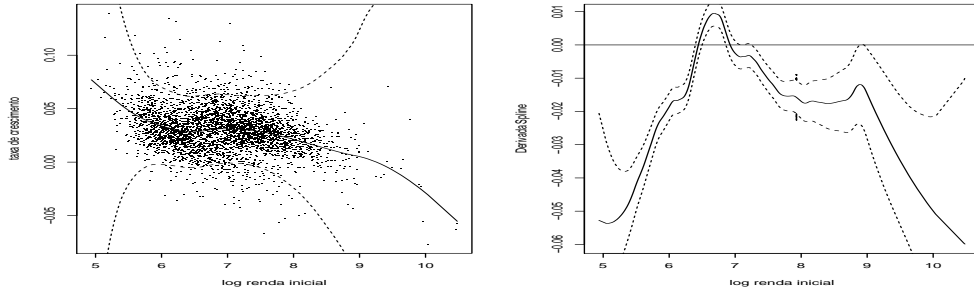
onde g é uma curva qualquer, x é o conjunto de dados e λ é um parâmetro de suavidade do ajuste, controlando a troca entre minimização do resíduo e a aspereza (*roughness*) do ajuste. Conforme Hardle [14], este problema de minimização tem uma única solução $\hat{m}_\lambda(x)$, dada por um polinômio cúbico chamado de *cubic spline*. Uma das vantagens deste método de interpolação é que ele fornece diretamente as primeiras derivadas da função $\hat{m}_\lambda(x)$. A primeira derivada pode ser interpretada como a medida da resposta da variável dependente Y em relação à uma variação da variável explicativa X, de forma análoga aos parâmetros de uma regressão linear. O sinal da derivada estimada será o nosso indicador de convergência, onde valores negativos da derivada indicam convergência e valores positivos divergência de renda.

Para a determinação do parâmetro de suavização λ utilizamos o critério de validação cruzada generalizada (*Generalized Cross Validation*), e na construção dos intervalos de confiança para o *Smoothing Spline* utilizamos a formulação de Wahba (1983), que visualizando este modelo como um modelo bayesiano, determinou que os intervalos de confiança para o *spline* são dados por:

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

²Barro e Sala-i-Martin [3] derivam a regressão de crescimento utilizada nos testes de Beta Convergência.

Figura 2: Regressão de Crescimento - *Smoothing Spline* -Brasil



onde $\hat{S}_\lambda(x_i)$ são os valores previstos pelo spline, $z_{\alpha/2}$ é o valor correspondente em uma distribuição normal ao nível de confiança desejado para o intervalo de confiança, e $a_{ii}(\lambda)$ são os elementos da diagonal da matriz de *leverages* $A(\lambda)$ definida pelo smoothing spline³. A variância σ^2 do smoothing spline é definida como:

$$\hat{\sigma}^2 = \frac{e'e}{tr(A(\lambda))} \quad (4)$$

onde e é o erro associado a cada valor na amostra e tr é o traço da matriz $A(\lambda)$ de *leverages*. O intervalo de confiança para a derivada do *smoothing spline* é obtido a partir dos valores previstos para o spline usando *delta method*. Na construção dos intervalos de confiança assumimos o nível de 5% de significância.

A existência de divergência para algumas faixas de renda seria dada por um relacionamento positivo entre a taxa de crescimento e a renda inicial para estas rendas, o que pode ser medido pela primeira derivada do smoothing spline. A formação de clubes de convergência seria dada pela existência de uma faixa de divergência, correspondente a valores positivos para a primeira derivada do spline, para os valores intermediários de renda. Isto pode ser interpretado como o desaparecimento de municípios com rendas intermediárias. Estes passariam a pertencer ao grupo de municípios de rendas altas.

As Figuras 2, 3, 4, 5, 6 e 7 mostram os *Smoothing Splines* e as primeiras derivadas estimados tendo como variável dependente a taxa de crescimento da renda entre 1970 e 1996 e como variável explicativa o log da renda no ano de 1970, para todos os municípios do Brasil e para os municípios das regiões em separado, bem como os intervalos de confiança associados. Estes gráficos mostram que dentro de cada região não encontramos indícios de divergência, mostrando que a formação dos clubes de convergência é devido ao deslocamento das rendas relativas entre as regiões e não de municípios dentro de cada região.

A regressão não paramétrica entre a taxa de crescimento e a renda inicial na forma de um *Smoothing Spline* para todos os municípios do Brasil (Figura 2) mostra que para as rendas entre aproximadamente 5 e 6 do log do Pib per capita, e para rendas acima de 7 do log do pib per capita, a relação entre taxa de crescimento e renda inicial é uma curva com inclinação negativa, como esperado pelo hipótese de Beta convergência. Entretanto se olharmos para o logaritmo das rendas em 1970 entre 6.3552 e 6.7640⁴, a curva ajustada e especialmente a derivada do *Spline* mostram a presença de divergência, indicada pela derivada com valores positivos e estatisticamente diferentes de zero. A faixa de divergência é consistente com os valores de rendas que tendem a desaparecer com a formação dos clubes de convergência, o que será confirmado nas seções 5 e 6 pelas estimações não paramétricas de densidade e a modelagem de dinâmicas de distribuição. Os resultados dos

³Para maiores referências sobre os componentes do estimador de Smoothing Spline ver a documentação do pacote ModReg para software R: <http://www.r-project.org>

⁴Correspondentes às rendas de US\$ 575.72 e US\$ 866.15 per capita em 1970 .

Figura 3: Regressão de Crescimento - *Smoothing Spline* - Norte

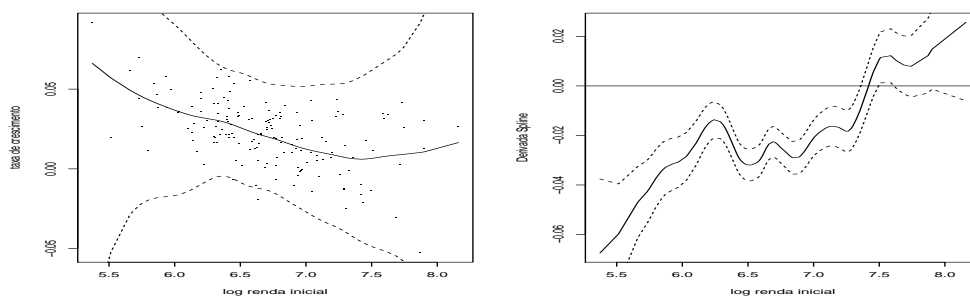


Figura 4: Regressão de Crescimento - *Smoothing Spline* - Nordeste

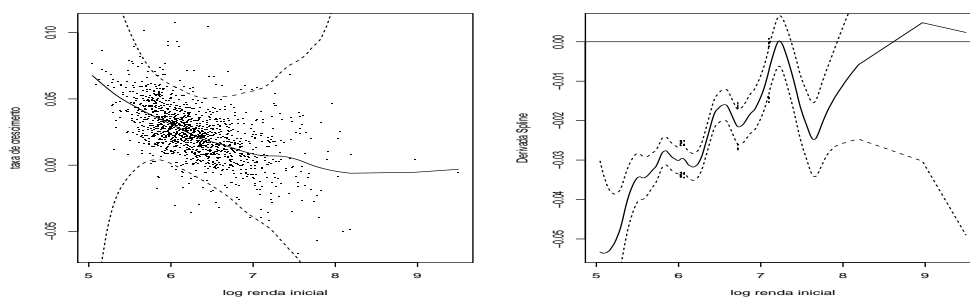
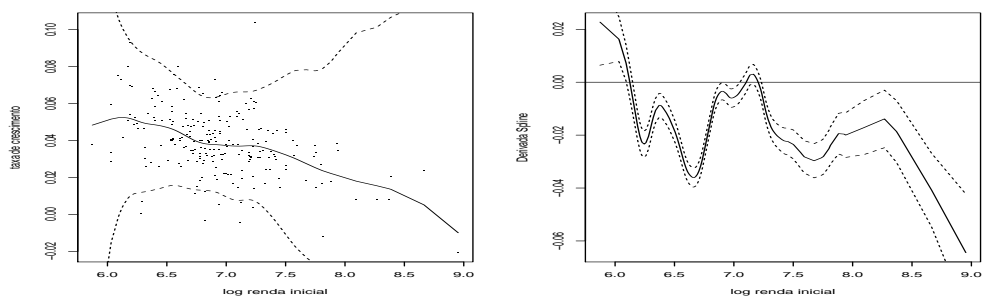


Figura 5: Regressão de Crescimento - *Smoothing Spline* - Centro-Oeste



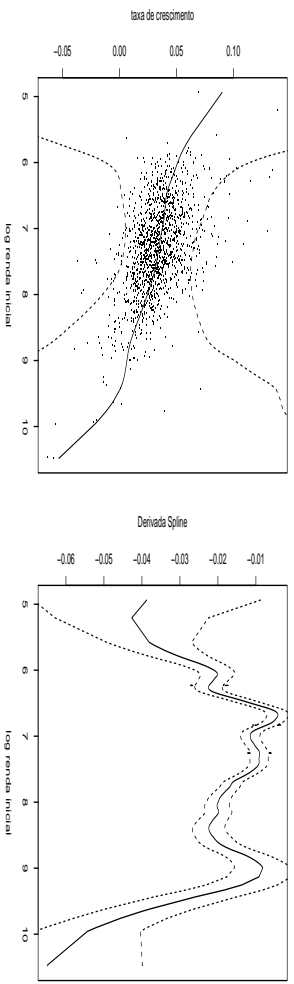


Figura 6: Regressão de Crescimento - *Smoothing Spline* - Sudeste

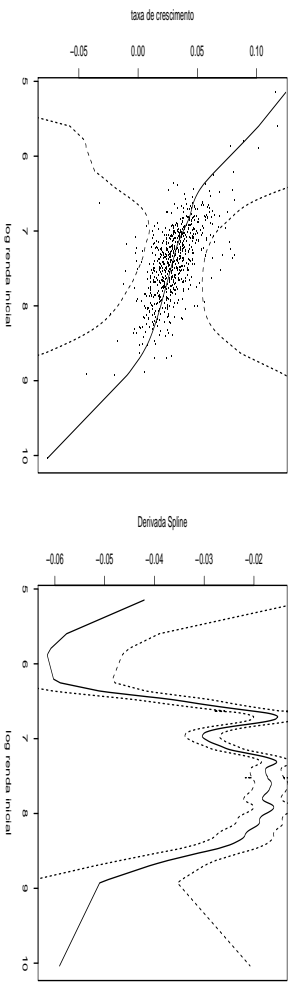


Figura 7: Regressão de Crescimento - *Smoothing Spline* - Sul

Smoothing Splines aplicados para cada região em separado não indicam a presença de divergência, sugerindo que a formação dos clubes de convergência no Brasil será causada por um deslocamento uniforme das rendas relativas per capita de cada região, conforme será visto na Seções 5 e 6.

O pressuposto de uma relação linear entre renda inicial e taxa de crescimento, utilizada nos testes tradicionais de convergência, mostra-se inadequado já que assumindo uma mesma taxa de convergência para todas os níveis de renda inicial, sofre do problema de regressão a média e não permite revelar a existência de faixas de divergência relativas a níveis específicos de renda. A estimação não paramétrica utilizando *Smoothing Splines* mostra que o nível de rendas intermediárias está divergindo, e podemos interpretar esta divergência como o desaparecimento dessas faixas intermediárias de renda como a formação de dois clubes de convergência.

5 Densidades Não Paramétricas

Uma forma de se analisar a distribuição das rendas relativas per capita é através da visualização das funções de densidade de probabilidade. Iremos estimar a função de densidade de probabilidade de forma não paramétrica através do uso de densidades estimadas usando funções de *núcleo*.

A definição de um *núcleo* é a de uma função contínua, limitada e simétrica, com a propriedade de que sua integral indefinida é igual a 1:

$$\int K(u)du = 1. \quad (5)$$

Esta propriedade permite construir um estimador para densidade, ao lembrarmos que a função de densidade de um escalar Z no ponto z_0 pode ser aproximada por:

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

e um estimador para $f(z_0)$ ser dado por:

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

Usando estas propriedades, a forma típica de um estimador de densidade por *núcleo* é dada por:

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

onde utilizamos com uma função de *núcleo* $K(u)$:

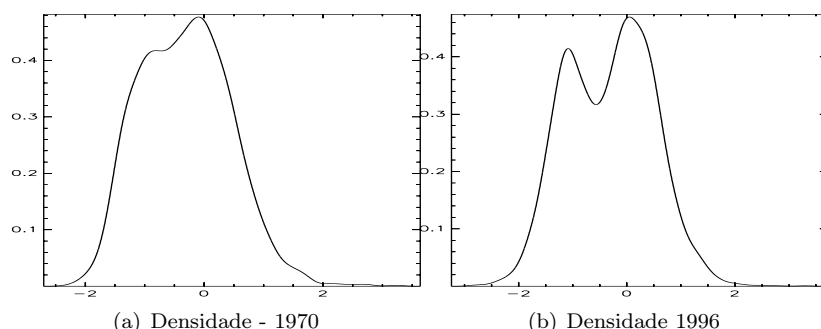
$$K(u) = \frac{1}{2} I_{(-1,1)}(u) \quad (9)$$

onde I é a função Indicador. O núcleo Gaussiano utilizado no artigo é definido como:

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

Um fator fundamental no uso de estimadores de densidade usando núcleo é a escolha do parâmetro h . O parâmetro h , conhecido como janela (*bandwidth*), controla a ponderação dada aos pontos $z_i \neq z_0$. O parâmetro h controla a vizinhança de pontos utilizados na estimação de $\hat{f}(z_0, z)$. Valores menores de h levam a um menor número de pontos utilizado para estimação da densidade de torno do ponto z_0 , o que faz com que a densidade estimada para os dados seja menos suave. Na definição do parâmetro h utilizamos a regra de Silverman para um núcleo Gaussiano, que corresponde a 0.9 vezes o mínimo entre o desvio padrão dos dados e a diferença entre o menor e o maior quartil dos dados, multiplicado pelo tamanho da amostra e elevando este valor a potência de 1/5. As propriedades desta regra podem ser visualizadas em Silverman (1986, pp. 48-49).

Figura 8: Sequência de Densidades - Brasil - 1970 e 1996



A Figura 8 mostra a evolução das densidades estimadas, utilizando um núcleo Gaussiano, para o log das rendas relativas per capita nos anos de 1970 e 1996. A renda relativa é construída dividindo-se o valor da renda per capita do município dividida pela média das rendas per capita de todos os municípios no mesmo ano, e tomando-se o logaritmo deste valor. Nesta normalização, o valor zero no eixo horizontal indica a renda per capita igual a média nacional, um valor de 0.69 equivale a duas vezes a média nacional, e assim por diante. Deste modo, a renda relativa em algum ponto deste eixo é o log da renda relativa em relação a média do Brasil naquele ano.

Podemos observar na seqüência de densidades para a renda relativa dos municípios do Brasil a formação de duas modas na distribuição, o que foi chamado por Quah [16] de *Twin Peaks*. Em 1970 podemos observar a existência de um início de bimodalidade. Neste ano os picos inferiores e superiores da distribuição correspondem a -0.84 e -0.05 do log da renda relativa (respectivamente rendas per capita de 0.43 e 0.94 vezes a média brasileira naquele ano).

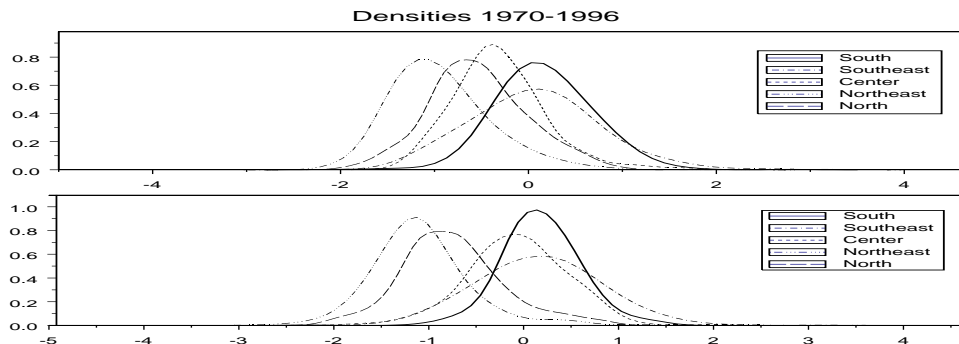
No ano de 1996 as posições do pico inferior e superior se deslocam para -1.08 e 0.06 (correspondentes a rendas relativas de 0.33 e 1.07 da média brasileira em 1996). Conforme a figura 9, os picos se tornam mais pronunciados e distantes em 1996 em comparação com o ano de 1970, sugerindo a formação de dois clubes de convergência para a renda relativa dos municípios brasileiros, um grupo formado pelos municípios ricos e outro composto por municípios pobres.

A Figura 9 mostra as densidades obtidas via núcleo das rendas relativas dos municípios das cinco regiões brasileiras para os anos de 1970 e 1996. Em ordem decrescente de renda relativa temos as regiões Sudeste, Sul, Centro-Oeste, Norte e Nordeste⁵. As duas principais mensagens da Figura 9 são a de que não há indicação de formação de clubes de convergência dentro de cada região, já que todas as densidades regionais são unimodais. Mais importante é o fato de que a distâncias entre as rendas relativas das regiões mais pobres (Norte e Nordeste) e as regiões mais ricas (Sudeste, Sul e Centro-Oeste) estão aumentando com o tempo. Em comparação com o ano de 1970, as densidades das regiões ricas se moveram para a direita, se tornando relativamente mais ricas. Os picos do Sudeste, Sul e Centro se moveram de 0.058, 0.048 e -0.446 em 1970 (ou seja 1.06, 1.05 e 0.64 vezes a média nacional), respectivamente, para 0.23, 0.029 e -0.1625 em 1996 (equivalente a 1.27, 1.03 e 0.85 vezes a renda média brasileira). O contrário ocorreu com as rendas das regiões mais pobres, Norte e Nordeste, que mudaram de -0.713 e -1.139 (0.49 e 0.32 vezes a média nacional), respectivamente, para -0.891 e -1.17 em 1996 (0.41 e 0.31 vezes a renda média no ano).

As densidades estimadas para o Brasil (Figura 8) e para as regiões (Figura 9), indicam que municípios das regiões Norte e Nordeste são em grande parte os formadores do pico de renda mais baixa, enquanto que os municípios das regiões Centro-Oeste, Sudeste e Sul formam o pico de renda mais elevada. Desta forma os municípios mais pobres em geral se tornaram relativamente mais

⁵O número de municípios analisados em cada região é Sudeste (1393), Sul (671), Centro-Oeste (226), Norte (160), e Nordeste (1331), conforme a tabela 6 presente no Apêndice.

Figura 9: Densidades - Regiões do Brasil - 1970 e 1996



pobres, enquanto que os municípios mais ricos se tornam relativamente mais ricos, o que equivale a definição da formação de clubes de convergência advogada por Quah [16].

5.1 Testes de Multimodalidade

Para verificar se a multimodalidade existente na densidade ajustada não parametricamente é estatisticamente significativa, utilizamos um teste de multimodalidade⁶ baseado no princípio do bootstrap proposto por Silverman [20], usando o algoritmo descrito por Efron e Tibshirani [7]. Dado que a densidade ajustada via núcleo não assume forma funcional ou distribuição, este teste de multimodalidade é baseado em encontrar via *bootstrap* uma distribuição de teste para a hipótese de m modas contra $m+1$ modas.

Como o número de modas encontradas na densidade estimada usando núcleo é função do bandwidth utilizado, Silverman [20] propõe usar a distancia entre o bandwidth que restringe m modas nos dados contra o bandwidth que determina $m+1$ modas como estatística de teste. Como o número de modas é uma função não crescente do bandwidth escolhido, o teste de multimodalidade se baseia em usar a distribuição de densidade ajustada dos dados em função do bandwidth mínimo necessário para induzir a hipótese nula de m modas.

Conforme Efron e Tibshirani [7], definimos a densidade ajustada como $\hat{f}(t, h_1)$ utilizando a equação 8, aonde t é o tamanho da amostra e h_1 a janela necessária para induzir o número de modas assumido na hipótese nula. A equação 8 artificialmente aumenta a variância da estimação, e por isso é necessário ajustá-la de forma a se tornar equivalente a da amostra, definindo uma nova densidade $\hat{g}(t, h_1)$, e a estatística de teste será construída utilizando o valor estimado para h_1 . Um valor maior de h_1 indica que um maior grau de suavização é necessário para induzir m modas na densidade em comparação com o valor ajustado pelo critério de Silverman para a janela de estimação.

O teste de hipótese baseado em replicações de *bootstrap* é obtido mantendo-se fixo o valor estimado para h_1 (janela mínima para se induzir m modas), e o nível de significância do teste é obtido pela probabilidade de que nas n replicações do bootstrap o valor mínimo h_1^* para se induzir m modas em cada replicação seja maior que o valor h_1 obtido com os dados observados. O nível de significância obtido via bootstrap é dado por:

$$NS_{boot} = Prob_{g(t^*, h_1)}\{h_1^* > h_1\} \quad (11)$$

Para obter a equação 11 utilizando-se replicações com a mesma variância dos dados originais usamos o *smooth bootstrap* de Efron e Tibshirani [7], pp. 231. Os níveis de significância foram

⁶Este teste de multimodalidade foi usado por Bianchi [5] para se testar a hipótese de convergência de renda para um grupo de 119 países entre os anos de 1970 e 1989. Bianchi [5] rejeita a hipótese de convergência em favor da formação de clubes de convergência.

Tabela 3: Testes de Multimodalidade

	NS 1 moda	NS 2 modas	h - Silverman	h - 1 moda	h - 2 modas
Brasil- 1970	0.8710	*	0.1319	0.1498	0.1289
Brasil - 1996	0.0474	0.7611	0.1398	0.3603	0.1047
Norte - 1970	0.4882	*	0.1528	0.1466	0.0898
Nordeste - 1970	0.2148	*	0.1042	0.3424	0.0684
Nordeste - 1996	0.2498	*	0.0878	0.1879	0.1722
Centro-Oeste 1970	0.6926	*	0.1340	0.1467	0.1062
Centro- Oeste 1996	0.6991	*	0.1505	0.1341	0.2889
Sudeste - 1970	0.7156	*	0.1399	0.1788	0.1540
Sudeste - 1996	0.7836	*	0.1320	0.1264	0.1091
Sul- 1970	0.5237	*	0.1168	0.1881	0.0857
Sul - 1996	0.7841	*	0.0898	0.0855	0.0459

obtidos com 2000 replicações de *bootstrap*.

Os resultados dos testes de multimodalidade (Tabela 3) aplicados ao log das rendas relativas para os municípios de todo o Brasil em 1996 mostram que obtemos um nível de significância empírico de 0.0474 para a hipótese nula de 1 moda, o que aponta que não a rejeitamos em apenas 4% das replicações. Isto indica que podemos rejeitar ao nível de significância de 5% que a distribuição das rendas relativas seja unimodal em favor da hipótese alternativa de bi-modalidade. Quando assumimos uma hipótese nula de 2 modas, o nível de significância empírico é de 0.7611, o que indica que não devemos rejeitá-la, sugerindo a formação de dois clubes de convergência para às rendas municipais no Brasil.

Os testes de multimodalidade também confirmam a evidência mostrada pelos gráficos das densidades ajustadas para as cinco regiões do país, de que cada região está convergindo para uma distribuição unimodal. O menor nível de significância para hipótese nula de unimodalidade em 1996 foi obtido para a região Nordeste com o valor de 0.2148, que visualmente é que apresenta uma densidade mais heterogênea. Este resultado é consistente com a interpretação de que a formação dos dois clubes de convergência para o Brasil é devido ao deslocamento das rendas relativas das regiões Norte e Nordeste para valores mais baixos e o deslocamento das rendas das Regiões Centro-Oeste, Sudeste e Sul para valores mais elevados, onde cada região se desloca mantendo um único pico.

6 Dinâmicas de Distribuição

Como a hipótese de convergência unimodal foi rejeitada pelos métodos não paramétricos, para modelar a evolução da distribuição relativa da renda per capita dos municípios brasileiros utilizamos a metodologia de Dinâmicas de Distribuição, que consiste em se modelar diretamente a evolução das distribuições relativas de renda como um processo de Markov de primeira ordem. Esta metodologia foi popularizadas pelos trabalhos de Quah [15, 17], e consiste no uso de Matrizes de Transição em uma formulação discreta e sua contraparte contínua através do uso de Núcleos Estocásticos.

A modelagem das Dinâmicas de Distribuição considera que a distribuição de densidade ϕ_t evolui de acordo com a seguinte equação:

$$\phi_{t+1} = M \cdot \phi_t, \quad (12)$$

sendo M um operador que mapeia a transição entre a distribuição das rendas existentes nos períodos t e t+1. Como a distribuição de densidade ϕ no período t só depende da densidade ϕ no período imediatamente anterior, isso é um processo de Markov de primeira ordem. Para capturar

a dinâmica das rendas relativas entre 1970 e 1996 precisamos de um operador M que determine a evolução do gráfico (a) para o gráfico (b) na Figura 8. A equação 12 pode ser vista como análoga a uma autoregressão de primeira ordem, onde substituímos pontos por distribuições completas.

A construção do operador M pode ser realizada ou assumindo-se que existe um número finito de estados para a distribuição ϕ_t , utilizando-se o modelo conhecido como Matrizes de Transição de Markov, ou então evitar a discretização e modelar M como uma variável contínua, no que é conhecido como núcleo Estocástico. A aplicação das matrizes de transição de Markov é realizada na seção 6.1, e a modelagem contínua usando Núcleos Estocásticos é realizada na seção 6.2.

6.1 Modelagem Discreta - Matrizes de Transição de Markov

Suponha que a probabilidade da variável s_t assumir algum valor particular j dependa somente do valor passado s_{t-1} conforme a seguinte equação:

$$P \{s_t = j | s_{t-1} = i, s_{t-2} = k, \dots\} = P \{s_t = j | s_{t-1} = i\} = P_{ij} \quad (13)$$

Este processo é descrito como uma cadeia de Markov de primeira ordem com n -estados, onde P_{ij} indica a probabilidade de que o estado i será seguido pelo estado j . Notando que

$$P_{i1} + P_{i2} + \dots + P_{in} = 1 \quad (14)$$

podemos construir o que é chamado de *matriz de transição*, onde a linha i e a coluna j , são a probabilidade de que o estado i será sucedido pelo estado j .

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix} \quad (15)$$

A utilização da cadeia de Markov para modelar a evolução da distribuição de renda relativa entre municípios brasileiros consiste na idéia de que cada estado desta matriz representa uma faixa de renda relativa. As matrizes de transição estimadas para as rendas relativas per capita dos municípios foram construídas de forma a ter nove estados, onde determinamos que nove faixas de renda seriam delimitadas pelo seguinte vetor de rendas relativas em relação ao valor médio da renda per capita do país⁷ no ano de análise: $\{0-0.25, 0.25-0.5, 0.5-0.75, 0.75-1, 1-1.25, 1.25-1.5, 1.5-1.75, 1.75-2, 2-\infty\}$ ⁸.

A probabilidade p_{ij} mede a proporção de municípios que estando no período anterior no regime i migra para o regime j no período corrente. Conforme Geweke, Marshall e Zarkin[12], o estimador de máxima verossimilhança para a probabilidade de transição p_{ij} é dado por:

$$\widehat{p}_{ij} = \frac{\sum m_{ij}}{\sum m_i}, \quad (16)$$

onde $\sum m_{ij}$ é o número de municípios que estavam na faixa de renda i no período anterior e migraram para a faixa de renda j no período corrente, e $\sum m_i$ é o total de municípios que estavam na faixa de renda i no período anterior.

Uma matriz de transição definida desta forma apresenta algumas características interessantes no estudo de mobilidade. A primeira é que, dadas às transições estimadas no período, as probabilidades de transição n períodos a frente podem ser previstas pela matriz de transição multiplicada por ela mesma n vezes, conforme Hamilton [13], cap 24. A segunda característica relevante é o fato de que as probabilidades de transição estimadas indicam as distribuições relativas de renda no longo prazo, conhecidas como distribuição ergódica.

⁷Ver discussão abaixo sobre os problemas da escolha ad hoc da discretização do número de estados da natureza.

⁸Esta mesma discretização é utilizada quando analisamos a transição dos municípios por região.

Tabela 4: Matriz de Transição - Brasil

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

Tabela 5: Distribuições Ergódicas

Renda	$-\infty - 0.25$	0.25-0.5	0.5-0.75	0.75-1	1-1.25	1.25-1.5	1.5-1.75	1.75-2	$2-\infty$
Norte	0.124	0.519	0.260	0.075	0.015	0.004	0.003	0.000	0.000
Nordeste	0.295	0.587	0.081	0.017	0.004	0.006	0.005	0.001	0.003
Centro-Oeste	0.001	0.014	0.056	0.092	0.233	0.144	0.180	0.092	0.187
Sudeste	0.003	0.042	0.085	0.137	0.129	0.131	0.106	0.080	0.287
Sul	0.001	0.016	0.069	0.252	0.246	0.176	0.107	0.062	0.070
Brasil	0.100	0.246	0.106	0.118	0.106	0.091	0.069	0.046	0.116

A distribuição ergódica pode ser encontrada ao notarmos que como a matriz de transição requer que cada linha da matriz some a unidade, um dos autovalores dessa matriz tem necessariamente o valor igual a um. Se os demais autovalores estão dentro do círculo unitário, a matriz de transição é dita ergódica e assim possui uma distribuição incondicional. Este vetor de distribuição incondicional, que em nosso caso vai representar as distribuições de longo prazo da renda relativa, é o autovetor associado com o autovalor unitário da matriz de transição.

A Tabela 4 mostra a matriz de transição estimada para os dados de rendas relativas dos municípios do Brasil. Na primeira coluna temos o número de municípios que estava em cada faixa de renda no ano de 1970. A matriz formada pelas linhas 2 de até 10 e colunas de 3 até 11 contém a matriz de probabilidades de transição entre as faixas de renda, e na última linha da matriz temos as distribuições ergódicas da rendas impostas pela matrizes de transição estimadas.

Os valores estimados mostram que existe uma mobilidade relativamente alta para as faixas de renda intermediária. Isso significa que quem era muito pobre tende a permanecer muito pobre (probabilidades de permanecer no mesmo nível de renda anterior de 0.39 e 0.56 para os dois extratos mais baixos de renda), e quem era muito rico tende a permanecer muito rico (probabilidade de 0.5 de permanência para a faixa mais elevada de renda), mas que as faixas intermediárias entre esses valores tendem mais a migrar para níveis inferiores ou superiores de renda do que permanecer no mesmo nível de renda, o que confirma a tendência do desaparecimento das faixas médias de renda.

As distribuições ergódicas da renda para o Brasil e para cada região em separado estão colocadas na Tabela 5. A distribuição de longo prazo mostra que para o Brasil temos dois picos, um de renda entre 0.25 e 0.5 vezes a média da renda no ano, com 0.246 do total de municípios do país, e outro pico com 0.118 dos municípios para rendas entre 0.75 e 1 vezes a média de rendas, sugerindo a formação de dois clubes de convergência. A distribuição da renda relativa em cada região é basicamente unimodal. Isto pode ser visto dado que para as regiões Nordeste e Norte, a distribuição de longo prazo se concentra na faixa de rendas entre 0.25 e 0.5 da renda média, enquanto que para

as regiões Centro-Oeste, Sudeste e Sul as distribuições de longo prazo se concentram nas rendas acima de 0.75 da renda média do país.

O resultado obtido sugere que a hipótese de convergência para um único ponto é rejeitada quando utilizamos as matrizes de transição de Markov. A evidência capturada pelas matrizes de transição de Markov sugerem que podemos considerar a existência de 2 picos de renda - um para os municípios mais pobres, entre 0.25 e 0.5 da renda per capita média do país, e outro para os municípios com renda acima de 0.75 do PIB per capita nacional. Estas evidências apontam para a existência de dois clubes de convergência para a renda relativa per capita entre municípios brasileiros.

Os resultados obtidos pelas matrizes de transição, formadas pela discretização do número de estados, são sujeitos a dois problemas graves. O primeiro é que a determinação do número de intervalos na matriz e os valores limites para cada intervalo são determinados de forma *ad hoc* pelo pesquisador, o que pode alterar sensivelmente os resultados obtidos. O segundo problema é que o processo de discretização pode eliminar a propriedade de dependência Markoviana existente nos dados, conforme Bulli [6], e é inerente ao processo de discretização a perda de informação, o que é conhecido na literatura de cadeias de Markov como o problema de *Alliasing*. A solução para este problema consiste em tornar a análise de transição contínua, evitando a discretização através do uso de densidades condicionais estimadas não parametricamente conhecidas como núcleo Estocástico, o que faremos em seguida.

6.2 Núcleo Estocástico

Para evitar os problemas associados a discretização na estimação de matrizes de transição, podemos estimar diretamente uma função de transição contínua entre as rendas relativas per capita em 1970 e 1996 de forma não paramétrica. Esta função de transição contínua recebe o nome de Núcleo Estocástico. Basicamente ela é a estimação de uma função de densidade condicional bivariada, aonde estamos condicionando a função nos valores da renda no ano inicial. Formalmente um núcleo estocástico é definido como:

Definição: Sejam $M_{(u,v)}$ e (R, \mathfrak{R}) espaços mensuráveis. Um **Núcleo Estocástico** em (M, \mathfrak{R}) é uma função $M_{(u,v)}(y, A) \times (R, \mathfrak{R}) \rightarrow [0, 1]$ tal que:

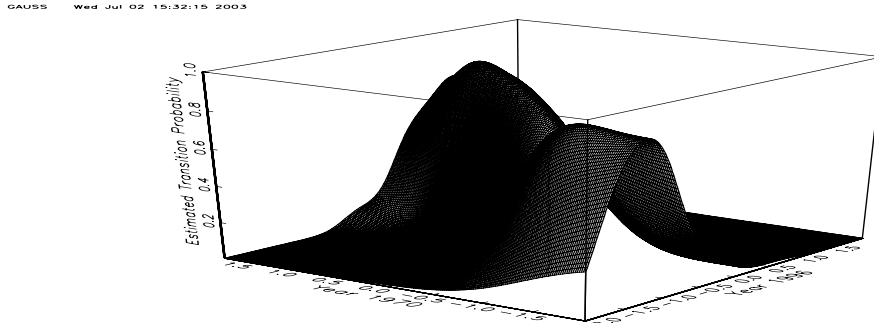
- para cada $y \in R$, $M(u, v)(y, A)$ é uma medida de probabilidade em (R, \mathfrak{R}) ;
- para cada $A \in \mathfrak{R}$, $M(u, v)$ é uma função mensurável em \mathfrak{R} ;
- para cada $A \in \mathfrak{R}$, é válido $u(A) = \int M_{(u,v)}(y, A) dv(y)$.

Condições a e b asseguram que o núcleo estocástico é um mapeamento bem definido para os espaços de probabilidade $M_{(u,v)}$ e (R, \mathfrak{R}) . O conceito principal do núcleo estocástico está na condição c. Dado um período inicial t , para uma dada renda y existe uma fração $dv(y)$ de economias com renda próxima a y . No período $t+n$, uma parte das economias que contidas em $dv(y)$ irão se mover para um sub-conjunto $A \subseteq R$. Normalizando esta fração de economias pelo número total de economias, teremos a definição do núcleo estocástico $M(u, v)(y, A)$. A integral $\int M_{(u,v)}(y, A) dv(y)$ representa o total de economias que independente da renda inicial estarão no sub-conjunto A de economias no período $t+n$. Nesta integral $M(u, v)(y, A)$ representa a total de economias que migraram de y para A , e $dv(y)$ é a ponderação associada a cada $M(\cdot)$ dada pela distribuição marginal de y . Quah [17] pp. 75-77 formaliza as demais condições necessárias. Desta forma o núcleo estocástico pode ser visualizado como a forma contínua da matriz de transição, onde temos um contínuo de linhas e colunas, ou seja, formamos um contínuo de estados.

A estimação do núcleo estocástico é realizada obtendo-se medidas empíricas para os elementos da integral $\int M_{(u,v)}(y, A) dv(y)$. A termo $\int M_{(u,v)}(y, A) dv(y)$ é obtido estimando-se não-parametricamente a densidade conjunta das rendas relativas nos períodos t e $t+n$ utilizando-se um *núcleo* bivariado, e esta densidade conjunta se torna um núcleo Estocástico ao normalizarmos pela distribuição marginal em t , que é a contraparte empírica de $dv(y)$. Uma propriedade importante é que a probabilidade de transição é independente do período de transição t , correspondendo a uma densidade de transição estacionária, conforme Quah [16] e Bulli [6].

A Figura 10 mostra o núcleo Estocástico estimado para os municípios do Brasil, e a Figura

Figura 10: Núcleo Estocástico - Brasil (1970-1996)



11 contém os Kernéis Estocásticos estimados para cada região em separado⁹. A interpretação da Figura 10 é a seguinte: a probabilidade de transição associada a mudança de um intervalo de rendas no período t para outro intervalo de rendas no período $t+n$ pode ser visualizada ao se calcular a probabilidade definida pelo núcleo estocástico na intersecção do intervalo de rendas no eixo t com o intervalo no eixo $t+n$. Esta projeção é análoga a célula formada pelo cruzamento intervalos de rendas nas linhas e colunas da matriz de transição.

A Figura 10 sugere que existem dois picos pronunciados no núcleo estocástico, mostrando que existem duas regiões de concentração de rendas, confirmando a formação dos dois clubes de convergência, um deles capturando a convergência para as rendas mais pobres e o outro pico capturando a convergência para a renda mais elevada. Esta figura também aponta que a probabilidade de transição estimada para as regiões de renda intermediária é muito baixa. Este resultado confirma o encontrado anteriormente na análise discreta, da tendência de desaparecimento das regiões de renda intermediária. Consistente os resultados mostrados nas Seções 5e 5.1, a hipótese de convergência para um único ponto não é válida para as rendas relativas per capita dos municípios brasileiros.

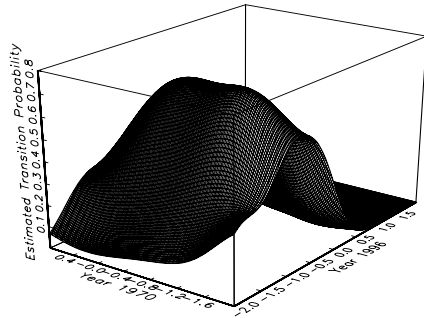
A estimação de kernéis estocásticos em separado para cada região do país (Figura 11) tem o objetivo de analisar se para cada região existe um único ponto de convergência, equivalente a apenas um pico no núcleo estocástico, se ou existiria alguma tendência a formação de clubes de convergência dentro de cada região.

A Figura 11 sugere que não existe em nenhum dos kernéis estocásticos estimados para cada região uma tendência significativa de formação de clubes de convergência. Este resultado é importante por mostrar que a formação dos dois clubes de convergência para os municípios do Brasil com um todo é causada pelo fato de que as regiões se movem para algum dos pontos de convergência. A outra hipótese possível seria a de que parte dos municípios de cada região se movessem para alguns dos clubes, o que seria mostrado pela existência de mais de um pico nos kernéis estocásticos estimados para cada região, o que não acontece segundo a Figura 11. Esta figura também mostra o deslocamento das rendas das regiões Norte e Nordeste para níveis mais baixos de rendas relativas, o que pode ser visto notando que existem maiores probabilidades de se migrar para regiões de renda relativa mais baixa em 1996. Este resultado contrasta com os obtidos para as regiões Centro-Oeste, Sudeste e Sul, que mostram maiores probabilidades de transição para níveis mais elevados de renda em 1996.

⁹Da mesma forma que a estimação do núcleo univariado, os Kernéis Estocásticos foram estimados com o uso de uma função de núcleo Gaussiana, e as janelas foram calculadas usando a regra de Silverman. Para facilitar a visualização, as figuras se concentram nas regiões com probabilidades mais elevadas de transição.

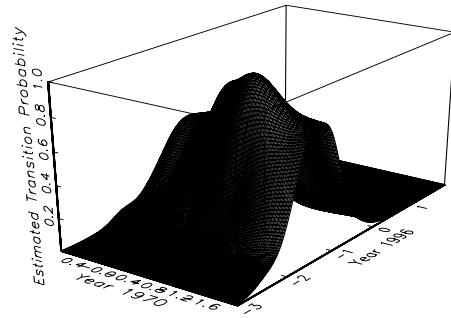
Figura 11: Núcleos Estocásticos - Regiões (1970-1996)

GAUSS Thu Jul 03 12:28:53 2003



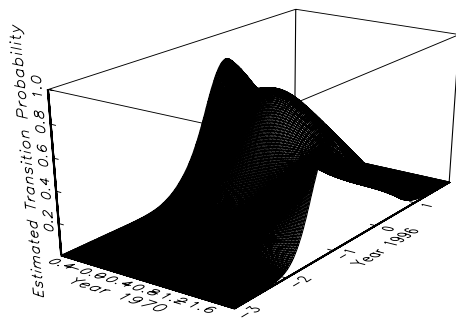
(a) Região Norte

GAUSS Thu Jul 03 12:33:02 2003



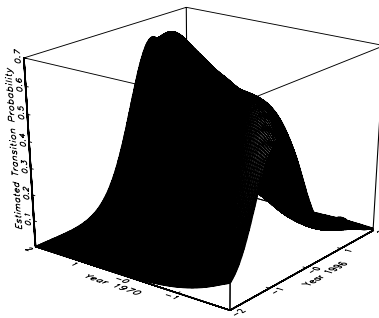
(b) Região Nordeste

GAUSS Thu Jul 03 12:37:33 2003



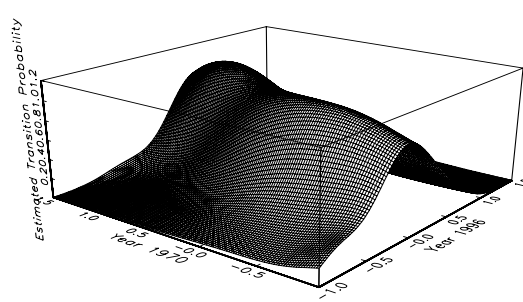
(c) Região Centro-Oeste

GAUSS Thu Jul 03 12:34:20 2003



(d) Região Sudeste

GAUSS Thu Jul 03 12:11:08 2003



(e) Região Sul

7 Conclusões

As análises realizadas neste artigo apontam para a formação de dois clubes de convergência de renda para os municípios brasileiros no período 1970-1996. Um clube de baixa renda formado pelas regiões Nordeste e Norte, e outro de renda mais elevada formado pelos municípios das regiões Centro-Oeste, Sudeste e Sul. As evidências mostram que o diferencial de renda relativa entre estes dois clubes de convergência se tornou ainda maior no ano de 1996 em relação aos valores em 1970.

A estimação das regressões de crescimento usando *Smoothing Splines* foi capaz de identificar a formação de clubes de convergência. Esta análise responde diretamente a hipótese básica da formação de clubes de convergência, que é a existência de níveis de renda tendem a desaparecer com o passar do tempo, com os municípios que estavam nesta faixa de renda intermediária migrando para níveis de rendas relativas superiores ou inferiores. Relaxando o pressuposto de uma mesma taxa de convergência para todos os municípios existentes na estimação por mínimos quadrados ordinários e possibilitando a estimação de uma taxa de convergência para cada nível de renda inicial, a estimação da regressão não paramétrica por *Smoothing Splines* permite a identificação de clubes de convergência. Estas estimações também sugerem que para cada região em separado não existe divergência, confirmando a hipótese de que a formação dos clubes é devido a um distanciamento uniforme da renda relativa das regiões Norte e Nordeste em relação aos valores das regiões Centro-Oeste, Sudeste e Sul.

A existência de dois picos na distribuição das rendas relativas para os municípios do Brasil foi mostrada pela estimação de densidades usando função de núcleo, e a esta bimodalidade foi confirmada com o uso de um teste para multimodalidade usando bootstrap. Este teste, aplicado em separado para cada região, mostra que em cada região não se pode rejeitar a hipótese nula de unimodalidade, confirmando que o processo de clubes de convergência é formado por um deslocamento uniforme das distribuições de renda relativa das regiões do Brasil.

Para capturar o processo de formação de clubes de convergência para os municípios do Brasil, modelamos a evolução temporal das rendas relativas no Brasil como um processo de Markov de primeira ordem, usando a metodologia de *Dinâmicas de Distribuição* advocadas por Quah [16]. Estimamos este processo em uma formulação discreta usando matrizes de transição, e na forma contínua usando o método de núcleo Estocástico. Os resultados sugerem que a dinâmica adequada para as rendas relativas para os municípios brasileiros não é o processo de convergência previsto pelo modelo neoclássico de crescimento, e sim a formação de dois clubes de convergência.

Os métodos não paramétricos utilizados se mostraram capazes de superar os problemas existentes nas estimações tradicionais de convergência, e são consistentes com os resultados de dois *clusters* de renda encontradas por Mossi et al [18], e também com os resultados obtidos por Ferreira et al [10]. Os resultados sugerem que os principais focos de pobreza relativa no país se localizam em primeiro lugar na região Nordeste, seguido da região Norte, e que as demais regiões estão em níveis mais elevados de renda. Os resultados apresentados aqui apontam que os novos modelos de crescimento econômico com modelos com armadilhas de pobreza, como os de Becker, Murphy e Tamura, podem ser uma representação adequada para a dinâmica da renda relativa dos municípios brasileiros.

Referências

- [1] Andrade, E., Laurini, M.P., Madalozzo, R. e Valls, P.L. 2002. Testing Convergence Across Municipalities Using Quantile Regression.
- [2] Azzoni, C.R. 2001. Economic Growth and Regional Income Inequality in Brazil. *The Annals of Regional Science* 35, pp 133-152
- [3] Barro, R. and Sala-i-Martin, X. 1992. Convergence. *Journal Of Political Economy* 100:2, pp. 223-251.

- [4] Becker, G., Murphy, K. M. e Tamura R. 1990. Human Capital, Fertility and Economic Growth. *Journal of Political Economy* 98:5, pp. 12-37.
- [5] Bianchi, M. 1997. Testing for Convergence - Evidence from Non Parametric Multimodality Tests. *Journal Of Applied Econometrics*, 12:4, pp. 393:410.
- [6] Bulli, S. 2001. Distribution Dynamics and Cross-Country Convergence: A New Approach. *Scottish Journal of Political Economy*, vol 48:2, pp. 226-243
- [7] Efron, B. e Tibshirani, R.J. (1993). *An Introduction to Bootstrap*. Chapman & Hall, Nova York.
- [8] Ferreira, A.F.H. 1998. Concentração Regional e Dispersão das Rendas Per Capita Estaduais: Um Comentário. *Texto para Discussão 121*, Cedeplar-UFMG.
- [9] Ferreira, A, H. e Diniz, C. C. 1995. Convergência Entre Rendas Per Capita Estaduais no Brasil. *Revista de Economia Política* 15:4, pp. 38-56.
- [10] Ferreira, A.F.H , Lanjouw, P. e Neri, Marcelo (2001). A Robust Poverty Profile for Brasil Using Multiple Data Sources. *World Bank Research Papers*.
- [11] Friedman, M. 1992. Do Old Fallacies Ever Die ?. *Journal of Economic Literature* 30, pp. 2129:2132.
- [12] Geweke, J. Marshall, R. C. e Zarkin, G. (1986). Mobility Indices in Continuous Time Markov Chains. *Econometrica* 54, pp 1407-1423.
- [13] Hamilton, J. D. 1994. *Time Series Analysis*. Princeton University Press.
- [14] Hardle, W. 1990. *Applied Nonparametric Regression*. Cambridge University Press.
- [15] Quah, D. 1993. "Galton's Fallacy and the Tests of the Convergence Hypothesis", *Scandinavian Journal of Economics*, 95, pp. 427-443.
- [16] Quah, D. (1996). Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *The Economic Journal* 106, pp 1045-1055.
- [17] Quah. D. (1998). *TsRf Reference Manual*.
- [18] Mossi, M.B., Aroca, P., Fernandez, I. J. e Azzoni, C.R. (2001). *Growth Dynamics and Space in Brazil*.
- [19] Ribeiro, E. P. e Pôrto Júnior, S. S (2000). Crescimento e Convergência - Uma Análise Empírica para a Região Sul. *Anais do Encontro de Economia da Região Sul 2000*.
- [20] Silverman, B.W. (1981). Using Kernel Density Estimates to Investigate Multimodality. *Journal of The Royal Statistical Society, Series B*, 43, 97-99.
- [21] Schwartsman, A. 1996. *Convergence Across Brazilian States*. Manuscript.
- [22] Solow, R. 1956. A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70, pp. 65:94.
- [23] Swan, T. W. 1956 Economic Growth and Capital Accumulation. *Economic Record* 32, pp. 334-361.
- [24] Wahba, G. 1983. Bayesian Confidence Intervals for the Cross Validated Smoothing Splines. *Journal of the Royal Statistical Society, Series B*, 45, 133-150.

Apêndice - Base de Dados

Tabela 6: Base de Dados - Municípios

Estado/ Região	Municípios existentes 1996	Municípios procedentes de mais que um município	Municípios procedentes de somente um município	Municípios sem origem	Municípios sem dados de PIB	Municípios existentes em 1970	Municípios sem dados de população em 1970	Total de municípios agrupados	Total de municípios usados nas estimações
PR	371	7	76	0	0	288	1	8	279
SC	259	2	59	0	0	198	0	5	193
RS	427	37	158	0	0	232	0	33	193
Sul	1057	47	292	0	0	718	1	46	671
ES	72	1	17	0	1	53	0	0	53
MG	755	1	31	0	0	723	0	10	713
RJ	81	0	18	0	0	63	0	0	63
SP	635	2	53	0	9	571	0	7	564
Sudeste	1543	4	199	0	10	1410	0	17	1393
GO	232	5	51	8	0	168	0	5	163
MT	115	32	47	2	0	34	1	10	23
MS	77	9	18	0	0	50	0	11	39
DF	1	0	0	0	0	1	0	0	1
Centro	425	46	116	10	0	253	1	26	226
MA	136	2	4	0	0	130	0	4	126
PI	148	4	30	1	0	113	0	7	106
CE	184	1	42	0	0	141	0	1	140
RN	152	1	1	0	0	150	1	3	146
PB	168	0	0	0	0	168	0	0	168
PE	177	2	9	1	0	165	0	7	158
AL	100	4	1	0	0	95	0	8	87
SE	75	0	1	0	0	74	0	1	73
BA	415	7	72	2	0	334	0	7	327
Nordeste	1555	21	160	4	0	1370	1	38	1331
AC	22	4	11	0	0	7	0	2	5
AP	15	4	6	0	0	5	0	1	4
AM	62	10	8	0	0	44	0	12	32
PA	128	10	35	0	0	83	0	12	71
RO	50	8	32	8	0	2	0	1	1
RR	8	2	4	0	0	2	0	1	1
TO	123	7	62	2	0	52	0	6	46
Norte	408	45	158	10	0	195	0	35	160
TOTAL	4998	163	845	24	10	3946	3	162	3781