

**UNIVERSIDADE DE BRASÍLIA
DEPARTAMENTO DE ECONOMIA**

**CONVERGÊNCIA DE RENDA *PER CAPITA*
ENTRE OS ESTADOS BRASILEIROS DE 1939 A 2004**

Ernani César e Silva Cabral

Tese apresentada à Universidade de Brasília como
Requisito para obtenção do título de doutor em
Ciência Econômicas.

Orientador: Prof. Dr. Stephen Anthony de Castro

Brasília, DF
2008

AGRADECIMENTOS

Gostaria em primeiro lugar de expressar meu profundo agradecimento ao professor e orientador Stephen de Castro, que se mostrou muito atencioso desde os momentos em que fora meu orientador da dissertação de Mestrado. Agradeço ao professor por toda a compreensão e paciência por esse longo período de formulação da Tese. Fico muito grato também pela ótima orientação em todas as etapas deste trabalho, como também pela ótima qualidade dos ensinamentos transmitidos seja como professor seja como orientador. Manifesto minha gratidão aos professores Roberto de Góes Ellery Júnior, Maria Eduarda Tannuri Pianto, Donald Matthew Pianto e Pedro Cavalcanti Ferreira cujas críticas contribuíram significativamente para o aperfeiçoamento da Tese. Fico grato também por todo o corpo docente do Departamento de Economia da Universidade de Brasília, por onde recebo ensinamentos de muita qualidade desde a época de graduação. Um agradecimento institucional ao Ministério do Planejamento, Orçamento e Gestão que me concedeu afastamento para estudos sob regime de dedicação exclusiva ao curso de doutoramento. Manifesto gratidão ao meu pai (*in memoriam*) que foi meu grande incentivador e melhor amigo, cujas orientações foram indispensáveis para minha evolução e equilíbrio. Fico grato também pelo suporte de meu irmão, Fred. Um agradecimento especial à minha mãe (*in memoriam*), que atravessou um momento recente muito difícil em sua vida, e me transmitiu sua qualidade de batalhadora e de dedicação a todos em sua volta. Estará sempre no meu coração.

Obrigado a todos!

RESUMO

Existe um extenso debate na literatura sobre a convergência de renda entre países ou regiões. Neste trabalho, busca-se avaliar a convergência de renda entre os estados brasileiros sob quatro diferentes óticas para o período de 1939 a 2004, e fazendo-se um destaque especial na análise dos períodos pré e pós milagre econômico. São realizadas estimativas β -convergência e σ -convergência para o corte transversal (*cross-section*), bem como se introduz uma variação no intervalo de tempo a fim de averiguar robustez dos resultados das regressões. Apesar de se verificar um aumento substancial de concentração de renda no período do milagre econômico, há uma nítida tendência de convergência de renda para o período pós milagre econômico. Aplica-se a abordagem das cadeias de Markov para o caso brasileiro, assim como são verificados os pressupostos necessários para utilização deste método. A matriz de transição representativa da amostra não se mostrou estatisticamente estável o que torna a técnica inapropriada para o caso brasileiro. Realiza-se também a decomposição da desigualdade a fim de se observar o efeito ultrapassagem, porém a alteração no ordenamento (*ranking*) dos estados brasileiros apresentou um papel diminuto em relação ao efeito de progressividade da renda. Por fim, utiliza-se a metodologia de dados de painel dinâmico de Phillips e Sul (2007) que possibilitou a constatação da existência de dois clubes de convergência para grupos de estados brasileiros, um baseado no Norte e Nordeste, regiões relativamente mais pobres, e outro no Centro-Sul.

ABSTRACT

In the literature an extensive debate exists on the convergence of income across countries or regions. In this thesis, we use four different methods to study income convergence across the Brazilian States for the period of 1939 through 2004. Separate analyses are performed of the periods before and after the so-called economic miracle, treated as the period 1968-1973. Cross-section estimates of β -convergence and σ -convergence are carried out. To check for robustness, small perturbations are introduced to the initial and terminal years. Despite a substantial increase in concentration of income during the period of economic miracle, clear trends of β -convergence and σ -convergence of income are found for the period after economic miracle. Markov Chains for the Brazilian States are constructed and the Markov assumptions tested statistically. The transition matrices did not exhibit statistical stability. This technique is thus shown to be inappropriate for the Brazilian case. Next, the change over time in income inequality across states is decomposed into three components representing β -convergence, σ -convergence and reranking (leapfrogging). Reranking was shown to have only a minimal effect on the changes in Gini. Finally, we use the dynamic panel-data methodology of Phillips and Sul (2007) to establish the existence of two convergence clubs for the period after the economic miracle, 1974-2004, one based in the relatively poor north and northeast regions, the other in the center-south.

SUMÁRIO

Introdução	1
1. Análise de β -Convergência e σ -Convergência	7
1.1. Introdução	7
1.2. Modelo	7
1.2.1. β -Convergência	7
1.2.2. Variações nos intervalos de tempo	10
1.2.3. σ -convergência	11
1.3. Estimação	14
1.4. Conclusão	21
2. Estimativa e Testes dos Pressupostos de Markov	22
2.1. Introdução	22
2.2. Abordagem das Cadeias de Markov	23
2.2.1. Modelo	23
2.2.2. Teste de homogeneidade no tempo	25
2.2.3. Teste de homogeneidade espacial	27
2.2.4. Outros testes	27
2.3. Estimativa e Avaliação dos Pressupostos	28
2.3.1. Teste de homogeneidade no tempo	29
2.3.2. Teste de homogeneidade espacial	30
2.3.3. Teste de propriedade de Markov	33
2.3.4. Teste de independência espacial	35
2.3.5. Teste de convergência nas regiões	37
2.4. Conclusão	38
3. Decomposição da Dinâmica da Distribuição Estadual Usando Gini	40
3.1. Introdução	40
3.2. Ilustrações dos Efeitos β -Convergência, σ -Convergência e Ultrapassagem	40
3.3. Metodologia	46
3.4. Análise dos Dados	49
3.5. Conclusão	51
4: Aferição de Convergência com Transições e Clubes	52
4.1. Introdução	52
4.2. Metodologia Econométrica	52
4.2.1. Representação do fator variável no tempo	52
4.2.2. Equilíbrio de longo prazo e convergência	54
4.2.3. Transição relativa	56
4.2.4. Testes de convergência	57
4.3. Análise Empírica	60
4.4. Conclusão	66
Conclusão Geral	67
Referências Bibliográficas	70
Apêndice	76

SUMÁRIO DE TABELAS

Capítulo 1

Tabela 1.1: Resultados das Regressões para Períodos Seleccionados	14
Tabela 1.2: Porcentagens de Coeficientes β Significativos e Negativos – Pré e Pós Milagre	15
Tabela 1.3: Porcentagens de Coeficientes β Significativos e Negativos – 1939 a 2004	15

Capítulo 2

Tabela 2.1: Matriz de Transição Anual Estimada para Estados Brasileiros entre os Anos de 1939 e 2003	29
Tabela 2.2: Matriz de Transição Anual Estimada para Estados Brasileiros entre os Anos de 1953 e 2003	30
Tabela 2.3: Teste de Homogeneidade Espacial entre as 5 Regiões Brasileiras de 1953 a 2003	31
Tabela 2.4: Teste de Propriedade de Markov para os Estados Brasileiros entre 1953 e 2003	33
Tabela 2.5: Teste de Dependência Espacial entre as Fronteiras dos Estados	36
Tabela 2.6: Distribuição Estacionária da Renda das Regiões Brasileiras de 1953 a 2003	38

Capítulo 3

Tabela 3.1: Decomposição da Desigualdade para Períodos Seleccionados	50
--	----

Capítulo 4

Tabela 4.1: Resultado das Regressões para Períodos Seleccionados	63
Tabela 4.2: Resultado das Regressões – Pré e Pós Milagre	64

SUMÁRIO DE GRÁFICOS E FIGURAS

GRÁFICOS

Capítulo 1

Gráfico 1.1: Evolução do Coeficiente de Variação Ponderado de Williamson e Índice de Theil no Período Pré-Milagre – 1939 a 1968	17
Gráfico 1.2: Evolução do Coeficiente de Variação Ponderado de Williamson e Índice de Theil no Período do Milagre – 1969 a 1973	18
Gráfico 1.3: Evolução do Coeficiente de Variação Ponderado de Williamson e Índice de Theil no Período Pós-Milagre – 1974 a 2004	19
Gráfico 1.4: Evolução do Coeficiente de Variação Ponderado de Williamson e Índice de Theil para Toda a Amostra – 1939 a 2004	20

Capítulo 3

Gráfico 3.1: Participação de cada Estado em Relação à Renda Média – 1939 e 2004	43
Gráfico 3.2: Efeito Ultrapassagem – 1939 e 2004	44
Gráfico 3.3: Efeito Ultrapassagem – 1939 e 1968	45
Gráfico 3.4: Efeito Ultrapassagem – 1974 e 2004	46

Capítulo 4

Gráfico 4.1: Evolução de H_t (Variância do Corte Transversal de h_{it}) – 1939 a 2004	61
Gráfico 4.2: Trajetórias Relativas de Transição, h_{it} – 1939 a 2004	62
Gráfico 4.3: Trajetórias Relativas de Transição, h_{it} – 1974 a 2004	65

Apêndice

Gráfico A.1: Evolução dos PIBs <i>per capita</i> para Estados Brasileiros – 1939 a 2004	78
---	----

FIGURAS

Capítulo 3

Figura 3.1: β -convergência, σ -convergência e sem ultrapassagem	41
Figura 3.2: β -convergência, σ -convergência e com ultrapassagem	41
Figura 3.3: β -convergência, não há σ -convergência e com ultrapassagem	42
Figura 3.4: Não há β -convergência, não há σ -convergência e sem ultrapassagem	42

Capítulo 4

Figura 4.1: Variação no Corte Transversal	58
---	----

INTRODUÇÃO

O conceito de convergência está relacionado à redução de desigualdade da renda *per capita* (renda média dos residentes) entre países ou regiões. Define-se β -convergência quando existe correlação parcial negativa entre o nível de renda inicial e o crescimento da renda no tempo e σ -convergência quando a dispersão da renda real *per capita* entre um grupo de economias diminui no tempo.

A denominação de convergência absoluta de renda adotada neste trabalho se refere à representação assintótica da razão das rendas *per capita* convergir no tempo para uma constante¹.

As principais abordagens econométricas que têm sido empregadas no estudo de convergência são o corte transversal (*cross-section*), as séries temporais e os dados de painel.

Na metodologia de corte transversal, os testes de convergência estão associados às investigações de β -convergência e σ -convergência. O estudo de β -convergência, investigado por Baumol (1986) e Barro e Sala-i-Martin (1992), é realizado mediante regressões de crescimento de renda ponderada no tempo sobre a renda inicial. Quando se inclui um conjunto de variáveis de controle às regressões, pesquisado por Barro (1991) e Barro e Sala-i-Martin (1991,1992), denomina-se β -convergência condicional. O exame de σ -convergência, analisado por Barro (1991) e Barro e Sala-i-Martin (1991,1992), verifica-se por análises de indicadores de medidas de dispersão na evolução da renda *per capita* das economias.

Por um lado, Friedman (1992) argumenta que o teste de convergência deve se concentrar na redução da variância entre os países (σ -convergência), por outro lado, Sala-i-Martin (1996a,b) sugere que tanto β -convergência quanto σ -convergência devem ser avaliados empiricamente.

¹ O conceito de convergência absoluta pode ser representado pela equação (6) do Capítulo 4.

Ferreira e Ellery Jr. (1996) buscam verificar a existência de convergência entre os PIBs *per capita* dos estados brasileiros bem como calcular a velocidade desse processo. Os dados utilizados foram obtidos do Anuário Estatístico do IBGE para os anos compreendidos entre 1970 e 1985, e foi realizada uma estimação para o ano de 1990.

A metodologia utilizada para esse trabalho foi baseada em Barro e Sala-i-Martin (1992). Foram avaliadas tanto a correlação negativa entre renda inicial e crescimento (β -convergência), quanto a evolução da dispersão do produto *per capita* dos estados (σ -convergência). Procurou-se ainda estimar a hipótese de rendimentos marginais decrescentes dos estados brasileiros.

Os resultados de Ferreira e Ellery Jr. (1996) constataam a existência de β -convergência e σ -convergência entre os PIBs *per capita* dos estados brasileiros para o período de 1970 a 1990.

A meia-vida (período de tempo necessário para reduzir a metade da distância) estimada para o caso brasileiro é em torno de 56,4 anos, o que revela uma lenta velocidade de convergência tendo em vista que esse valor é de aproximadamente 35 anos seja para os estados dos Estados Unidos seja para regiões da Europa, tal como estimado por Barro e Sala-i-Martin (1991).

Ferreira e Diniz (1995) encontram igualmente β -convergência e σ -convergência para os PIBs *per capita* dos estados brasileiros para o período entre 1970 e 1985. No entanto, Azzoni (2001), ao analisar para o período compreendido entre 1939 a 1996, encontra a presença de β -convergência e σ -convergência dos PIBs *per capita* dos estados apenas quando se restringe às regiões brasileiras.

Os testes de cortes transversais têm sido alvos de críticas. Segundo Bernard e Durlauf (1996), os testes de corte transversal impõem poucas restrições no comportamento de crescimento entre países se comparado à técnica de séries temporais uma vez que se negligencia o estado estacionário de longo prazo das economias. Quah (1996) alega que tais testes geram resultados viesados ao ignorar a dinâmica de mudança na distribuição de renda. Quah (1993a) mostrou formalmente que um coeficiente negativo dos níveis

iniciais em regressões de corte transversal pode ser consistente com o aumento de dispersão, cujo mecanismo ficou conhecido como “falácia de Galton”.

Outra forma de se testar convergência é mediante a aplicação de técnicas de séries temporais, onde o conceito de convergência é identificado não pelo confronto da renda inicial e o crescimento em um período t , mas na relação das previsões de longo prazo da renda *per capita*. Os testes mais comuns, realizados por Bernard e Durlauf (1995), Oxley e Greasley (1995), Li e Papell (1999) e Lee *et alii* (2005), são a verificação de raiz unitária e a análise de cointegração.

Bernard e Durlauf (1996) e Durlauf e Quah (1999) criticam o método de séries temporais uma vez que essa técnica parte do pressuposto que o processo estocástico gerador dos dados é invariante no tempo. Desse modo, se as economias estiverem em transição aos estados estacionários, compromete-se a avaliação de convergência definida por uma média zero ou ausência de raiz unitária.

Ainda que se introduza quebra estrutural nos testes de raiz unitária e se utiliza a técnica *vector autoregression* (VAR), vide Bernard e Durlauf (1995), ocorre uma tendência de rejeição da hipótese de convergência tendo em vista à falta de poder nos testes.

Bernard e Durlauf (1996) sugerem que um avanço em relação às abordagens de corte transversal (*cross-section*) e séries temporais seria uma integração da informação de transição do estudo de corte transversal e da análise de estado estacionário da técnica de séries temporais. Segundo os autores, a sugestão seria a abordagem de Quah (1993a,b) que estima uma função geral de transição de Markov para os dados e inferem uma distribuição limítrofe de corte transversal.

A adoção das cadeias heterogêneas de Markov, introduzida por Quah (1993b), consiste em um modelo de probabilidades de transição. Maiores detalhes para construção desta metodologia e avaliação dos pressupostos são expostos no Capítulo 2.

Ferreira (1998) faz a estimação de matrizes de transição de Markov para os dados de PIB *per capita* estaduais brasileiros para os anos de 1970 a 1995. Os resultados apontam para a tendência de concentração nas categorias médias de renda, o desaparecimento das

categorias de renda acima de 120% da média nacional e a manutenção na distribuição de renda dos extratos pobres na matriz de transição.

Em Quah (1997), o autor faz uma modificação do modelo discretizado de Markov ao estimar a densidade de distribuição pelo método de suavização por núcleo (*kernel smoothing*). Nesse caso, procura-se evitar a arbitrariedade do modelo discreto de suposição do número de intervalos, que pode levar a resultados viesados. O autor verifica que pela persistência e estratificação da distribuição da amostra, a metodologia aponta para a formação de clubes de convergência, referidos como “twin peaks”, onde ocorre a polarização das rendas *per capita* das economias pobres e ricas.

Mossi *et alii* (2003) utiliza matrizes de transição e núcleos estocásticos na análise da evolução das rendas *per capita* estaduais para os anos entre 1939 e 1998. Os principais resultados são a importância da geografia para o crescimento regional e a tendência de formação de dois clubes de convergência, um de baixa renda formado pelos estados da região norte e nordeste e outro de alta renda, pelos estados da região centro-oeste, sul e sudeste.

Da mesma forma, Gondim *et alii* (2007) utilizam a metodologia de núcleos estocásticos e encontram igualmente a propensão à existência de clubes de convergência no Brasil entre os anos de 1970 e 2000. Chega-se a conclusão que os principais condicionantes para a formação de clubes no Brasil são as variáveis de localização geográfica e nível inicial de escolaridade.

Outra maneira de se testar convergência é por meio de dados de painel. Os principais trabalhos a respeito de dados de painel são divididos em: a) testes de raiz unitária com tendência estocástica, tal como Evans e Karras (1996), Evans (1998) e Fleissig e Strauss (2001); b) avaliações que supõem taxas de crescimento homogêneas entre os países participantes da amostra, vide Kocenda (2001) e Kutan e Yigit (2004); c) exames que empregam técnicas de painel dinâmico sob a suposição de especificidade não observável de heterogeneidade de determinados países, por exemplo, Islam (1995) e Lee *et alii* (1997).

Segundo Durlauf e Quah (1999), existem vantagens e desvantagens na análise de dados de painel. No caso de painéis dinâmicos, torna-se mais clara a correlação entre as variáveis dependentes defasadas e o resíduo não observável. Além disso, há uma maior liberdade na especificação do modelo uma vez que é possível observar a estrutura dos erros em função da omissão de variáveis.

A desvantagem apontada por Durlauf e Quah (1999) na análise de dados de painel seria o excesso de liberdade dada aos parâmetros a fim de melhorar o ajuste nos dados, sem uma justificativa da teoria econômica.

Phillips e Sul (2007) propõem uma nova metodologia de dados de painel que permite a possibilidade de vários caminhos de transição. No modelo, a variação do fator crescimento se dá de maneira não linear e há a presença de componentes comuns e individuais das economias em transição. A grande vantagem dessa técnica é inferir sobre a formação de clubes de convergência.

Neste trabalho, o objetivo é fazer uma análise de convergência usando quatro diferentes metodologias para a maior amostra possível dos estados brasileiros, que data de 1939 a 2004. Uma vez que se observa um aumento significativo dos PIBs *per capita* da maioria dos estados brasileiros a partir do período do milagre econômico (1969-1973), Gráfico A.1 do Apêndice, busca-se enfatizar na maior parte das análises deste trabalho os períodos pré e pós milagre econômico.

Deve-se atentar ao fato que os dados dos PIBs estaduais brasileiros apresentam algumas inconsistências para o período pré-1970, mencionado por Ferreira (1996). Não obstante, os ajustes na série se limitaram aos descritos no apêndice deste trabalho.

No Capítulo 1, são realizados os testes β -convergência e σ -convergência para períodos selecionados, assim como se introduz uma variação no intervalo de tempo com o propósito de averiguar a robustez dos resultados das regressões.

Tendo em vista que as regressões realizadas para verificar β -convergência são alvos de críticas tal como mencionado anteriormente, procura-se, no Capítulo 2, aplicar a abordagem das cadeias de Markov para o caso dos estados e regiões brasileiras.

Ademais, são estudados os pressupostos para a estabilidade estatística da matriz de transição.

No Capítulo 3, faz-se uma decomposição da dinâmica da distribuição das rendas *per capita* estaduais usando o coeficiente de Gini com o objetivo de avaliar o processo de convergência. Emprega-se a técnica de Kakwani (1977), de tributação progressiva e redistribuição de renda, como forma de decompor a desigualdade em efeito progressivo ou regressivo da renda e efeito reordenamento. Por fim, aplica-se no Capítulo 4 a metodologia introduzida por Phillips e Sul (2007) de dados de painel dinâmico que permite a identificação da formação de clubes de convergência, bem como avaliar medidas de convergência ao longo das trajetórias das economias.

1. Análise de β -Convergência e σ -Convergência

1.1. Introdução

A proposta deste capítulo é fazer análises de β -convergência e σ -convergência para os dados dos estados brasileiros de 1939 a 2004, especialmente para os períodos pré e pós milagre econômico (1969-1973)². Além disso, aplica-se um método de variação nos intervalos de tempo a fim de averiguar robustez dos resultados das regressões.

1.2. Modelo

A próxima subseção é destinada a exibir um modelo para aferição de β -convergência, e na subseção seguinte, será apresentado o método de variação nos intervalos de tempo. Na subseção 1.3.3, é mostrada a relação entre β e σ convergência.

1.2.1. β -Convergência

King e Rebelo (1989) e Barro e Sala-i-Martin (1990) fazem a seguinte representação para o modelo neoclássico de crescimento. A função de produção para um produtor representativo é

$$Y = F(K, Le^{gt}) \quad (1)$$

tal que Y é o fluxo de produto, K é o estoque de capital, e^{gt} representa o efeito exógeno do progresso técnico que amplia a força de trabalho, L .

Exclui-se a possibilidade de escolha entre trabalho e lazer e supõe-se pleno emprego, resultando ambos que a força de trabalho se torna equivalente ao tamanho populacional. Uma das principais características de um modelo neoclássico de crescimento é o retorno decrescente na reprodução do capital. Assim, assume-se que $F(\cdot)$ exibe retorno positivo

² Define-se neste trabalho que o período do milagre econômico compreende os anos entre 1969 e 1973. Existe uma pequena variação na literatura aos anos que se referem o período do milagre econômico. A título de exemplo, Giambiagi (1988) entende que o milagre econômico se refere ao período entre 1968 e 1973, ao passo que Lago (1990), entre 1967 e 1973.

e produto marginal negativo para cada insumo, bem como se adota a hipótese de retornos constantes de escala.

A equação (1) pode ser reescrita seguindo a função de produção Cobb-Douglas:

$$\hat{y} = f(\hat{k}) = (\hat{k})^\alpha \quad (2)$$

onde o símbolo $\hat{\cdot}$ representa a quantidade por unidade de trabalho efetivo, $L e^{gt}$, e $f(\cdot)$ satisfaz as condições $f' > 0$ e $f'' < 0$, tal que $0 < \alpha < 1$.

O produto em uma economia fechada pode ser utilizado para investimento, \dot{K} , ou consumo, C . Sendo que δ representa a taxa de depreciação do capital, a evolução de \hat{k} segue de acordo com a equação:

$$\dot{\hat{k}} = f(\hat{k}) - \hat{c} - (n + g + \delta)\hat{k} \quad (3)$$

tal que n é a taxa de crescimento exógena da força de trabalho (ou população).

Uma família representativa de horizonte infinito busca maximizar a função de utilidade dada por:

$$U = \int_0^{\infty} u(c) e^{nt} e^{-\rho t} dt \quad (4)$$

onde $c = C/L$ e ρ é a taxa constante de preferência intertemporal.

A condição de primeira ordem de maximização da equação (4) resulta em

$$\frac{\dot{c}}{c} = \left(\frac{1}{\theta} \right) \cdot [f'(k) - \delta - \rho] \quad (5)$$

tal que θ significa a elasticidade de substituição intertemporal.

Supondo tecnologia do tipo Cobb-Douglas, King e Rebelo (1989) realizam uma aproximação em torno do estado estacionário da log-linearização das equações (3) e (5) e chegam a seguinte equação:

$$\log [\hat{y}(t)] = \log [\hat{y}(0)] \cdot e^{-\gamma t} + \log (\hat{y}^*) \cdot (1 - e^{-\gamma t}) \quad (6)$$

Vale ressaltar que o parâmetro positivo γ , que reflete a velocidade de ajuste ao estado estacionário, submete-se a seguinte restrição dos parâmetros do modelo:

$$2\gamma = \left\{ h^2 + 4 \left(\frac{1-\alpha}{\theta} \right) (\rho + \delta + \theta g) \left[\frac{\rho + \delta + \theta g}{\alpha} - (n + \delta + g) \right] \right\}^{\frac{1}{2}} - h$$

onde $h = \rho - n - (1 - \theta)g > 0$.

A partir da equação (6), é possível obter a taxa média de crescimento do produto *per capita* (y) compreendido entre os intervalos t_0 e t_{0+T} :

$$\left(\frac{1}{T} \right) \cdot \log \left[\frac{y(t_0 + T)}{y(t_0)} \right] = g + \frac{(1 - e^{-\gamma T})}{T} \cdot \left\{ \log \left[\frac{y^*}{y(t_0)} \right] \right\} \quad (7)$$

Em termos econométricos, a equação (7) pode ser representada por:

$$\left(\frac{1}{T} \right) \log \left(\frac{y_{it_T}}{y_{it_0}} \right) = \alpha + \beta \log(y_{it_0}) + u_{i_0,T} \quad (8)$$

tal que α e β sejam constantes, y_{it_T} e y_{it_0} sejam, respectivamente, a renda *per capita* na região i no tempo T e 0 , e $u_{i_0,T}$ representa o termo de erros médios (u_{it}) entre os períodos 0 e T .

Se β for menor que zero, uma renda *per capita* inicial baixa vai resultar em taxas médias de crescimento maiores nos anos seguintes. Portanto, os dados investigados exibem β -

convergência se o β estimado por estimação de mínimos quadrados (EMQ/OLS) for negativo.

Uma variação desse modelo, formulada por Mankiw, Romer e Weil (1992), surge quando se implementa variáveis de controle, X , à equação (8).

$$\left(\frac{1}{T}\right)\log\left(\frac{y_{it_T}}{y_{it_0}}\right) = \alpha + \beta \log(y_{it_0}) + \Pi X_j + u_{i_0,T} \quad (9)$$

No caso de $\beta < 0$, diz-se que há β -convergência condicional de renda. O nível de produto *per capita* de equilíbrio de longo prazo pode diferir, desde que cada economia convirja para seu próprio estado estacionário. No entanto, as economias pertencentes a um mesmo grupo, definido pelas variáveis condicionais, a diferença entre os produtos tende a diminuir.

1.2.2. Variações nos intervalos de tempo

A idéia de se introduzir uma variação nos intervalos de tempo tem como objetivo conferir robustez aos resultados das regressões de análise de β -convergência. São realizadas mais regressões em torno do período inicial e final considerado para permitir uma melhor avaliação da consistência dos resultados das regressões.

A nomenclatura para os intervalos de tempo é descrita da seguinte forma:

$t_{j,+/-1}$: $t_{j,-1}$, $t_{j,0}$, $t_{j,1}$

$t_{j,+/-2}$: $t_{j,-2}$, $t_{j,-1}$, $t_{j,0}$, $t_{j,1}$, $t_{j,2}$

$t_{j,+/-3}$: $t_{j,-3}$, $t_{j,-2}$, $t_{j,-1}$, $t_{j,0}$, $t_{j,1}$, $t_{j,2}$, $t_{j,3}$

...

$t_{j,+/-i}$: $t_{j,-i}$, $t_{j,-(i-1)}$, ..., $t_{j,-1}$, $t_{j,0}$, $t_{j,1}$, ..., $t_{j,(i-1)}$, $t_{j,i}$

tal que $j = 0$ representa o intervalo inicial, $j = 1$, o final e i o número de vizinhanças escolhido.

Para o caso de o número de vizinhanças escolhido, i , seja até três, as regressões a serem feitas são:

$i = 1$

$t_{j,+/-1}$: 9 regressões ($t_{0,-1,t_{1,-1}}$; $t_{0,-1,t_{1,0}}$; $t_{0,-1,t_{1,1}}$; $t_{0,0,t_{1,-1}}$; $t_{0,0,t_{1,0}}$; $t_{0,0,t_{1,1}}$; $t_{0,1,t_{1,-1}}$; $t_{0,1,t_{1,0}}$; $t_{0,1,t_{1,1}}$)

$i = 2$

$t_{j,+/-2}$: 25 regressões ($t_{0,-2,t_{1,-2}}$; $t_{0,-2,t_{1,-1}}$; $t_{0,-2,t_{1,0}}$; $t_{0,-2,t_{1,1}}$; $t_{0,-2,t_{1,2}}$; $t_{0,-1,t_{1,-2}}$; $t_{0,-1,t_{1,-1}}$; $t_{0,-1,t_{1,0}}$; $t_{0,-1,t_{1,1}}$; $t_{0,-1,t_{1,2}}$; $t_{0,0,t_{1,-2}}$; $t_{0,0,t_{1,-1}}$; $t_{0,0,t_{1,0}}$; $t_{0,0,t_{1,1}}$; $t_{0,0,t_{1,2}}$; $t_{0,1,t_{1,-2}}$; $t_{0,1,t_{1,-1}}$; $t_{0,1,t_{1,0}}$; $t_{0,1,t_{1,1}}$; $t_{0,1,t_{1,2}}$; $t_{0,2,t_{1,-2}}$; $t_{0,2,t_{1,-1}}$; $t_{0,2,t_{1,0}}$; $t_{0,2,t_{1,1}}$; $t_{0,2,t_{1,2}}$)

$i = 3$

$t_{j,+/-3}$: 49 regressões ($t_{0,-3,t_{1,-3}}$; $t_{0,-3,t_{1,-2}}$; $t_{0,-3,t_{1,-1}}$; $t_{0,-3,t_{1,0}}$; $t_{0,-3,t_{1,1}}$; $t_{0,-3,t_{1,2}}$; $t_{0,-3,t_{1,3}}$; $t_{0,-2,t_{1,-3}}$; $t_{0,-2,t_{1,-2}}$; $t_{0,-2,t_{1,-1}}$; $t_{0,-2,t_{1,0}}$; $t_{0,-2,t_{1,1}}$; $t_{0,-2,t_{1,2}}$; $t_{0,-2,t_{1,3}}$; $t_{0,-1,t_{1,-3}}$; $t_{0,-1,t_{1,-2}}$; $t_{0,-1,t_{1,-1}}$; $t_{0,-1,t_{1,0}}$; $t_{0,-1,t_{1,1}}$; $t_{0,-1,t_{1,2}}$; $t_{0,-1,t_{1,3}}$; $t_{0,0,t_{1,-3}}$; $t_{0,0,t_{1,-2}}$; $t_{0,0,t_{1,-1}}$; $t_{0,0,t_{1,0}}$; $t_{0,0,t_{1,1}}$; $t_{0,0,t_{1,2}}$; $t_{0,0,t_{1,3}}$; $t_{0,1,t_{1,-3}}$; $t_{0,1,t_{1,-2}}$; $t_{0,1,t_{1,-1}}$; $t_{0,1,t_{1,0}}$; $t_{0,1,t_{1,1}}$; $t_{0,1,t_{1,2}}$; $t_{0,1,t_{1,3}}$; $t_{0,2,t_{1,-3}}$; $t_{0,2,t_{1,-2}}$; $t_{0,2,t_{1,-1}}$; $t_{0,2,t_{1,0}}$; $t_{0,2,t_{1,1}}$; $t_{0,2,t_{1,2}}$; $t_{0,2,t_{1,3}}$; $t_{0,3,t_{1,-3}}$; $t_{0,3,t_{1,-2}}$; $t_{0,3,t_{1,-1}}$; $t_{0,3,t_{1,0}}$; $t_{0,3,t_{1,1}}$; $t_{0,3,t_{1,2}}$; $t_{0,3,t_{1,3}}$)

Generalizando, o número de regressões é igual a:

$$(2i+1)^2 \tag{10}$$

A constatação de β -convergência em torno de uma vizinhança i ocorre se o número coeficientes β significativos e menores que zero representarem a maioria absoluta (acima de 50%) das regressões realizadas.

1.2.3. σ -convergência

Nesta subseção, busca-se demonstrar, conforme metodologia de Young *et alii* (2007), que β -convergência é condição necessária mas não suficiente para σ -convergência.

Sala-i-Martin (1996) apresenta uma pequena variação ao modelo exposto na subseção 1.2.1. Suponha que se verifique β -convergência para as economias $i = 1, \dots, N$. O logaritmo natural da i -ésima economia pode ser aproximado por:

$$\log(y_{it}) = a + (1 - \beta) \log(y_{i,t-1}) + u_{it} \quad (11)$$

onde $0 < \beta < 1$ e u_{it} possua média zero e variância finita, σ_u^2 , e seja independente sobre t e i .

Como se trata de β -convergência, presume-se que a seja uma constante entre as economias, ou seja, as trajetórias de crescimento equilibrado são idênticas. Young *et alii* (2007) fazem a ressalva que supor um modelo de convergência condicional teria impacto apenas no aumento do conjunto de cenários em que β -convergência não implica necessariamente em σ -convergência.

A equação (11) pode ser reescrita como:

$$\log\left(\frac{y_{it}}{y_{i,t-1}}\right) = a - \beta \log(y_{i,t-1}) + u_{it} \quad (11)'$$

Ao contrário da equação (8), $\beta > 0$ implica na correlação negativa entre crescimento e logaritmo da renda inicial.

A variância da amostra do logaritmo da renda em t é dada por

$$\sigma_t^2 = \left(\frac{1}{N}\right) \sum_{i=1}^N [\log(y_{it}) - \mu_t]^2 \quad (12)$$

tal que μ_t é a média da amostra do logaritmo da renda.

A variância da amostra é próxima à da população quando N é grande, e, desse modo, deriva-se a evolução de σ_t^2 por meio da equação (11):

$$\sigma_t^2 \cong (1 - \beta)^2 \sigma_{t-1}^2 + \sigma_u^2 \quad (13)$$

Portanto, a equação em diferenças (13) somente é estável se $0 < \beta < 1$. Se $\beta \leq 0$, a variância aumenta no tempo. E se $\beta = 1$, a variância vai permanecer constante, ao passo que se $\beta > 1$, a correlação parcial entre o logaritmo da renda e seus valores nos períodos anteriores iriam oscilar entre valores negativos e positivos (pouco sentido econômico).

Adotando-se $0 < \beta < 1$, a variância de estado estacionário é:

$$(\sigma^2)^* = \frac{\sigma_u^2}{[1 - (1 - \beta)^2]} \quad (14)$$

Nota-se que a dispersão no corte transversal (*cross-section*) diminui com β e aumenta com σ_u^2 . Fazendo as substituições entre (13) e (14), obtém-se:

$$\sigma_t^2 = (1 - \beta)^2 \sigma_{t-1}^2 + [1 - (1 - \beta)^2](\sigma^2)^* \quad (15)$$

A equação (15) é uma equação em diferenças linear de primeira ordem com coeficientes constantes. A solução é dada por:

$$\sigma_t^2 = (\sigma^2)^* + (1 - \beta)^{2t} [\sigma_0^2 - (\sigma^2)^*] + c(1 - \beta)^{2t} \quad (16)$$

onde c é uma constante arbitrária.

Sendo que $0 < \beta < 1$, temos que $|1 - \beta| < 1$, o que implica:

$$\lim_{t \rightarrow \infty} (1 - \beta)^{2t} = 0 \quad (17)$$

A última relação garante a estabilidade de σ_t^2 visto que:

$$\lim_{t \rightarrow \infty} \sigma_t^2 = (\sigma^2)^* \quad (18)$$

1.3. Estimação

A Tabela 1.1 a seguir mostra os resultados das regressões realizadas de acordo com a equação (8) para os períodos selecionados.

Tabela 1.1
Resultados das Regressões para Períodos Selecionados

Período	Coefficiente $\hat{\beta}$	Desvio-padrão	Estatística t	Probabilidade
1939 a 2004	-0,0023	0,0021	-1,15	0,26
1939 a 1968	-0,0061	0,0038	-1,58	0,13
1974 a 2004	-0,0129	0,0022	-5,75	0,00
1969 a 1973	0,1012	0,0377	2,69	0,01

Observa-se na Tabela 1.1 que não se verifica β -convergência nos períodos selecionados exceto para o período pós-milagre (1974-2004) tendo em vista que é o único intervalo em que se verifica o coeficiente $\hat{\beta}$ significativo (prob. < 0,01) e negativo. É interessante notar que no período do milagre (1969 a 1973) ocorre um efeito justamente contrário à β -convergência. Um $\hat{\beta}$ significativo e positivo corresponde ao crescimento médio de estados mais pobres a taxas inferiores aos estados mais ricos. Ou seja, há um aumento na desigualdade nesse período.

Na Tabela 1.2 a seguir, são apresentadas as porcentagens de coeficientes β significativos e negativos de acordo com os níveis de significância (α). A vantagem dessa metodologia, exposta na subseção 1.2.2, é variar o intervalo de tempo nas regressões de tal maneira a se obter mais informações sobre determinado período de tempo da amostra.

Tabela 1.2

Porcentagens de Coeficientes β Significativos e Negativos – Pré e Pós Milagre

Nível de significância (α)	Pré-Milagre 1940 \pm 1;1967 \pm 1	Pós-Milagre 1975 \pm 1;2003 \pm 1
$\alpha = 1\%$	44,4%	100,0%
$\alpha = 5\%$	66,7%	100,0%

Como se nota na Tabela 1.2, não há uma clara constatação de β -convergência para o período pré-milagre econômico e nível de significância igual a 1% uma vez que quatro das nove (44,4%) regressões³ em torno das vizinhanças de um ano para mais e para menos apresentam coeficientes significativos e menores que zero. No entanto, seis das nove regressões (66,7%) indicam β -convergência quando o nível de significância é igual a 5%. Para o período pós-milagre, todas as nove regressões, seja para o nível de significância igual a 1% ou 5%, apontam para a β -convergência.

Por fim, cabe fazer a análise de variação no intervalo de tempo para todo o período amostral, de 1939 a 2004. As porcentagens de coeficientes β significativos e menores que zero são exibidas na Tabela 1.3.

Tabela 1.3

Porcentagens de Coeficientes β Significativos e Negativos – 1939 a 2004

Nível de significância (α)	$t_{j,+/-1}$ 1940 \pm 1;2003 \pm 1	$t_{j,+/-2}$ 1941 \pm 2;2002 \pm 2
$\alpha = 1\%$	Nenhum	20,0%
$\alpha = 5\%$	55,6%	68,0%

³ Por exemplo, as nove regressões de $t \pm 1$ de 1940-1967 são realizadas de acordo com os seguintes intervalos: entre 1939 e 1966, 1939 e 1967, 1939 e 1968, 1940 e 1966, 1940 e 1967, 1940 e 1968, 1941 e 1966, 1941 e 1967, e 1941 e 1968.

Para um nível de significância de 1%, nenhuma das nove regressões apontou para β -convergência quando se varia um ano no intervalo de tempo para todo período amostral e apenas 5 das 25 regressões, ou seja, 20% indicaram β -convergência quando se varia dois anos no intervalo de tempo. Por outro lado, quando se considera um nível de significância de 5%, a maioria absoluta das regressões seja para uma ou duas variações no intervalo de tempo indicaram a presença de β -convergência.

A análise de σ -convergência na qual consiste verificar se a dispersão dos níveis dos PIBs *per capita* das economias diminui com o tempo será realizada graficamente, e terá como referência o coeficiente de variação ponderado de Williamson e o índice de Theil.

Williamson (1965) propõe um indicador baseado na definição do desvio padrão que mede a dispersão dos níveis de renda *per capita* regionais em relação à média nacional, onde os desvios apresentados por cada região são ponderados pela sua participação na população nacional. Desse modo, o coeficiente ponderado de Williamson (V_w) pode ser apresentado conforme a seguinte equação.

$$V_w = \frac{\sqrt{\sum_{i=1}^N \left(\frac{Y_i}{P_i} - \frac{Y}{P} \right)^2 \frac{P_i}{P}}}{\frac{Y}{P}} \quad (19)$$

onde Y_i é a renda do estado i , P_i é a população do estado i , Y é a renda nacional, P é a população nacional e N é o número de estados que compõem o país.

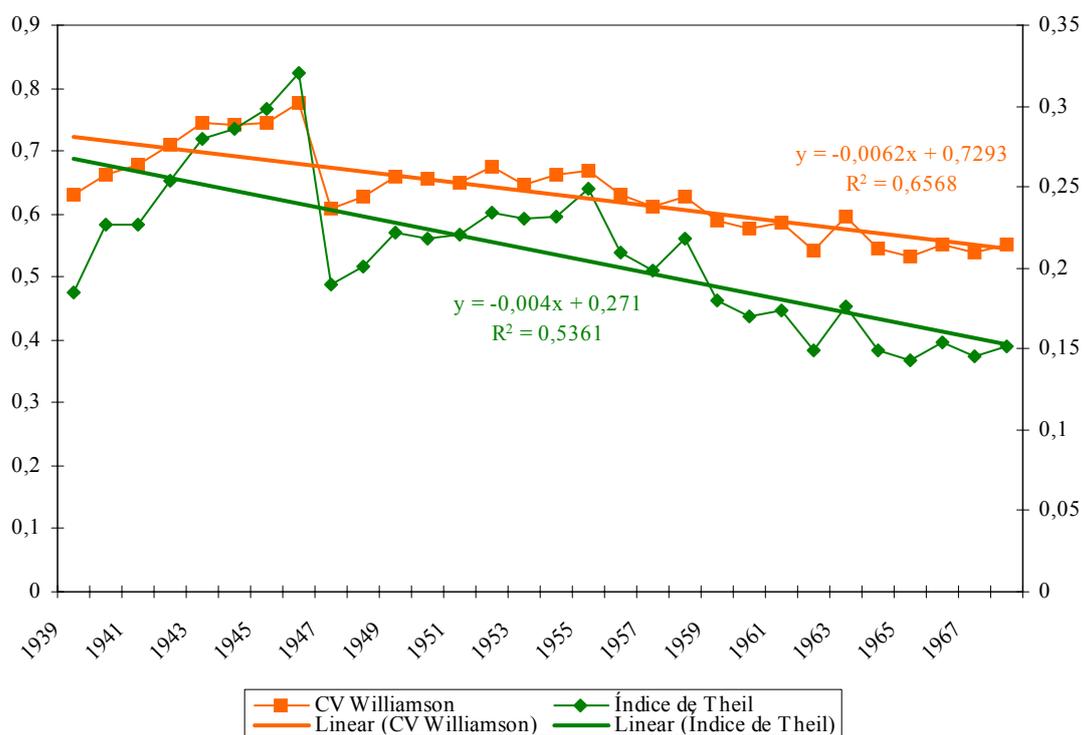
O índice de Theil tem sido bastante utilizado nas avaliações de desigualdades regionais. Para um país dividido em N estados, e definindo p_i (P_i / P) e y_i (Y_i / Y) como as frações de população e renda para cada estado i , o índice de Theil de desigualdade interestadual (J_E) é apresentado pela seguinte equação.

$$J_E = \sum_{i=1}^N p_i \log \left(\frac{p_i}{y_i} \right) \quad (20)$$

Sobre a evolução da desigualdade regional brasileira, destacam-se os trabalhos de Azzoni (2001), Silva e Medina (1999), Ferreira (1998), Azzoni (1997) e Ferreira e Diniz (1995).

Exibe-se no Gráfico 1.1 a seguir a evolução do coeficiente de variação ponderado de Williamson e do índices de Theil para o período pré milagre econômico. É importante mencionar que como existem dois eixos de ordenadas nos quatro gráficos subseqüentes, o eixo da esquerda corresponde aos valores do coeficiente de variação ponderado de Williamson, e o eixo da direita representa os valores do índice de Theil.

Gráfico 1.1
Evolução do Coeficiente de Variação Ponderado de Williamson e Índice de Theil no Período Pré-Milagre – 1939 a 1968

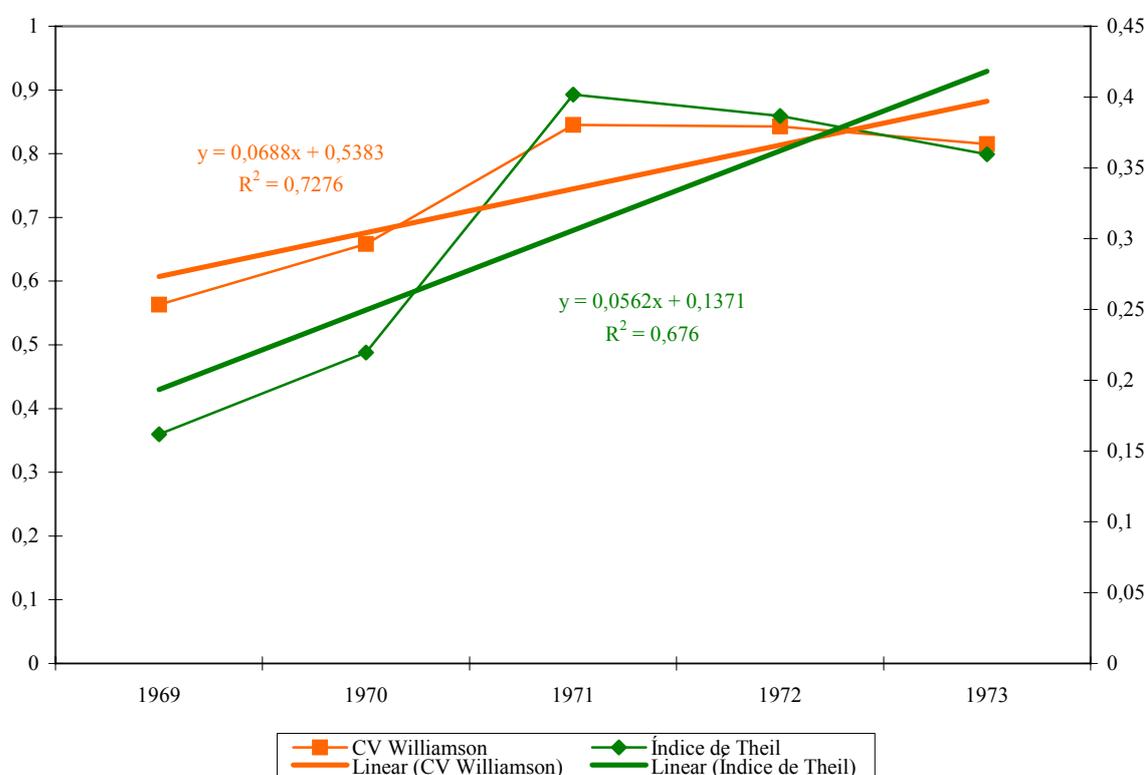


Verifica-se que de 1939 a 1968 ambos os indicadores apresentaram uma leve queda. Nesse período, o coeficientes de variação ponderado de Williamson e o índice de Theil

reduziram, respectivamente, 12,45% e 17,56%. Ademais, foram estimadas linhas de tendência para os dois indicadores cujas inclinações negativas refletem a propensão de queda das variáveis no período mencionado.

No Gráfico 1.2, é mostrada a evolução dos indicadores no período do milagre econômico.

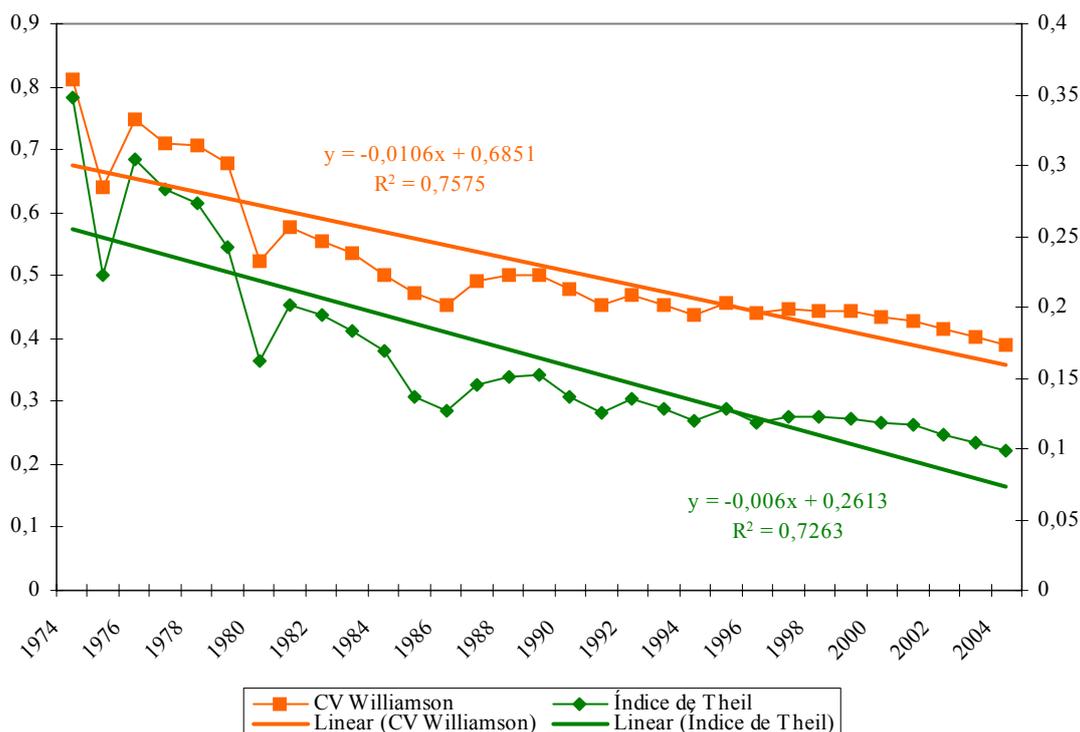
Gráfico 1.2
Evolução do Coeficiente de Variação Ponderado de Williamson e Índice de Theil no Período do Milagre – 1969 a 1973



Entre os anos de 1969 e 1973, o coeficiente de variação ponderado de Williamson aumentou 44,77% ao passo que o índice de Theil elevou 122,23%. Isto é, o período do milagre econômico se caracterizou por uma forte concentração de renda *per capita* ao nível estadual.

O Gráfico 1.3 a seguir destaca o período pós milagre econômico.

Gráfico 1.3
Evolução dos Coeficientes de Variação Ponderado de Williamson e Índices de Theil no Período Pós-Milagre – 1974 a 2004

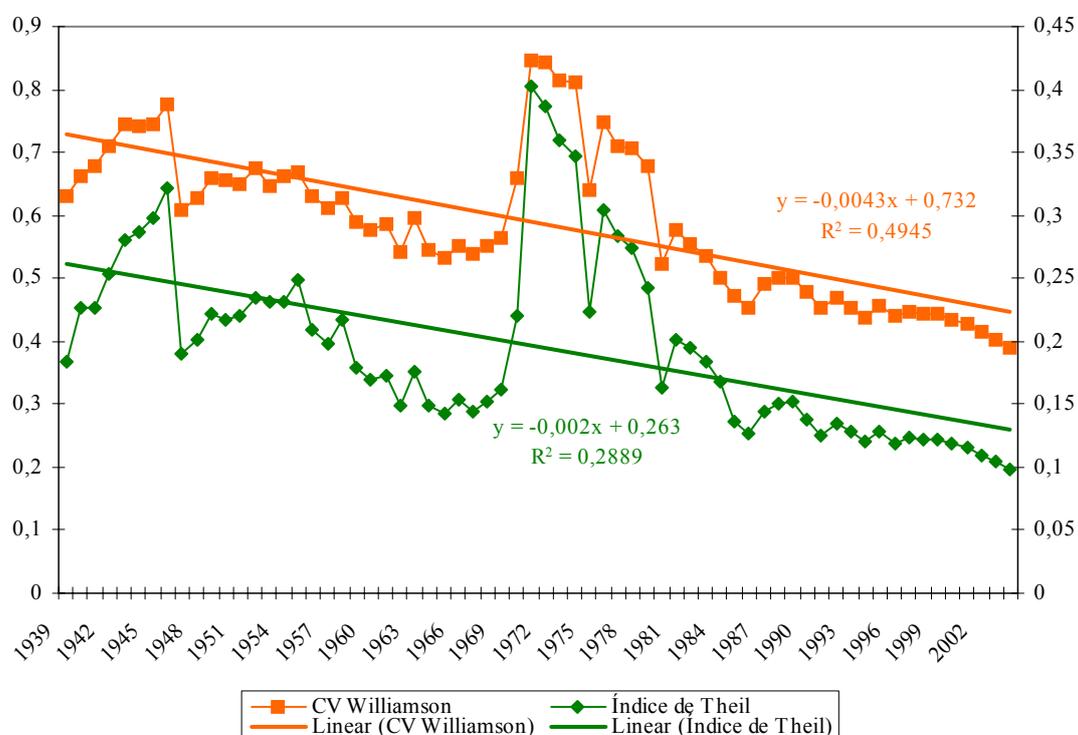


Constata-se uma redução do coeficiente de variação ponderado de Williamson de 51,88% e uma queda de 71,73% do índice de Theil para o período pós milagre. As inclinações negativas das linhas de tendência dos coeficientes de variação reforçam a idéia de queda do indicador no mencionado período.

Por fim, o Gráfico 1.4 revela a evolução dos indicadores para todo o período da amostra de 1939 a 2004.

Gráfico 1.4

Evolução do Coeficiente de Variação Ponderado de Williamson e Índice de Theil para Toda a Amostra – 1939 a 2004



De 1939 a 2004, o coeficiente de variação ponderado de Williamson apresentou uma baixa de 38,11% e o índice de Theil, uma redução de 46,66%. Apesar de ocorrer elevações dos indicadores em alguns períodos (especialmente no período do milagre), pode-se dizer que de um modo geral houve uma queda na dispersão dos PIBs *per capita* estaduais para todo período amostral. As retas de tendência cujas inclinações são negativas apontam igualmente para a diminuição dos referidos indicadores.

1.4. Conclusão

A abordagem neoclássica tradicional de convergência, que se baseia na metodologia da subseção 1.2.1, permitiu verificar β -convergência apenas para o período pós-milagre econômico. No entanto, se houver variação no intervalo de tempo em torno de uma vizinhança, tal como exposto na subseção 1.2.2, e considerar um nível de significância de 5%, constata-se β -convergência não só para os períodos pré e pós milagre econômico, mas também para todo o período amostral de 1939 a 2004.

O processo de σ -convergência na qual ocorre a redução da dispersão dos PIBs *per capita* estaduais foi estudado de acordo com os indicadores coeficiente de variação ponderado de Williamson e índice de Theil. Foi verificada a redução gradativa de ambos os indicadores nos períodos analisados, com exceção do período do milagre econômico. Portanto, pode-se concluir que se confirmou β -convergência e σ -convergência para os períodos pré e pós milagre econômico, bem como se constatou a mesma hipótese para todo o período amostral (1939 a 2004).

2. Estimativa e Testes dos Pressupostos de Markov

2.1. Introdução

A teoria das cadeias de Markov surgiu como instrumento alternativo ao modelo de regressões ilustrado por Barro e Sala-i-Martin (1991, 1992). Quah (1993b) critica o método de regressões de Barro ao mencionar a falácia de Galton. Segundo Quah, uma inclinação negativa para a reta ajustada dos dados de crescimento médio e renda inicial não significa que exista convergência ou que o grau de dispersão das rendas *per capita*s entre diversas regiões tenha diminuído. Na verdade, ocorre que, numa dada amostra, há uma tendência de ajuste para a média caso se imponha a restrição de erros estocásticos bem comportados, independentes, com média zero e distribuição normal. Ademais, ele afirma que a regressão de crescimento é compatível com qualquer padrão de comportamento da dispersão da renda, seja um padrão divergente ou constante da distribuição da renda.

A justificativa da adoção das cadeias de Markov para o estudo de convergência foi elucidada por Bernard e Durlauf (1996), *“Overall, our results suggest that neither testing framework is likely to yield unambiguous conclusions with respect to competing models of growth. At a minimum, our work shows that the bulk of cross-section evidence on convergence which has thus far appeared can be construed as consistent with some versions of the new growth theory. Additionally, the time series results accepting the no convergence null may be due to transitional dynamics in the data. An important advance over both approaches, though, would lie in the integration of the transition information in the cross-section approach with the steady state information in the time series approach to create a more general empirical methodology. One possibility is to follow the approach in Quah (1993a; 1993b) and estimate a general Markov transition function for the data and then infer the limiting distribution of the cross-section.”*

O pressuposto básico da abordagem das cadeias de Markov é que a lei de movimento do processo seja invariante no tempo e não apresente memória. A particularidade desse método é que a distribuição estacionária da renda obtida por meio de estimações de probabilidades de transição indica a direção de convergência de todo sistema.

Bickenbach e Bode (2001) argumentam, todavia, que o poder da abordagem das cadeias de Markov não foi suficientemente estudado em modelos empíricos. Ou seja, os pressupostos de invariância no tempo e a ausência de memória têm sido ignorados na aplicação desse método. O presente artigo busca seguir a metodologia de Bickenbach e Bode para investigar a validade desses pressupostos para a realidade brasileira quando se considera o período de 1939 a 2003.

Apresenta-se na seção 2.2 a teoria das cadeias de Markov bem como os testes de homogeneidade temporal, espacial e de propriedade de Markov. Os testes empíricos para o caso brasileiro são realizados na seção 2.3. E as conclusões são exibidas na seção 2.4.

2.2. Abordagem das Cadeias de Markov

2.2.1. Modelo

Uma cadeia de Markov é um processo estocástico tal que a probabilidade p_{ij} de uma variável aleatória X no estado j em qualquer ponto no tempo $t + 1$ dependa somente do estado i em que se encontra no tempo t , porém seja independente de momentos anteriores no tempo. Assim,

$$\begin{aligned} P \{X(t+1) = j \mid X(0) = i_0, \dots, X(t-1) = i_{t-1}, X(t) = i\} \\ = P\{X(t+1) = j \mid X(t) = i\} = p_{ij} \end{aligned}$$

Se o processo for constante no tempo, a cadeia de Markov é completamente determinada pela matriz de transição de Markov,

$$\Pi = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix}, \quad p_{ij} \geq 0, \quad \sum_{j=1}^N p_{ij} = 1 \quad (1)$$

cujos parâmetros sintetizam todas N^2 probabilidades de transição p_{ij} ($i, j = 1, \dots, N$), e a distribuição inicial $h_0 = (h_{10}, h_{20}, \dots, h_{N0})$, $\sum_j h_{j0} = 1$, descreve a probabilidade inicial de vários estados.

Seja X a renda relativa *per capita* regional, definida como $y_{rt} = Y_{rt} / [(1/R) \sum_r Y_{rt}]$ para a região r e período t ($r = 1, \dots, R; t = 0, \dots, T$). Divide-se toda série da renda relativa *per capita* em N classes (estados). A probabilidade de transição de Markov é definida como a probabilidade p_{ij} que uma região seja membro de uma classe de renda j em $t+1$, dado que se situava na classe i em t . A segunda linha de uma matriz de transição (1) relata a probabilidade que um membro da segunda menor classe de renda ($i=2$) vai se rebaixar para a menor classe de renda ao longo de um período de transição (p_{21}), permanecer na mesma classe (p_{22}), mover uma classe acima (p_{23}) e assim por diante. Uma vez alocada em outra classe de renda, a região vai se comportar de acordo com a distribuição de probabilidade relevante para essa classe. O vetor de probabilidade inicial h_0 descreve a distribuição regional de renda no início do primeiro período de transição em $t = 0$.

Tendo em vista a suposição que todo processo seja invariante no tempo, a matriz de transição pode reproduzir a evolução da distribuição regional da renda sobre qualquer horizonte finito ou infinito do tempo. Assim, após m períodos de transição (de t para $t + m$), a distribuição regional de renda pode ser calculada pela simples multiplicação da matriz de transição m vezes por ela mesma, isto é, $h_{t+m} = h_t \prod^m$. Ademais, se a cadeia de Markov for regular então a distribuição converge em direção à distribuição de renda estacionária h^* a qual seja independente da distribuição inicial de renda h_0 . A comparação entre a distribuição inicial de renda (h_0) e a distribuição estacionária (h^*) fornece indícios para concluir se as rendas *per capita* das regiões tendem a convergir ou divergir. Quando se observa que na distribuição estacionária apresenta maiores freqüências nas classes de renda média do que a distribuição inicial, evidencia-se um processo de convergência, e, da mesma maneira, altas freqüências nas classes mais baixas e altas da distribuição estacionária apontam para divergência.

A matriz de transição pode ser estimada pelo método da Máxima Verossimilhança. Suponha que exista apenas um período de transição, dada a distribuição inicial $h = n_i / n$, tal que n_{ij} seja o número absoluto de transições observadas de i para j . Então, a maximização

$$\log L = \sum_{i,j=1}^N n_{ij} \log p_{ij} \quad \text{sujeita a } \sum_j p_{ij} = 1, p_{ij} \geq 0 \quad (2)$$

em relação a p_{ij} fornece

$$\hat{p}_{ij} = n_{ij} / \sum_j n_{ij} \quad (3)$$

como estimador assintoticamente não viesado e normalmente distribuído de Máxima Verossimilhança de p_{ij} . Os desvios padrões dos estimadores podem ser estimados por

$$\hat{\sigma}_{\hat{p}_{ij}} = (\hat{p}_{ij}(1 - \hat{p}_{ij}) / n_i)^{1/2}.$$

Vale lembrar que o grau de confiança das probabilidades de transição depende de dois aspectos: (i) que o processo gerador dos dados seja do tipo de Markov, atendendo suas propriedades de invariância no tempo e ausência de memória, e (ii) que se disponha de número suficientemente grande de observações.

Empiricamente, as estimações das cadeias de Markov são sujeitas ao *trade-off* entre o aumento do número de observações para a obtenção de estimadores confiáveis e a crescente probabilidade de violação da propriedade de Markov. Se o período em consideração for muito longo, há o risco de se capturar quebras estruturais e fornecendo resultados da matriz de transição viesados. Com relação à área de abrangência, deve-se balancear os seguintes aspectos. Se por um lado, a fragmentação em muitas regiões possibilita o surgimento de significativas correlações entre as variáveis, por outro lado, a aglomeração de regiões omite a verificação de comportamentos particulares que sejam relevantes para análise.

2.2.2. Teste de homogeneidade no tempo

O teste de homogeneidade no tempo (estacionariedade) pode ser verificado ao dividir toda amostra em T períodos, e fazendo o teste se as matrizes de transição estimadas de cada T sub-amostra diferem significativamente da matriz estimada para toda a amostra.

Assim, testa-se $H_0: \forall t: p_{ij}(t) = p_{ij} (t = 1, \dots, T)$ contra a alternativa de as probabilidades de transição serem diferentes entre períodos: $H_a: \exists t: p_{ij}(t) \neq p_{ij}$. A estatística qui-quadrado consiste⁴:

$$Q^{(T)} = \sum_{t=1}^T \sum_{i=1}^N \sum_{j \in B_i} n_i(t) \frac{(\hat{p}_{ij}(t) - \hat{p}_{ij})^2}{\hat{p}_{ij}} \sim \chi^2 \left(\sum_{i=1}^N (a_i - 1)(b_i - 1) \right) \quad (4)$$

tal que \hat{p}_{ij} significa a probabilidade de transição da i -ésima para a j -ésima classe estimada para toda a amostra (comunhão de todos os T períodos), e $\hat{p}_{ij}(t)$ corresponde a probabilidade de transição estimada para a t -ésima sub-amostra. Visto que a suposição de que $\hat{p}_{ij}(t)$ seja mutuamente independente entre as sub-amostras sob H_0 , os N^2 parâmetros podem ser estimados de maneira similar a (3), de modo que $\hat{p}_{ij}(t) = n_{ij}(t) / n_i(t)$. Denomina-se $n_i(t)$ como o número absoluto de observações que se localiza na i -ésima classe da t -ésima sub-amostra. Vale dizer que são consideradas somente as probabilidades de transição positivas em toda a amostra, ou seja, $B_i = \{j: \hat{p}_{ij} > 0\}$, as transições nas quais não existirem observações são excluídas.

A estimativa $Q^{(T)}$ tem uma distribuição qui-quadrado assintótica cujos graus de liberdade sejam iguais aos termos da soma em $Q^{(T)}$, exceto àqueles nas quais $n_i(t) = 0$, subtraído do número de probabilidades de transição estimadas $\hat{p}_{ij}(t)$, onde ambos sejam corrigidos pelas restrições $\sum_j p_{ij}(t) = 1$ e $\sum_j p_{ij} = 1$. Os graus de liberdade podem ser calculados conforme $\sum_i a_i (b_i - 1) - (b_i - 1)$ onde b_i ($b_i = |B_i|$) é o número de entradas positivas na i -ésima linha da matriz para toda a amostra, e a_i é o número de sub-amostras (t) nas quais as observações para a i -ésima linha esteja disponível ($a_i = |A_i|$; $A_i = \{t: n_i(t) > 0\}$).

⁴ Supõe-se no teste que cada linha i da matriz de transição para toda a amostra contenha pelo menos duas probabilidades de transição não-nulas.

2.2.3. Teste de homogeneidade espacial

O teste de homogeneidade espacial implica que $H_0: \forall r: p_{ij}(r) = p_{ij}$, ($r = 1, \dots, R$) pode ser testado contra a hipótese alternativa H_a de probabilidades de transição variando entre regiões, isto é, $H_a: \exists r: p_{ij}(r) \neq p_{ij}$, na qual

$$Q^{(R)} = \sum_{r=1}^R \sum_{i=1}^N \sum_{j \in B_i} n_i(r) \frac{(\hat{p}_{ij}(r) - \hat{p}_{ij})^2}{\hat{p}_{ij}} \sim \chi^2 \left(\sum_{i=1}^N (c_i - 1)(b_i - 1) \right) \quad (5)$$

tal que $c_i = |C_i|$; $C_i = \{r: n_i(r) > 0\}$.

2.2.4. Outros testes

O teste qui-quadrado é bastante flexível no sentido que pode testar se um estado i (classe de renda) viola as suposições de homogeneidade do i -ésimo estado no tempo. Dessa forma, o teste consiste em $H_0: \forall t: p_{ji}(t) = p_{ji}$ ($t = 1, \dots, T$) contra a não-estacionariedade de $H_a: \exists t: p_{ji}(t) \neq p_{ji}$, cuja estatística é representada por

$$Q_i^{(T)} = \sum_{t \in D_i} \sum_{j \in B_i} n_i(t) \frac{(\hat{p}_{ij}(t) - \hat{p}_{ij})^2}{\hat{p}_{ij}} \sim \chi^2((d_i - 1)(b_i - 1)) \quad (6)$$

onde $d_i = |D_i|$; $D_i = \{t: n_i(t) > 0\}$, e $b_i = |B_i|$, $B_i = \{j: \hat{p}_{ij} > 0\}$.

De maneira similar, esse teste pode ser igualmente aplicado para regiões em que $H_0: \forall r: p_{ji}(r) = p_{ji}$ ($r = 1, \dots, R$) contra $H_a: \exists r: p_{ji}(r) \neq p_{ji}$, e a estatística seja

$$Q_i^{(R)} = \sum_{r \in E_i} \sum_{j \in B_i} n_i(r) \frac{(\hat{p}_{ij}(r) - \hat{p}_{ij})^2}{\hat{p}_{ij}} \sim \chi^2((e_i - 1)(b_i - 1)) \quad (7)$$

a qual $e_i = |E_i|$; $E_i = \{r: n_i(r) > 0\}$.

Note que a estatística (6) é similar a (4), da mesma maneira que (7) é semelhante a (5), sendo que a diferença de (6) e (7) é a comparação de apenas uma linha entre as matrizes de transição, ao passo que (4) e (5) utilizam todas as linhas para efeitos de análise. Conseqüentemente, as estatísticas (4) e (5) podem ser derivadas de (6) e (7) quando se faz o somatório de Q_i em todos os estados, ou seja, $Q^{(T)} = \sum_i Q_i^{(T)}$, e $Q^{(R)} = \sum_i Q_i^{(R)}$.

O teste de propriedade de Markov tem o objetivo de verificar se o processo em consideração apresenta memória. Assim, o teste consiste em verificar se as probabilidades de transição das regiões sejam independentes do estado k ($k = 1, \dots, N$) quando situadas no tempo $t - 1$. Esse teste permite avaliar se ocorre a autocorrelação serial de primeira ordem. Isto é, uma região pertencente a alguma classe de renda em t se comporte de maneira diferenciada de acordo com seu estado em $t - 1$.

Por fim, existe ainda o teste para avaliar se as probabilidades de transição são independentes no espaço. O teste para independência espacial emprega o conceito de sub-amostras tal como proposto por Rey (2001b). Uma vez definida os estados i nas quais se possam dividir a amostra em N classes de acordo com a renda da região no tempo t , Rey sugere ainda um novo conjunto de estados s para a renda média relativa dos vizinhos da região. Dessa forma, uma sub-amostra $s = 1$ representa vizinhança com renda mais baixa, e $s = S$ a vizinhança com renda mais alta.

Tanto o teste de propriedade de Markov quanto o de independência espacial são efetuados mediante as estatísticas (4) e (5) conforme os devidos ajustes da amostra. Em seguida, todos esses testes são realizados para o caso dos estados brasileiros.

2.3. Estimativa e Avaliação dos Pressupostos

Seguindo a metodologia da seção 2.2, calculou-se o PIB *per capita* relativo⁵. Dessa maneira, separou-se a amostra (1300 observações) em cinco classes de renda com iguais freqüências (quintil) para a elaboração da Tabela 2.1 a seguir.

⁵ Bickenbach e Bode (2001) sugerem a normalização pela média nacional para dirimir os efeitos de choques e tendências na série.

Tabela 2.1 – Matriz de Transição Anual Estimada para Estados Brasileiros entre os Anos de 1939 e 2003

classe de renda	limite superior	distribuição inicial		probabilidades de transição (t para $t+1$)				
		absoluta	relativa	1	2	3	4	5
1	0,2742	260	0,2	0,896	0,104	0,000	0,000	0,000
2	0,4905	260	0,2	0,062	0,812	0,127	0,000	0,000
3	0,9064	260	0,2	0,004	0,070	0,810	0,116	0,000
4	1,6761	260	0,2	0,000	0,000	0,047	0,870	0,083
5	4,5286	260	0,2	0,000	0,000	0,004	0,036	0,960
distribuição estacionária				0,195	0,200	0,199	0,202	0,203

Quando a distribuição estacionária (última linha da Tabela 2.1) apresenta frequências que sejam maiores nas classes de renda médias e menores nas classes de renda extremas, isso indica uma tendência de convergência de renda se o processo for considerado de Markov. Aparentemente, a matriz de transição não sugere convergência tendo em vista que as maiores frequências são observadas nas classes 4 e 5.

2.3.1. Teste de homogeneidade no tempo

O teste de homogeneidade no tempo consistiu na divisão de 64 períodos de transição em 12 intervalos de 5 anos cada, iniciando-se no ano de 1944. Foram então estimadas 12 matrizes de probabilidade de transição ($T=12$) a fim de compará-las simultaneamente com a matriz que engloba todo período da amostra (Tabela 1). Aplicando o teste (4), obtem-se $Q = 196,79$ e 52 graus de liberdade, que revela a diferença estatística significativa entre as matrizes de transição das sub-amostras e de toda amostra (prob. $< 0,01$).

No estudo de Bickenbach e Bode (2001), os autores encontraram quebras estruturais para o caso dos estados dos EUA para o período de 1929 a 2000, cujos testes de homogeneidade acusaram diferenças estatísticas significativas entre as matrizes de transição. Não obstante, eles obtiveram homogeneidade temporal de 5 anos quando selecionaram a amostra de 1950 a 1995.

No presente estudo, apesar de não se observar homogeneidade temporal na amostra de 1939 a 2003, verificou-se essa homogeneidade quando se atribuiu à amostra de 1953 e

2003. A matriz de transição foi reestimada para a nova amostra (Tabela 2.2) cujos testes realizados mostraram $Q = 143,81$ e 57 graus de liberdade, indicando a igualdade estatística entre as sub-amostras e toda a amostra (prob. > 0,9). Ou seja, não há quebra estrutural no sub-período 1953-2003.

Tabela 2.2 – Matriz de Transição Anual Estimada para Estados Brasileiros entre os Anos de 1953 e 2003

classe de renda	limite superior	distribuição inicial		Probabilidades de transição (t para $t+1$)				
		absoluta	relativa	1	2	3	4	5
1	0,3309	204	0,2	0,853	0,137	0,010	0,000	0,000
2	0,5355	204	0,2	0,083	0,760	0,152	0,005	0,000
3	0,9798	204	0,2	0,005	0,080	0,804	0,101	0,010
4	1,6588	204	0,2	0,000	0,005	0,035	0,869	0,090
5	∞	204	0,2	0,000	0,000	0,000	0,052	0,948
distribuição estacionária				0,192	0,200	0,200	0,204	0,204

2.3.2. Teste de homogeneidade espacial

Pode-se pensar em várias situações em que ocorre a heterogeneidade espacial, onde cada estado apresente seu processo independente de Markov, ou que o processo gerador indique aglomeração em *clusters* dos estados.

Dessa forma, procura-se avaliar se as cinco regiões brasileiras possuem características especiais quando comparadas ao país como um todo. A matriz de transição que engloba todas as regiões é a mesma da Tabela 2.2 cujas informações se repetem na Tabela 2.3 e se observam também os dados desagregados por região. Várias inferências podem ser obtidas a partir da Tabela 2.3.

Em primeiro lugar, quando se faz a comparação das matrizes de transição de cada região em relação à amostra inteira por meio da estatística (5), a hipótese nula de igualdade das matrizes é fortemente rejeitada mesmo a um nível de significância de 1% uma vez que se observa o valor $Q = 164,6$ e 115 graus de liberdade, cujo valor de probabilidade é inferior a 0,01.

Tabela 2.3 – Teste de Homogeneidade Espacial entre as
5 Regiões Brasileiras de 1953 a 2003

TODAS AS REGIÕES									
classe de renda	obs.	probabilidades de transição					teste de homogeneidade		
		1	2	3	4	5	GL	Q_i	prob.
1	204	0,853	0,137	0,010	0,000	0,000	8	24,2	0,002
2	204	0,083	0,760	0,152	0,005	0,000	27	22,1	0,734
3	199	0,005	0,080	0,804	0,101	0,010	52	46,5	0,688
4	199	0,000	0,005	0,035	0,869	0,090	24	33,7	0,091
5	194	0,000	0,000	0,000	0,052	0,948	4	38,1	0,000
						total	115	164,6	0,002

NORTE									
classe de renda	obs.	probabilidades de transição					teste de homogeneidade		
		1	2	3	4	5	GL	Q_i	prob.
1	23	0,696	0,217	0,087	0,000	0,000	2	2,3	0,316
2	28	0,143	0,750	0,071	0,036	0,000	6	3,3	0,772
3	17	0,059	0,059	0,588	0,235	0,059	12	14,3	0,284
4	20	0,000	0,050	0,150	0,650	0,150	6	17,0	0,009
5	12	0,000	0,000	0,000	0,250	0,750	1	14,5	0,000
						total	27	51,3	0,003

NORDESTE									
classe de renda	obs.	probabilidades de transição					teste de homogeneidade		
		1	2	3	4	5	GL	Q_i	prob.
1	172	0,878	0,122	0,000	0,000	0,000	2	24,2	0,000
2	99	0,121	0,737	0,141	0,000	0,000	6	12,7	0,048
3	113	0,000	0,044	0,894	0,053	0,009	12	23,0	0,027
4	65	0,000	0,000	0,046	0,954	0,000	3	11,0	0,012
5	1	0,000	0,000	0,000	1,000	0,000	0	20,4	-
						total	23	91,4	0,000

CENTRO-OESTE									
classe de renda	obs.	probabilidades de transição					teste de homogeneidade		
		1	2	3	4	5	GL	Q_i	prob.
1	8	0,875	0,125	0,000	0,000	0,000	2	0,1	0,953
2	22	0,000	0,682	0,318	0,000	0,000	3	8,2	0,042
3	14	0,000	0,357	0,429	0,214	0,000	8	26,1	0,001
4	40	0,000	0,000	0,025	0,850	0,125	6	1,4	0,968
5	16	0,000	0,000	0,000	0,188	0,813	1	9,3	0,002
						total	20	45,1	0,001

SUL									
classe de renda	obs.	probabilidades de transição					teste de homogeneidade		
		1	2	3	4	5	GL	Q_i	prob.
1	0	-	-	-	-	-	-	-	-
2	21	0,000	0,762	0,238	0,000	0,000	3	3,4	0,330
3	38	0,000	0,105	0,816	0,079	0,000	8	1,5	0,993
4	21	0,000	0,000	0,000	0,762	0,238	3	9,1	0,028

5	70	0,000	0,000	0,000	0,029	0,971	1	1,5	0,221
						total	15	15,5	0,416
SUDESTE									
classe de renda	obs.	probabilidades de transição					teste de homogeneidade		
		1	2	3	4	5	GL	Q_i	prob.
1	1	0,000	1,000	0,000	0,000	0,000	0	6,5	-
2	34	0,029	0,882	0,088	0,000	0,000	6	3,9	0,688
3	17	0,000	0,059	0,706	0,235	0,000	8	4,8	0,775
4	53	0,000	0,000	0,000	0,906	0,094	3	3,1	0,381
5	95	0,000	0,000	0,000	0,011	0,989	1	7,4	0,006
						total	18	25,8	0,105

Segundo, testa-se a homogeneidade espacial de estado (ou classe) mediante a estatística (7), ao se observar o comportamento por dentro de uma classe de renda. Os resultados podem ser encontrados no canto superior direito da Tabela 2.3, denominado de teste de homogeneidade. Verifica-se que apenas nas classes de renda mais alta e baixa, 1 e 5, a evolução da renda nas regiões apresenta comportamento diferenciado nessas faixas de renda dado um nível de significância de 1%.

Em terceiro lugar, comparou-se a matriz de transição de cada região em relação ao restante da amostra (4 regiões restantes). A estatística (7) é utilizada sob a suposição de $R = 2$ que pode ser observada na última linha de cada região na Tabela 2.2, ao lado de “total”. Nota-se que as regiões Sul e Sudeste possuem comportamento na evolução de renda semelhante ao restante das regiões (prob. > 0,1), ao passo que nas regiões Norte, Nordeste e Centro-Oeste apresentam comportamento na evolução da renda distinto às demais regiões (prob. < 0,01).

Por fim, cabe fazer análise por dentro da classe de renda de cada região. Quando se examina a classe de renda igual a 5, as regiões Norte e Sudeste apresentam comportamento nessa faixa de renda distinto às demais regiões (prob. < 0,01). Assim, observando as probabilidades de transição dessas regiões, a probabilidade de rebaixamento da região Sudeste para a classe 4 em $t+1$ dado que estava nessa classe 5 em t é inferior a 2%, por outro lado, na região Norte essa probabilidade ($p_{54|NORTE}$) é de 25%.

2.3.3. Teste de propriedade de Markov

Observa-se na Tabela 2.4 a seguir as estatísticas do teste de propriedade de Markov (ausência de memória). O teste que engloba as cinco classes de renda, calculado de acordo com a equação (4), é ilustrado no canto superior direito da Tabela 2.4 cujo resultado de $Q = 425,3$ e 73 graus de liberdade rejeita a hipótese nula de igualdade entre as matrizes (prob. $< 0,01$). Portanto, a evolução da distribuição de renda entre os estados brasileiros não pode ser considerada como independente do passado. Além disso, as avaliações de cada classe de renda quando comparadas ao restante da amostra indica igualmente a rejeição da hipótese nula de igualdade entre as matrizes (prob. $< 0,01$), sendo indicado como a última linha de cada classe de renda. Portanto, a evolução da distribuição de renda *per capita* no Brasil não pode ser considerada como independente do passado.

Tabela 2.4 – Teste de Propriedade de Markov para os Estados Brasileiros entre 1953⁶ e 2003

classe de renda		obs.	probabilidades de transição					teste propr. Markov		
<i>t-1</i>	<i>t</i>		1	2	3	4	5	GL	Q	prob.
todas	1	192	0,844	0,146	0,010	0,000	0,000	6	9,5	0,146
todas	2	201	0,085	0,751	0,159	0,005	0,000	21	28,7	0,121
todas	3	193	0,005	0,073	0,808	0,104	0,010	28	154,9	0,000
todas	4	200	0,000	0,005	0,040	0,865	0,090	15	202,2	0,000
todas	5	194	0,000	0,000	0,000	0,052	0,948	3	30,0	0,000
total								73	425,3	0,000

classe de renda		obs.	probabilidades de transição					teste propr. Markov		
<i>t-1</i>	<i>t</i>		1	2	3	4	5	GL	Q	prob.
1	1	174	0,862	0,126	0,011	0,000	0,000	2	32,3	0,000
1	2	28	0,250	0,679	0,071	0,000	0,000	6	20,0	0,003
1	3	2	0,500	0,500	0,000	0,000	0,000	3	7,3	0,062
1	4	0	-	-	-	-	-	-	-	-
1	5	0	-	-	-	-	-	-	-	-
total								11	59,7	0,000

⁶ Houve redução do número de observações se comparado à Tabela 2.2 uma vez que o ano de 1953 foi considerado apenas como referência de tempo (*t-1*).

classe de renda			probabilidades de transição					teste propr. Markov		
$t-1$	t	obs.	1	2	3	4	5	GL	Q	Prob
2	1	17	0,706	0,294	0,000	0,000	0,000	2	4,1	0,131
2	2	155	0,065	0,794	0,135	0,006	0,000	6	23,5	0,001
2	3	30	0,000	0,367	0,600	0,033	0,000	8	202,2	0,000
2	4	1	0,000	1,000	0,000	0,000	0,000	-	1,0	-
2	5	-	-	-	-	-	-	-	-	-
total								16	230,7	0,000

classe de renda			probabilidades de transição					teste propr. Markov		
$t-1$	t	obs.	1	2	3	4	5	GL	Q	prob.
3	1	1	0,000	1,000	0,000	0,000	0,000	0	6,1	-
3	2	16	0,000	0,500	0,500	0,000	0,000	3	20,0	0,000
3	3	156	0,000	0,013	0,865	0,109	0,013	9	76,8	0,000
3	4	20	0,000	0,000	0,050	0,900	0,050	6	0,6	0,996
3	5	2	0,000	0,000	0,000	0,500	0,500	1	9,2	0,002
total								19	112,7	0,000

classe de renda			probabilidades de transição					teste propr. Markov		
$t-1$	t	obs.	1	2	3	4	5	GL	Q	prob.
4	1	0	-	-	-	-	-	-	-	-
4	2	2	0,000	0,500	0,500	0,000	0,000	3	1,9	0,599
4	3	5	0,000	0,000	0,600	0,400	0,000	4	5,6	0,234
4	4	169	0,000	0,000	0,041	0,870	0,089	6	6,2	0,401
4	5	18	0,000	0,000	0,000	0,278	0,722	1	40,6	0,000
total								14	54,2	0,000

classe de renda			probabilidades de transição					teste propr. Markov		
$t-1$	t	obs.	1	2	3	4	5	GL	Q	prob.
5	1	0	-	-	-	-	-	-	-	-
5	2	0	-	-	-	-	-	-	-	-
5	3	0	-	-	-	-	-	-	-	-
5	4	10	0,000	0,000	0,000	0,800	0,200	3	2,1	0,548
5	5	174	0,000	0,000	0,000	0,023	0,977	1	63,6	0,000
total								4	65,7	0,000

Algumas observações adicionais sobre a Tabela 2.4 são descritas a seguir:

- (i) os estados que se situam em $t-1$ na classe 1 e permanecem na mesma classe em t , passam a ter apenas 12,6% de probabilidade de subirem para a classe 2 em $t+1$. Por outro lado, se um estado estiver em $t-1$ na classe 1, subir para a classe 2 em t , esse estado terá 25% de probabilidade de voltar para a classe 1 em $t+1$. Tal circunstância indica a tendência de continuidade da pobreza dos estados de menor renda.
- (ii) na classe de renda $t-1$ igual a 2, os estados que caírem para a classe 1 em t , passam a ter a probabilidade de 70,6% de continuar na classe de renda mais baixa em $t+1$.

- (iii) na classe de renda $t-1$ igual a 4, os estados que subirem para classe 5 em t , passaram a ter a probabilidade de 72,2% de permanecerem na classe de renda mais alta em $t+1$.
- (iv) os estados que estavam em $t-1$ na classe 5 e continuaram nessa classe em t , passam a ter a probabilidade de 97,7% de permanecerem nas mesma classe em $t+1$. Visualiza-se, portanto, o quadro de continuidade da riqueza de alguns estados diante da alta probabilidade de $\hat{p}_{5|5}$.

A conclusão desta subseção é a importância do passado na evolução da distribuição da renda *per capita* dos estados brasileiros. Dessa maneira, o pressuposto básico de ausência de memória utilizado nas abordagens de Markov é severamente violado para a amostra de 1953 a 2003.

2.3.4. Teste de independência espacial

Apesar de a abordagem de Markov pressupor observações estocasticamente independentes, a dinâmica da renda pode ser afetada por efeitos de transbordamentos (*spillovers effects*) de regiões vizinhas. Os artigos de Rey e Montouri (1999), Rey (2001a) e Rey (2001b) mostram a forte dependência de vizinhança entre os estados dos EUA.

O teste de independência espacial foi feito mediante a metodologia da seção 2.2.4. Os estados foram classificados de acordo com 5 classes de renda para o período de 1953 a 2003 e categorizados conforme a renda média de seus vizinhos. Assim, uma observação alocada em $s = 1$ ocorre se a renda média *per capita* dos estados vizinhos se situa no primeiro quintil da renda mais baixas da amostra. O teste consiste, portanto, se as probabilidades de transição encontradas sejam iguais à matriz da Tabela 2.2.

Quando se considerou toda a amostra, o resultado do teste de dependência espacial foi o $Q=89,1$ e 118 graus de liberdade, cujo valor de probabilidade de erro corresponde a 0,978. Dessa maneira, aceita-se a hipótese nula de independência espacial e conclui-se que a renda média dos estados vizinhos não é fundamental para explicar a evolução da renda de cada estado.

Verificou-se que o efeito de transbordamento ocorre somente nos estados cujas vizinhanças possuem a renda média $s = 3$, quando se considera um nível de significância igual ou inferior a 5%. Em todas as demais rendas da vizinhança, os testes mostraram a independência espacial dos estados brasileiros. Vale dizer que os estados cujas vizinhanças possuem a renda média $s=3$ são Alagoas, Amazonas, Bahia e Pará.

Uma observação importante do significado das probabilidades da matriz de transição das sub-amostras pode ser descrito da seguinte maneira. Um estado com renda *per capita* alta (classe de renda 5) cujos vizinhos sejam também ricos (classe de renda 5) tem a probabilidade de reduzir para a classe de renda 4 mais baixa se comparada ao caso de toda amostra, em que não existe a restrição de se fixar o vizinho rico ($\hat{p}_{54|5} = 0,020 < \hat{p}_{54} = 0,052$).

Tabela 2.5 – Teste de Dependência Espacial entre as Fronteiras dos Estados

TODA AMOSTRA									
classe de renda	obs.	probabilidades de transição					teste de homogeneidade		
		1	2	3	4	5	GL	Q	prob.
1	204	0,853	0,137	0,010	0,000	0,000	8	13,0	0,112
2	204	0,083	0,760	0,152	0,005	0,000	30	14,5	0,992
3	199	0,005	0,080	0,804	0,101	0,010	52	20,4	1,000
4	199	0,000	0,005	0,035	0,869	0,090	24	19,0	0,750
5	194	0,000	0,000	0,000	0,052	0,948	4	22,1	0,000
						Total	118	89,1	0,978

VIZINHANÇA COM A MENOR RENDA (s=1)									
classe de renda	obs.	probabilidades de transição					teste de homogeneidade		
		1	2	3	4	5	GL	Q	prob.
1	65	0,846	0,154	0,000	0,000	0,000	2	1,2	0,538
2	46	0,130	0,696	0,174	0,000	0,000	6	3,4	0,758
3	66	0,000	0,061	0,894	0,045	0,000	8	7,2	0,515
4	23	0,000	0,000	0,043	0,957	0,000	3	2,8	0,423
5	0	-	-	-	-	-	-	-	-
						total	19	14,6	0,745

VIZINHANÇA COM A RENDA ABAIXO DA MÉDIA(s=2)

classe de renda	obs.	probabilidades de transição					teste de homogeneidade			
		1	2	3	4	5	GL	Q	prob.	
1	81	0,901	0,099	0,000	0,000	0,000	2	4,0	0,137	
2	49	0,082	0,816	0,102	0,000	0,000	6	1,9	0,925	
3	23	0,000	0,043	0,826	0,087	0,043	12	6,5	0,891	
4	35	0,000	0,000	0,029	0,886	0,086	6	0,3	0,999	
5	12	0,000	0,000	0,000	0,250	0,750	1	14,5	0,000	
							total	27	27,2	0,453

VIZINHANÇA COM A RENDA MÉDIA (s=3)

classe de renda	obs.	probabilidades de transição					teste de homogeneidade			
		1	2	3	4	5	GL	Q	prob.	
1	57	0,807	0,158	0,035	0,000	0,000	2	0,6	0,731	
2	40	0,150	0,700	0,125	0,025	0,000	6	4,7	0,587	
3	46	0,022	0,043	0,761	0,152	0,022	12	5,4	0,943	
4	45	0,000	0,022	0,111	0,800	0,067	6	34,2	0,000	
5	12	0,000	0,000	0,000	0,250	0,750	1	14,5	0,000	
							total	27	59,4	0,000

VIZINHANÇA COM A RENDA ACIMA DA MÉDIA(s=4)

classe de renda	obs.	probabilidades de transição					Teste de homogeneidade			
		1	2	3	4	5	GL	Q	prob.	
1	1	0,000	1,000	0,000	0,000	0,000	0	6,5	-	
2	48	0,021	0,813	0,167	0,000	0,000	6	3,9	0,696	
3	24	0,000	0,208	0,625	0,167	0,000	8	11,2	0,192	
4	57	0,000	0,000	0,000	0,895	0,105	3	3,6	0,310	
5	70	0,000	0,000	0,000	0,029	0,971	1	1,5	0,221	
							total	18	26,6	0,086

VIZINHANÇA COM A MAIOR RENDA (s=5)

classe de renda	obs.	probabilidades de transição					Teste de homogeneidade			
		1	2	3	4	5	GL	Q	prob.	
1	0	-	-	-	-	-	-	-	-	
2	21	0,000	0,762	0,238	0,000	0,000	3	3,4	0,330	
3	40	0,000	0,100	0,800	0,100	0,000	8	1,1	0,998	
4	39	0,000	0,000	0,000	0,846	0,154	3	5,2	0,156	
5	100	0,000	0,000	0,000	0,020	0,980	1	5,4	0,020	
							total	15	15,2	0,439

2.3.5. Teste de convergência nas regiões

A Tabela 2.6 ilustra a distribuição estacionária da renda das regiões brasileiras de maneira a permitir uma avaliação preliminar do processo de convergência da renda no Brasil caso os pressupostos da abordagem de Markov sejam atendidos. Tendo como ponto de partida uma distribuição de 0,2 em cada classe de renda na distribuição do Brasil, observa-se na Tabela 2.6 (última linha) que houve uma ligeira alteração na

distribuição de renda no incremento das classes de renda mais elevadas, 4 e 5, no período de 1953 a 2003.

Tabela 2.6 – Distribuição Estacionária da Renda das Regiões Brasileiras de 1953 a 2003

	classe de renda				
	1	2	3	4	5
Norte	0,210	0,280	0,170	0,210	0,130
Nordeste	0,362	0,220	0,262	0,153	0,002
Centro-Oeste	0,070	0,210	0,140	0,400	0,180
Sul	0,000	0,133	0,240	0,140	0,487
Sudeste	0,005	0,160	0,075	0,265	0,495
BRASIL	0,192	0,200	0,200	0,204	0,204

No entanto, os dados da tabela sugerem que algumas regiões ficaram estagnadas em classes de renda mais baixa ao passo que outras permaneceram em estratos mais elevados da renda. As regiões Norte e Nordeste possuem participações majoritariamente em classes de renda mais baixas e ocorrendo o inverso para as regiões Sul e Sudeste. Por sua vez, a região Centro-Oeste se aproxima das regiões Sul e Sudeste ao possuir uma participação de apenas 7% na classe de renda mais baixa, e de 40% na classe de renda média alta (igual a 4). Portanto, os dados compreendidos entre os anos de 1953 a 2003 indicam a divergência das rendas *per capita* regionais quando os pressupostos de Markov são atendidos.

2.4. Conclusão

Este artigo seguiu a metodologia de Bickenbach e Bode (2001) para testar os pressupostos da abordagem das cadeias de Markov, bem como analisar alguns resultados das inferências estatísticas.

Inicialmente, verificou-se mediante o teste de homogeneidade no tempo com intervalos de cinco anos uma quebra estrutural na amostra de 1939 a 2003. Não obstante, obteve-se a homogeneidade no tempo quando foi considerado o período de amostra compreendido entre os anos de 1953 a 2003.

Segundo, foi realizado o teste de homogeneidade espacial aglomerando os estados de acordo com suas respectivas regiões. Observou-se a rejeição da hipótese de que a matriz de probabilidades de transição brasileira é significativa para representar as regiões. Ou seja, cada região possui sua própria distribuição de evolução da renda *per capita*.

Terceiro, o teste de propriedade de Markov foi de suma importância para a avaliação dos pressupostos de Markov. Foi constatada a importância do passado sobre a evolução da distribuição de renda uma vez que os testes em todas as classes de renda rejeitaram fortemente (prob. < 0,01) a hipótese de ausência de memória. Além disso, as probabilidades de transição indicaram a tendência de um estado pobre permanecer na classe de renda mais baixa e de um estado rico prosseguir na classe de renda mais alta.

Em quarto, foram avaliados os efeitos de transbordamentos (*spillovers effects*) da renda de regiões vizinhas sobre a renda de cada estado. Surpreendentemente, os testes indicaram para a independência espacial dos estados brasileiros na explicação da evolução da renda *per capita*. Isto é, a renda média dos estados vizinhos não é importante para explicar a evolução da renda de cada estado.

Por último, a distribuição estacionária da renda das regiões brasileiras exibiu um quadro em que a evolução da renda *per capita* não caminha para um processo de convergência quando se emprega a metodologia de Markov. As regiões Norte e Nordeste se situam principalmente nas classes de renda mais baixa, ao passo que as regiões Centro-Oeste, Sul e Sudeste, nas classes de renda mais alta.

De acordo com os resultados encontrados, nota-se que vários pressupostos para a utilização da abordagem de Markov foram infringidos, tal como a autoregressividade da renda no tempo e a particularidade de cada região brasileira (heterogeneidade espacial).

Se, por um lado, os testes realizados condenam os pressupostos para a adoção do processo de Markov, por outro lado, esclarecem alguns aspectos sobre a distribuição da evolução dos PIBs *per capita* estaduais, tal como comportamento regional da série assim como a presença de autocorrelação serial (efeito memória).

3. Decomposição da Dinâmica da Distribuição Estadual Usando Gini

3.1. Introdução

Neste capítulo, aplica-se o método de O'Neill e Van Kerm (2004) que propõem uma técnica de decomposição da dinâmica distributiva utilizando Gini. Os referidos autores aproveitam o instrumental desenvolvido para tributação e redistribuição de renda, em especial Kakwani (1977), da área de economia do setor público.

Examina-se σ -convergência como a variação do coeficiente de Gini no tempo. Ademais, realiza-se uma decomposição do coeficiente de Gini, proposta por Jenkins e Van Kerm (2003), na qual se separa de maneira aditiva os efeitos de β -convergência e de ultrapassagem. Ou seja, a vantagem desta metodologia em relação ao índice de Theil e o coeficiente de variação ponderado de Williamson, expostos no capítulo 1, é a percepção do efeito ultrapassagem ocorrida na evolução temporal da distribuição.

A próxima seção será destinada a ilustrar as diversas possibilidades que surgem da decomposição e mostrar alguns gráficos dos efeitos β -convergência e ultrapassagem dos estados brasileiros. A metodologia será apresentada na terceira seção. Na última seção, são exibidos os resultados encontrados para o caso dos PIBs *per capita* dos estados brasileiros referente aos anos de 1939 a 2004, e fazendo-se um destaque para o período do milagre econômico.

3.2. Ilustrações dos Efeitos β -Convergência, σ -Convergência e Ultrapassagem

Inicialmente, são apresentadas algumas figuras com o intuito de mostrar as diversas possibilidades da dinâmica distributiva. As Figuras 3.1 e 3.2 a seguir exibem as situações onde existem β -convergência e σ -convergência, com a diferença que o efeito ultrapassagem (ou reordenamento) só ocorre na segunda figura.

Figura 3.1
 β -convergência, σ -convergência e sem ultrapassagem

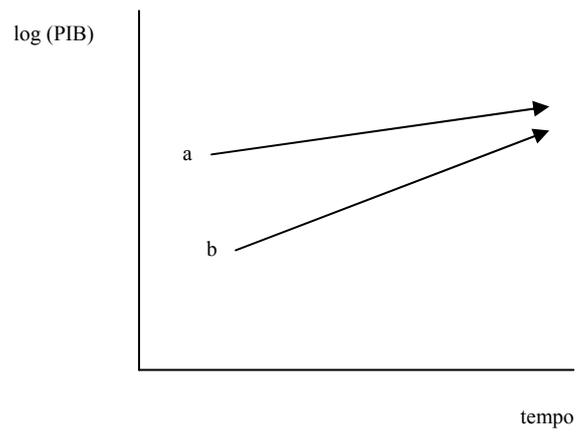
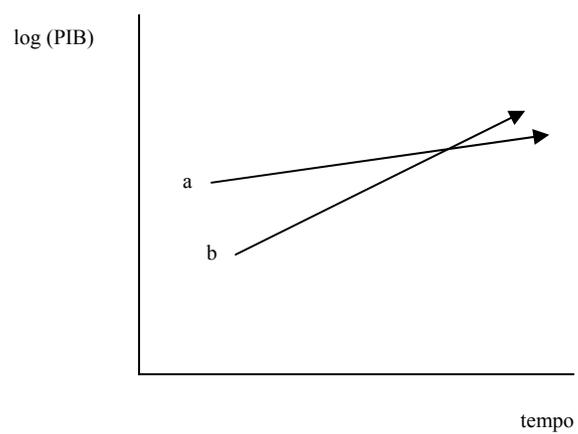


Figura 3.2
 β -convergência, σ -convergência e com ultrapassagem



Nas Figuras 3.3 e 3.4 que se seguem, são exibidas situações em que não há σ -convergência. A distinção entre elas é que na Figura 3.3 se verifica tanto o efeito β -convergência quanto o efeito ultrapassagem, ao passo que na Figura 3.4 nenhum desses efeitos acontece.

Figura 3.3
 β -convergência, não há σ -convergência e com ultrapassagem

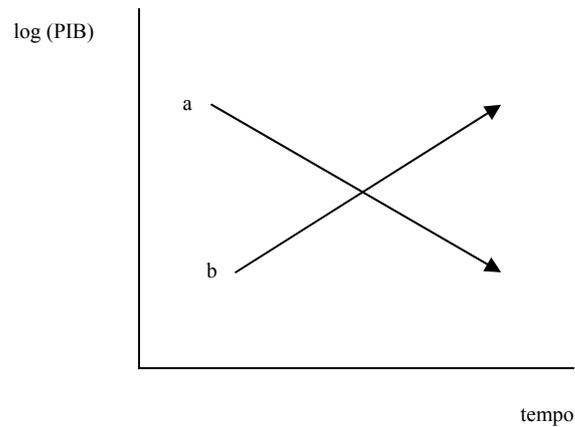
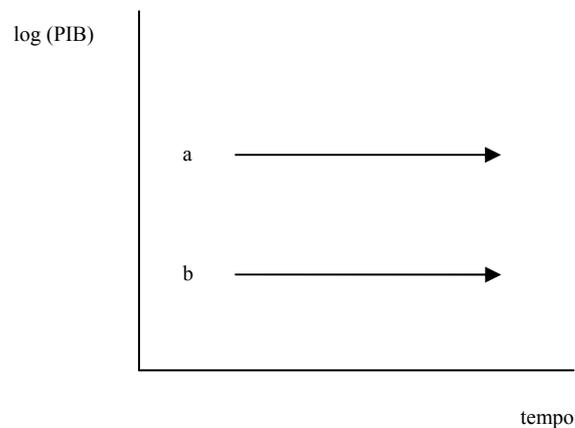


Figura 3.4
Não há β -convergência, não há σ -convergência e sem ultrapassagem

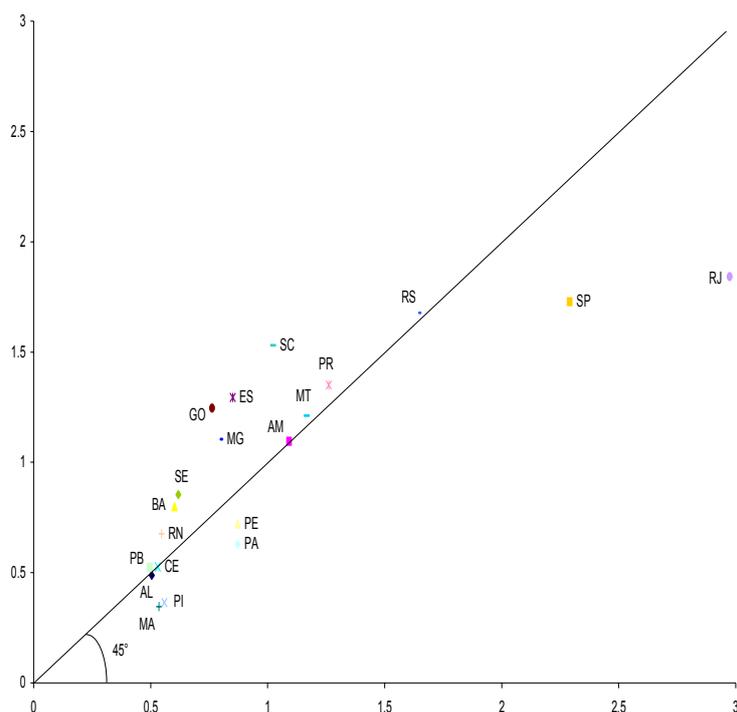


Vale dizer que a decomposição do índice de Gini, apresentada na seção 3.3, é capaz de distinguir qualquer hipótese representada nas figuras. Em seguida, são expostos dois gráficos que mostram a possibilidade de β -convergência e o efeito ultrapassagem para o caso dos PIBs *per capita* dos estados brasileiros.

O Gráfico 3.1 mostra as participações de cada estado em relação à renda média para os anos de 1939 e 2004. No eixo das abscissas, divide-se o PIB *per capita* do estado *i* do ano de 1939 pelo PIB *per capita* médio do mesmo ano. Ocorre a mesma divisão para o eixo das ordenadas, mas referente ao ano de 2004.

Gráfico 3.1

Participação de cada Estado em Relação à Renda Média – 1939 e 2004

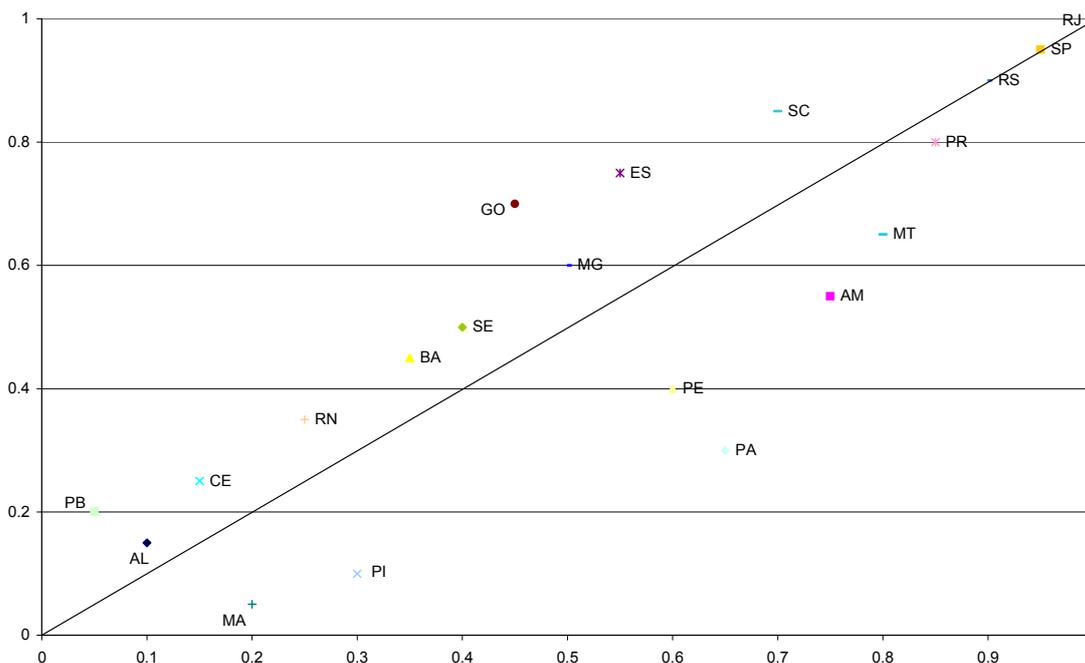


Os estados mais ricos são representados por valores maiores que 1, e os mais pobres, inferiores a 1. A linha de 45° reproduz a situação de crescimento proporcional de renda. A importância do Gráfico 3.1 é expor o efeito β -convergência, pois mostra a evolução proporcional da renda de cada estado. Assim, os estados acima da linha de 45° apresentam crescimento mais do que proporcional que a renda média, que se destacam

os estados de Santa Catarina, Goiás e Espírito Santo, e ocorre o contrário para os estados situados abaixo da linha de 45°, que se distinguem os estados do Rio de Janeiro e São Paulo.

O Gráfico 3.2 exibe o efeito de reordenamento, ou ultrapassagem, entre os anos de 1939 e 2004. Foi construído um índice de classificação de estados mais ricos (estes possuem o valor mais próximo de 1) para os mais pobres (mais próximo de 0) para os anos de 1939, representado pelo eixo das abscissas, e 2004, eixo das ordenadas. Desse modo, os estados que se situam acima da linha de 45° elevaram seu posto em relação aos demais no período mencionado, e vice versa para os países abaixo da linha.

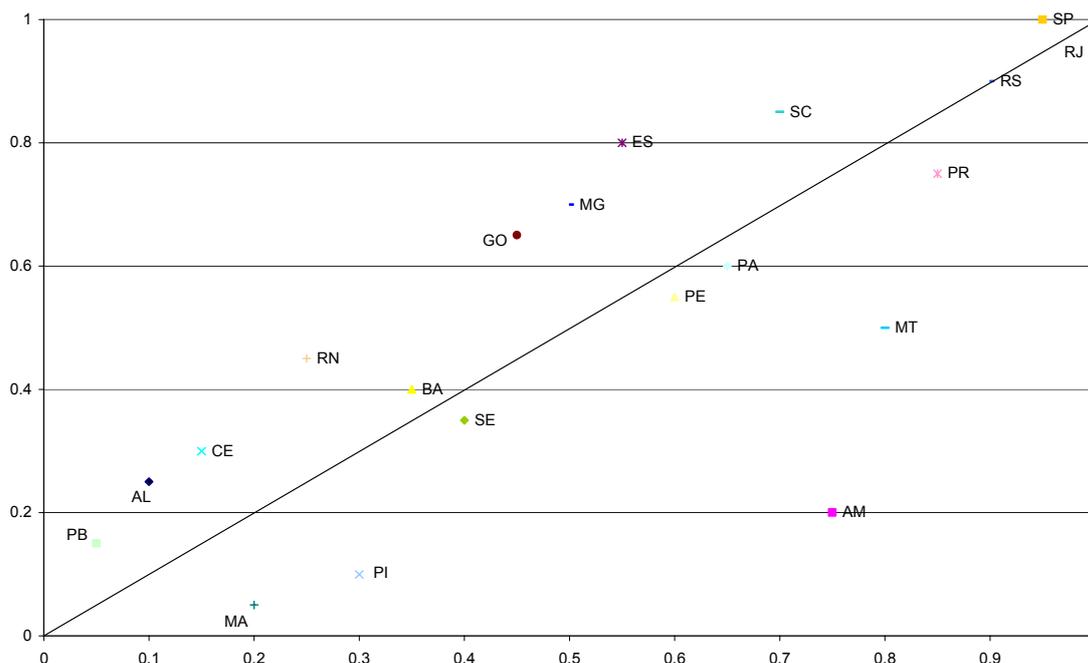
Gráfico 3.2
Efeito Ultrapassagem – 1939 e 2004



Três estados não variaram seu posto, Rio de Janeiro, São Paulo e Rio Grande do Sul, e seguem de 1939 a 2004 como os três estados mais ricos. Goiás ganhou 5 posições, ao passo que Pará perdeu 7 posições no período.

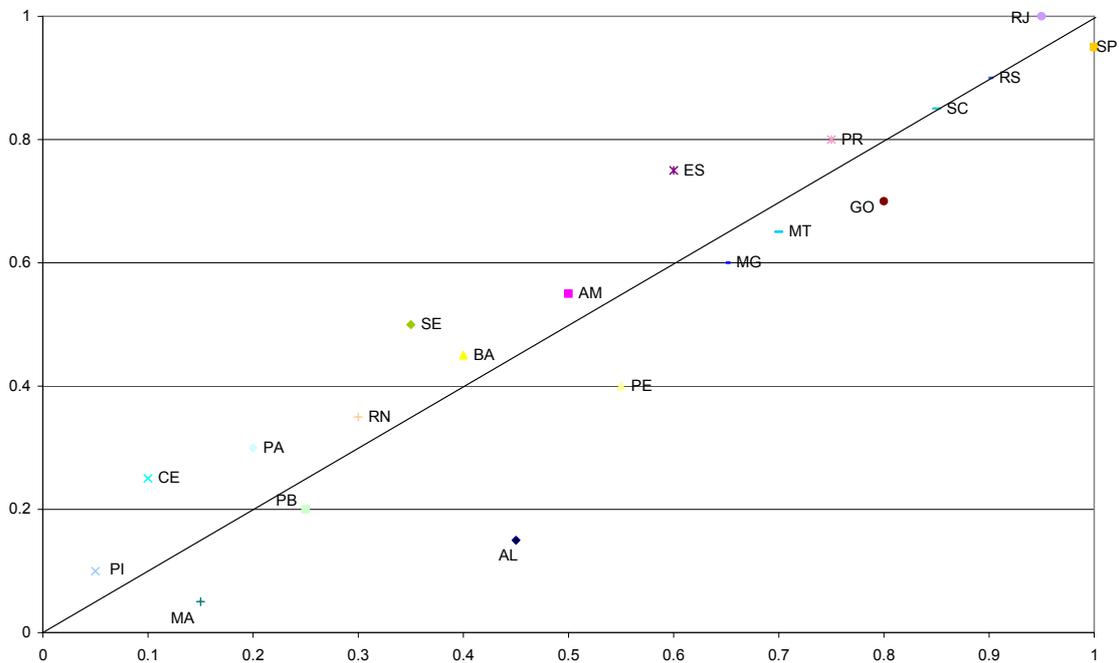
Os Gráficos 3.3 e 3.4 a seguir mostram igualmente o efeito ultrapassagem para os períodos pré e pós milagre econômico.

Gráfico 3.3
Efeito Ultrapassagem – 1939 e 1968



Para o período pré milagre econômico, os destaques são as quedas de 11 posições para o estado do Amazonas e 6 posições para o Mato Grosso. O estado do Espírito Santo foi o que apresentou o melhor desempenho ao ganhar 5 posições.

Gráfico 3.4
Efeito Ultrapassagem – 1974 e 2004



No período pós milagre econômico, o estado do Alagoas perdeu 6 posições e Pernambuco 3 posições. Os estados do Ceará, Espírito Santo e Sergipe apresentaram as maiores ultrapassagens no período, sendo que cada um ganhou 3 posições.

3.3. Metodologia

A decomposição da dinâmica da distribuição sugerida por Jenkins e Van Kerm (2003) será realizada a partir do índice de Gini generalizado, ou S-Gini. O coeficiente S-Gini para um dado ano é:

$$G(v) = 1 - \int [v(1-p)^{v-1} \frac{x}{\mu} f(x) dx] \quad (1)$$

tal que v mede a aversão à desigualdade, $f(x)$ é a função densidade de probabilidade da renda x , $p = F(x)$ é a função de distribuição acumulada da renda e μ é a renda média.

O coeficiente tradicional de Gini é medido quando $\nu = 2$. Os valores de $\nu > 2$ fornecem uma ponderação social maior para indivíduos mais pobres e o contrário acontece para $\nu < 2$.

Uma maneira alternativa de se expressar a equação (1) é:

$$G(\nu) = -\frac{\nu}{\mu} \text{cov}_p(x, (1-p)^{\nu-1}) \quad (2)$$

onde cov_p significa a covariância em relação a p .

A vantagem da equação (2) é a possibilidade de se estimar o coeficiente de Gini generalizado mediante um simples processo de regressão. Schechtman e Yitzhaki (2005) encontram uma equação equivalente à (2), mas voltada para o caso de dados discretos.

$$\hat{G}_1(\nu) = -\frac{\nu}{\mu} \sum_{i=1}^M p_i (x_i - \mu) [(1 - \hat{\pi}_i)^{\nu-1} - m] \quad (3)$$

tal que M é o número de observações, $p_i = 1 / M$, $\pi_i = p_1 + p_2 + \dots + p_i$, $\hat{\pi}_i = (\pi_{i-1} + \pi_i) / 2$, e $m = \sum_{i=1}^M p_i (1 - \hat{\pi}_i)^{\nu-1}$.

A avaliação social que fundamenta o índice de Gini é a média ponderada de cada renda individual relativa (em relação à média, x / μ). A ponderação social, $w(p; \nu) = \nu(1-p)^{1-\nu} > 0$, é uma função decrescente da ordenação individual na hierarquia social. A medida de desigualdade de S-Gini é a diferença na avaliação social entre o caso hipotético de igualdade de renda e a renda relativa efetivamente observada. A amplitude de $G(\nu)$ varia de zero, em que há igualdade de renda, a um, onde ocorre a desigualdade máxima.

Considere a mudança do S -Gini entre um ano base (0) e o ano final (1) para uma população fixa de indivíduos. Seja $f(x_0, x_1)$ a função densidade de probabilidade conjunta para a renda nos anos 0 e 1, a mudança em $G(v)$ se descreve como:

$$\Delta G(v) \equiv G_1(v) - G_0(v) \quad (4)$$

$$\Delta G(v) = - \iint \left[\left(w(p_1; v) \frac{x_1}{\mu_1} \right) - \left(w(p_0; v) \frac{x_0}{\mu_0} \right) \right] f(x_0, x_1) dx_0 dx_1 \quad (5)$$

tal que os subscritos 0 e 1 identificam os anos inicial e final, p_0 e p_1 são dados pelas distribuições marginais cumulativas e $w(p_i; v)$ é o peso social.

A partir da equação (5), é possível deduzir que existem dois fatores subjacentes a uma mudança na desigualdade de renda: mudanças nas rendas individuais relativas e variações nas ponderações sociais, a qual este último depende do ordenamento recebido na distribuição de renda. Os dois tipos de mudanças podem não ser independentes tendo em vista que uma grande mudança na renda relativa geralmente está associada a um aumento no ordenamento (*ranking*) e, conseqüentemente, a uma redução da ponderação social.

Cabe notar, no entanto, que as mudanças na renda e no ordenamento não são perfeitamente correlacionadas. Pode ocorrer, por exemplo, um caso de preservação da média com difusão de rendas de tal maneira a reduzir a renda daqueles que tinham a renda relativa menor que um e aumentar a renda daqueles com renda relativa maior que um, e ainda assim não afetar o ordenamento. O ordenamento é igualmente preservado se todas as rendas variam seja de maneira uniforme, proporcional ou em termos absolutos. Além disso, mesmo se a renda individual relativa de um agente permanecer constante, sua ponderação social varia uma vez que ocorrem mudanças nas rendas relativas de outros.

A fim de decompor o índice de Gini generalizado, é necessário dar conhecimento a algumas variáveis. Define-se $C_1^0(v)$ como o coeficiente de concentração generalizado

para as rendas do ano 1 calculado a partir da ordenação do ano 0. Essama-Nssah (2002) deriva a equação do coeficiente de concentração generalizado como a seguir.

$$C_1^0(v) = -\frac{v}{\mu_{x_1}} \text{cov}[x_1, (1-q)^{v-1}] \quad (6)$$

onde q se refere à distribuição acumulada de x_0 .

O índice de mobilidade na forma de reordenamento, $R(v)$, representa-se da seguinte maneira.

$$R(v) = \iint [w(p_0; v) - w(p_1; v)] \left(\frac{x_1}{\mu_1} \right) f(x_0, x_1) dx_0 dx_1 = G_1(v) - C_1^0(v) \quad (7)$$

E a medida de progressividade no crescimento da renda se traduz na equação que se segue.

$$P(v) = \iint w(p_0; v) \left[\frac{x_1}{\mu_1} - \frac{x_0}{\mu_0} \right] f(x_0, x_1) dx_0 dx_1 = G_0(v) - C_1^0(v) \quad (8)$$

A partir das definições de $R(v)$ e $P(v)$, é possível chegar a seguinte decomposição de Gini.

$$\Delta G(v) = R(v) - P(v) \quad (9)$$

Observa-se, por meio da equação (9), que a desigualdade é reduzida por um crescimento progressivo da renda a não ser que seja compensada por uma mobilidade concomitante da renda (reordenamento ou ultrapassagem).

3.4. Análise dos Dados

A amostra de dados é composta por 20 estados e o período total de abrangência é de 1939 a 2004. Observa-se na Tabela 3.1 os resultados da decomposição da desigualdade

para os períodos pré-milagre (1939 a 1968), milagre (1969-1973), pós-milagre (1974-2004) e de toda a amostra.

Tabela 3.1
Decomposição da Desigualdade para Períodos Seleccionados

Ano Inicial	Ano Final	Grau de Aversão à Desigualdade (ν)	Gini Inicial	Gini Final	Mudança no Gini (ΔG)	Reordenamento $R(\nu)$	Progressividade $P(\nu)$
1939	1968	1,5	0.2015	0.1803	-0.0212	0.0198	0.0410
1939	1968	2,0	0.2995	0.2719	-0.0277	0.0332	0.0609
1939	1968	2,5	0.3533	0.3253	-0.0280	0.0435	0.0715
1969	1973	1,5	0.1893	0.2955	0.1062	0.0151	-0.0911
1969	1973	2,0	0.2852	0.4396	0.1544	0.0282	-0.1261
1969	1973	2,5	0.3407	0.5190	0.1783	0.0381	-0.1402
1974	2004	1,5	0.2873	0.1600	-0.1273	0.0071	0.1345
1974	2004	2,0	0.4275	0.2629	-0.1645	0.0128	0.1773
1974	2004	2,5	0.5054	0.3332	-0.1722	0.0178	0.1900
1939	2004	1,5	0.2015	0.1600	-0.0415	0.0187	0.0602
1939	2004	2,0	0.2995	0.2629	-0.0366	0.0342	0.0708
1939	2004	2,5	0.3533	0.3332	-0.0201	0.0455	0.0656

As variações no coeficiente de Gini (ΔG) refletem mudanças na dispersão de renda sobre o tempo, e pode ser considerado como medida de σ -convergência. O valor negativo de ΔG corresponde a uma redução da desigualdade e seu valor positivo, um aumento na desigualdade. Conforme se nota na tabela, verificou-se uma redução da desigualdade em todos os períodos estudados exceto para o período do milagre econômico. Ou seja, houve um aumento na desigualdade de renda entre os estados brasileiros apenas no período do milagre econômico.

As variações negativas mais significativas do coeficiente de Gini ocorreram no período pós-milagre econômico (1974-2004), em que as variações desse coeficiente foram superiores a 30%. Por outro lado, as variações positivas ocorridas no período do milagre foram de alta magnitude, chegando a porcentagens superiores a 50% de incremento, revelando uma alta concentração de renda nesse período.

Cabe notar que o grau de aversão à desigualdade (ν), representado pela terceira coluna da tabela, não alterou o resultado das análises. É válido lembrar que $\nu = 2$ corresponde ao coeficiente de Gini regular, $\nu < 2$ é quando se atribui peso relativamente maior aos mais ricos e $\nu > 2$, aos mais pobres.

O termo de progressividade, $P(\nu)$, captura o grau de redução da desigualdade de renda no tempo como resultado de taxas de crescimento maiores entre os estados mais pobres. Age como uma medida distributiva de crescimento de renda pró-pobre. Essa definição se equivale ao conceito de β -convergência. Nota-se na tabela que o crescimento regressivo da renda, pró-rico, só ocorreu no período do milagre econômico.

O componente de reordenamento, $R(\nu)$, retrata o efeito ultrapassagem ou reordenamento na mobilidade de renda. O papel do reordenamento foi diminuto na dinâmica da distribuição dos estados brasileiros visto que seu valor absoluto foi sempre inferior ao efeito de crescimento progressivo da renda. Assim, $R(\nu)$ foi insuficiente para compensar a redução na desigualdade de renda ocasionada por β -convergência, com exceção do período do milagre onde houve aumento na desigualdade cujo fator preponderante foi o crescimento regressivo da renda ($P(\nu) < 0$).

3.5. Conclusão

A vantagem da decomposição sugerida por Jenkins e Van Kerm (2003) é se atentar para o efeito ultrapassagem, que possa levar a uma possível alteração no ordenamento (*ranking*) das rendas *per capita* dos estados. A metodologia desenvolvida no Capítulo 1 não capta o efeito ultrapassagem, o que impede fazer a distinção entre as situações dispostas nas Figuras 3.1 e 3.2.

Verificou-se no presente estudo uma redução da desigualdade em todos os períodos avaliados exceto para o período do milagre econômico. Assim, houve um aumento na desigualdade de renda entre os estados brasileiros apenas no período do milagre econômico. O efeito de reordenamento ou ultrapassagem apresentou um papel diminuto nas análises e não foi suficiente para compensar o efeito seja progressivo ou regressivo do crescimento da renda dos estados brasileiros.

4. Aferição de Convergência com Transições e Clubes

4.1. Introdução

A idéia deste capítulo é introduzir e aplicar a metodologia de dados de painel dinâmico de Phillips e Sul (2007) para avaliar a convergência dos PIBs *per capita* dos estados brasileiros. O método dos referidos autores se propõe a representar o comportamento de economias em transição, tal que se permita um amplo conjunto de caminhos temporais de transição e suponha heterogeneidade de cada economia.

O modelo possui componentes tanto específicos quanto comuns e se baseia na evolução dinâmica não linear do fator. A aplicação em micro dados de painel permite uma decomposição na qual o comportamento idiossincrático varie tanto no tempo como entre sessões, assim como mantendo perfil ordinário do painel por meio de um componente de crescimento desconhecido.

O aspecto ordinário do painel significa que quando os componentes idiossincráticos de variação heterogênea no tempo convergem para uma constante, há uma definição de convergência similar ao conceito de σ -convergência.

A seção seguinte busca explicar o modelo desenvolvido por Phillips e Sul (2007) para averiguar tanto a convergência absoluta quanto em clubes. Não foram apresentadas as propriedades assintóticas dos estimadores uma vez que as demonstrações podem ser encontradas nos artigos dos referidos autores. A aplicação do modelo ao caso dos estados brasileiros é ilustrada na seção 4.3, e as conclusões exibidas na seção 4.4.

4.2. Metodologia Econométrica

4.2.1. Representação do fator variável no tempo

Define-se um arcabouço de dados de painel para a variável X_{it} , onde $i = 1, \dots, N$ e $t = 1, \dots, T$, tal que N represente o número de países (ou regiões) e T a extensão temporal da amostra. Assume-se que a variável X_{it} possa ser decomposta em um componente

sistemático, g_{it} , a qual incorpora os componentes comuns que resultam na dependência do corte transversal, e outro transitório, a_{it} , tal como expressa na equação a seguir.

$$X_{it} = g_{it} + a_{it} \quad (1)$$

Phillips e Sul (2007) fazem uma transformação não linear na equação (1) de maneira que os componentes comuns, μ_t , e idiossincráticos, δ_{it} , do painel sejam isolados.

$$X_{it} = \left(\frac{g_{it} + a_{it}}{\mu_t} \right) \mu_t = \delta_{it} \mu_t, \text{ para qualquer } i \text{ e } t \quad (2)$$

Se, por exemplo, μ_t representar um componente comum de tendência no painel, então δ_{it} mede a participação relativa da economia i no tempo t em μ_t . Desse modo, δ_{it} reflete a distância idiossincrática entre a economia i e o componente comum, μ_t .

Quando aplicada a microdados de painel, a decomposição permite flexibilidade no comportamento idiossincrático tanto no tempo quanto sobre o corte transversal (*cross-section*), além de preservar o aspecto comum do desconhecido componente ordinário de crescimento. O aspecto comum revela que quando os componentes idiossincráticos convergem no tempo para uma constante, os dados de painel se sustentam de maneira similar ao conceito de σ -convergência.

Pode-se ainda ilustrar a contribuição da nova decomposição sugerida a partir de um modelo simples de único fator,

$$X_{it} = \delta_i \mu_t + \varepsilon_{it} \quad (3)$$

onde δ_i mede a distância idiossincrática entre o fator comum, μ_t , e a parcela sistemática de X_{it} . Assim, o modelo busca capturar a evolução da variável individual X_{it} em relação a μ_t por meio de dois componentes idiossincráticos: o elemento sistemático (δ_i) e o erro (ε_{it}).

Existe uma ambigüidade na interpretação do fator comum, μ_t , tendo em vista que a mesma pode refletir tanto o comportamento das economias como um todo quanto reproduzir a reação de uma economia.

Phillips e Sul (2007) fazem uma extensão na equação (3) ao permitir que o elemento sistemático idiossincrático varie no tempo. Desse modo, admite-se o comportamento heterogêneo dos agentes bem como a evolução desse comportamento no tempo mediante o coeficiente δ_{it} . Permite-se ainda que δ_{it} possua um componente aleatório, a qual incorpora o elemento ε_{it} de (3), e que possa ocasionar uma eventual convergência de δ_{it} no tempo em relação ao fator comum μ_t . Isso explica a razão pelo qual o teste de convergência absoluta se concentra em δ_{it} , ao invés de μ_t . Assim, a equação (3) pode ser igualmente transformada na (2).

Pode-se ilustrar o argumento com o modelo de crescimento neoclássico que permite progresso tecnológico heterogêneo, tal como a equação (4) a seguir:

$$\log y_{it} = \log y_i^* + (\log y_{i0} - \log y_i^*) e^{-\beta_{it}} + \log A_{it} = a_{it} + \log A_{it} \quad (4)$$

onde $\log y_{it}$ é o logaritmo da renda real *per capita* da economia i no tempo t , $\log y_i^*$ é o nível de estado estacionário do logaritmo da renda real *per capita* efetiva, $\log y_{i0}$ é o logaritmo da renda real inicial *per capita* efetiva, β_{it} é a variação no tempo da velocidade de convergência, $\log A_{it}$ significa o logaritmo da acumulação de tecnologia para a economia i no tempo t e a_{it} denota os componentes transitórios. A relação entre as equações (2) e (4) pode ser sumarizada em (5), em que se introduz o componente comum de crescimento, μ_t , entre as economias e o parâmetro de transição, δ_{it} .

$$\log y_{it} = \left(\frac{a_{it} + \log A_{it}}{\mu_t} \right) \mu_t = \delta_{it} \mu_t, \quad (5)$$

4.2.2. Transição relativa e convergência absoluta

Uma importante característica da representação de fator variável no tempo é apresentar uma nova abordagem à modelagem de convergência absoluta. De modo geral, a

aplicação de séries temporais na macroeconomia tem se realizado pelo uso extensivo de métodos de cointegração para o estudo de longo prazo, e por análises de estacionariedade para o curto prazo.

Na presente abordagem, o emprego de tendências estocásticas comuns aceita movimentos conjuntos de longo prazo do painel, sem exigir a existência de cointegração, e, inclusive, permitindo a captura de efeitos transitórios. Em particular, os fatores de carregamento idiossincráticos, δ_{it} , fornecem um mecanismo de comportamento heterogêneo entre economias assim como a possibilidade de que um período de transição numa trajetória que seja eventualmente governada por alguma tendência estocástica comum de longo prazo.

De acordo com a teoria de séries temporais, se duas variáveis macroeconômicas, X_{it} e X_{jt} , possuem tendências estocásticas e se acredita que estejam em equilíbrio de longo prazo, então as séries são hipoteticamente consideradas cointegradas e são realizados testes para verificar a premissa. Os testes de cointegração são usualmente semiparamétricos com relação à dinâmica de curto prazo e dependem de um extenso período de tempo.

Daí, necessita-se de métodos econométricos novos para analisar as duas dinâmicas simultaneamente. Suponha no contexto descrito pela equação (2) em que os coeficientes de carregamento δ_{it} converjam lentamente para δ no tempo, porém os dados disponíveis de análise sejam limitados. A diferença entre duas séries temporais no painel é dada por $X_{it} - X_{jt} = (\delta_{it} - \delta_{jt}) \mu_t$. Se μ_t possui raiz unitária não estacionária e $\delta_{it} \neq \delta_{jt}$, então X_{it} não vai ser geralmente cointegrado com X_{jt} .

Haja vista que δ_{it} e δ_{jt} convergem para δ quando $t \rightarrow \infty$, pode-se deduzir que X_{it} e X_{jt} são assintoticamente cointegráveis. No entanto, se a velocidade de divergência de μ_t for maior que a velocidade de convergência de δ_{it} , o resíduo $(\delta_{it} - \delta_{jt}) \mu_t$ pode reter características não estacionárias e os testes padrões de cointegração terão pouco poder de detectar o movimento conjunto assintótico. Portanto, torna-se indispensável a utilização de outro método econométrico para avaliar os movimentos conjuntos das variáveis em um contexto de convergência absoluta e heterogeneidade das economias.

Uma simples maneira porém intuitiva forma de se definir equilíbrio relativo de longo prazo ou convergência absoluta entre regiões/países é exibi-la em termos da razão ao invés da diferença. Ou seja, o equilíbrio relativo de longo prazo existe entre os X_{it} se

$$\lim_{k \rightarrow \infty} \frac{X_{it+k}}{X_{jt+k}} = 1 \text{ para todo } i \text{ e } j. \quad (6)$$

No contexto da equação (2), essa última condição é equivalente a convergência dos coeficientes dos fatores de carregamento,

$$\lim_{k \rightarrow \infty} \delta_{it+k} = \delta \quad (7)$$

4.2.3. Transição relativa

Já que μ_t é um fator comum na equação (2), pode-se eliminar o mesmo quando se realiza a seguinte divisão,

$$h_{it} = \frac{X_{it}}{\frac{1}{N} \sum X_{it}} = \frac{\delta_{it}}{\frac{1}{N} \sum \delta_{it}} \quad (8)$$

A equação (8) é chamada de parâmetro de transição relativa que mede o coeficiente de carregamento idiossincrático δ_{it} em relação à média do painel no tempo t . Tal como δ_{it} , o parâmetro h_{it} fornece igualmente o caminho de transição para uma economia i , porém isso ocorre em relação à média do painel.

As propriedades imediatamente observadas do parâmetro h_{it} são que a média do corte transversal é, por definição, igual a um, e que se o coeficiente de carregamento idiossincrático δ_{it} tende para δ , então h_{it} converge para a unidade. Assim, no longo prazo, a variância do corte transversal de h_{it} converge para zero (equação 9).

$$\text{Se } \delta_{it} \rightarrow \delta, \text{ então } \sigma_t^2 = \frac{1}{N} \sum_{i=1}^N (h_{it} - 1)^2 \rightarrow 0, \text{ quando } t \rightarrow \infty \quad (9)$$

4.2.4. Testes de convergência absoluta

Para realizar os testes de convergência absoluta, torna-se necessário padronizar X_{it} de tal maneira que se obtenha uma função limítrofe que incorpore tanto os componentes comuns quanto os caminhos de transição.

Suponha que $t = [Tr]$, tal que t represente a parte inteira de Tr , e r seja a fração efetiva da amostra T que corresponda a observação t . O valor de r é calculado para cada t , tal que quando $t = T$, $r = 1$, e quando $t = 0$, $r = 0$. Isto é, r pode ser escrito como $r(t)$, $r(0) = 0$ ou $r(T) = 1$.

Em Phillips e Sul (2005 e 2007), são exibidos maiores detalhes da padronização de X_{it} . A partir disso, chega-se a seguinte função limítrofe da equação de transição relativa (8):

$$h_{it}\left(\frac{t}{T}\right) = \frac{d_{iT}^{-1} X_{it}}{\frac{1}{n} \sum_{j=1}^n d_{jT}^{-1} X_{jt}} = \frac{\delta_{iT}\left(\frac{t}{T}\right)}{\frac{1}{n} \sum_{j=1}^n \delta_{jT}\left(\frac{t}{T}\right)} \quad (10)$$

Por sua vez, a equação (11) mostra $h_i(r)$ que representa a forma limítrofe da curva de transição relativa para a economia i .

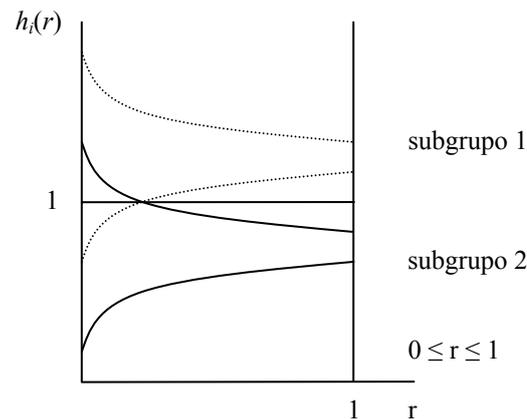
$$h_{it}\left(\frac{[Tr]}{T}\right) \xrightarrow{p} h_i(r) = \frac{\delta_i(r)}{\frac{1}{n} \sum_{j=1}^n \delta_j(r)} \quad \text{quando } T \rightarrow \infty \quad (11)$$

Os testes de convergência se baseiam na análise das curvas de transição ajustadas, $\hat{h}_i(r)$, que indicam a possibilidade de seja convergência transitória para alguns sub-grupos de economias seja divergência para outros.

Em um contexto de convergência absoluta, a variação do corte transversal de $\hat{h}_i(r)$ tende para zero quando t vai para o infinito (vide equação 9). No entanto, uma redução

da variação de $\hat{h}_i(r)$ não implica em convergência absoluta (no sentido de tendência para uma distribuição degenerada) tal como se observa na Figura 4.1.

Figura 4.1
Variação no Corte Transversal



A Figura 4.1 nos mostra que apesar de ocorrer a redução da variação no corte transversal existe convergência localizada para dois subgrupos, ou seja, divergência absoluta. O teste de convergência absoluta deve levar em consideração a possibilidade de haver clubes de convergência.

O teste de hipótese pode ser descrito de acordo com a hipótese nula,

$$H_0: \delta_i = \delta \text{ e } \alpha \geq 0,$$

contra a hipótese alternativa

$$H_a: \delta_i \neq \delta \text{ para todo } i, \text{ ou } \alpha < 0.$$

O primeiro passo para efetuar na prática o teste de convergência consiste em construir a razão H_1/H_t da variância do corte transversal tal que

$$H_t = \frac{1}{N} \sum_{i=1}^N (h_{it} - 1)^2, \quad h_{it} = \frac{X_{it}}{N^{-1} \sum_{i=1}^N X_{it}} \quad (12)$$

O segundo passo baseia-se em realizar a seguinte regressão com o ajuste consistente de autocorrelação e heterocedasticidade dos resíduos (*heteroskedasticity and autocorrelation consistent* – HAC):

$$\log(H_1/H_t) - 2 \log L(t) = \hat{a} + \hat{b} \log t + \hat{u}_t, \quad \text{para } t = [rT], [rT] + 1, \dots, T \text{ com } r > 0 \quad (11)$$

onde $L(t) = \log(t + 1)$, o coeficiente estimado de $\log t$ é $\hat{b} = 2\hat{\alpha}$ onde $\hat{\alpha}$ é a estimação de α em H_0 . A sugestão baseada em simulações de Monte Carlo de Phillips e Sul (2007) é que se adote $r = 0,3$ (valor satisfatório em termos de tamanho e potência) e que a regressão comece a partir do número inteiro de $t = [rT]$.

Sob a hipótese de convergência, $h_{it} \rightarrow 1$ (equação (8)) e $H_t \rightarrow 0$ (equação (10)) quando $t \rightarrow \infty$ para um dado N . E a partir da equação (11), $\log(H_1/H_t)$ diverge para ∞ , seja para $2 \log L(t)$ quando $\alpha = 0$, ou para $2\alpha \log t$ quando $\alpha > 0$. Ou seja, ocorre convergência (a variável dependente diverge) seja para $\alpha = 0$ ou $\alpha > 0$. Visto que α é um escalar, a hipótese nula pode ser averiguada mediante teste t unicaudal.

Quando se considera a hipótese de divergência sob H_a , isto é $\alpha < 0$, H_t tende a convergir para uma quantidade positiva sob $t \rightarrow \infty$. Assim, a variável dependente $\log(H_1/H_t) - 2 \log L(t)$ diverge para $-\infty$ sob H_a .

A rejeição da hipótese de convergência não implica na inexistência de convergência em subgrupos do painel. Há várias possibilidades que possam ocorrer quando se rejeita a hipótese nula. Pode-se citar como exemplo uma situação de grupos de convergência em torno de pontos distintos de equilíbrio, isto é, diferentes trajetórias de estado estacionário, ou mesmo quando se verifica tanto grupos de convergência quanto outras economias de comportamento divergente no painel.

Quando se suspeita que existem grupos (ou clubes) de convergência, torna-se primordial identificar esses agrupamentos ao se determinar o número de grupos e separar cada membro em seu respectivo grupo.

Phillips e Sul (2007) desenvolvem um algoritmo baseado em repetidas regressões de $\log t$ para classificar os membros nos grupos. Os passos desse algoritmo podem ser resumidos a seguir.

1º passo: Ordenar os membros do painel de acordo com a última observação;

2º passo: Calcular a estatística t de convergência, t_k , para uma seqüência de regressões $\log t$ baseada nos k maiores membros (passo 1), sendo $2 \leq k \leq N$. A escolha do tamanho do grupo vai depender do valor máximo do teste t_k , onde $t_k > -1,65$.

3º passo: Selecionar os membros a comporem os grupos adicionando um integrante por vez. Incluir um novo membro no grupo se a estatística t associada for maior que zero (posição conservadora). Deve-se confirmar que o grupo satisfaz o critério de convergência.

4º passo: Os membros que não forem selecionados no passo 3 devem formar um grupo complementar. Aplica-se uma regressão $\log t$ para testar o comportamento desse novo conjunto de membros. Se ocorrer a convergência para o novo grupo, então há a formação de um segundo grupo. Caso contrário, repita os passos de 1 a 3 para averiguar se há mais sub grupos convergentes. Se nenhum outro grupo for encontrado, então esses membros apresentam comportamento divergente.

4.3. Análise Empírica

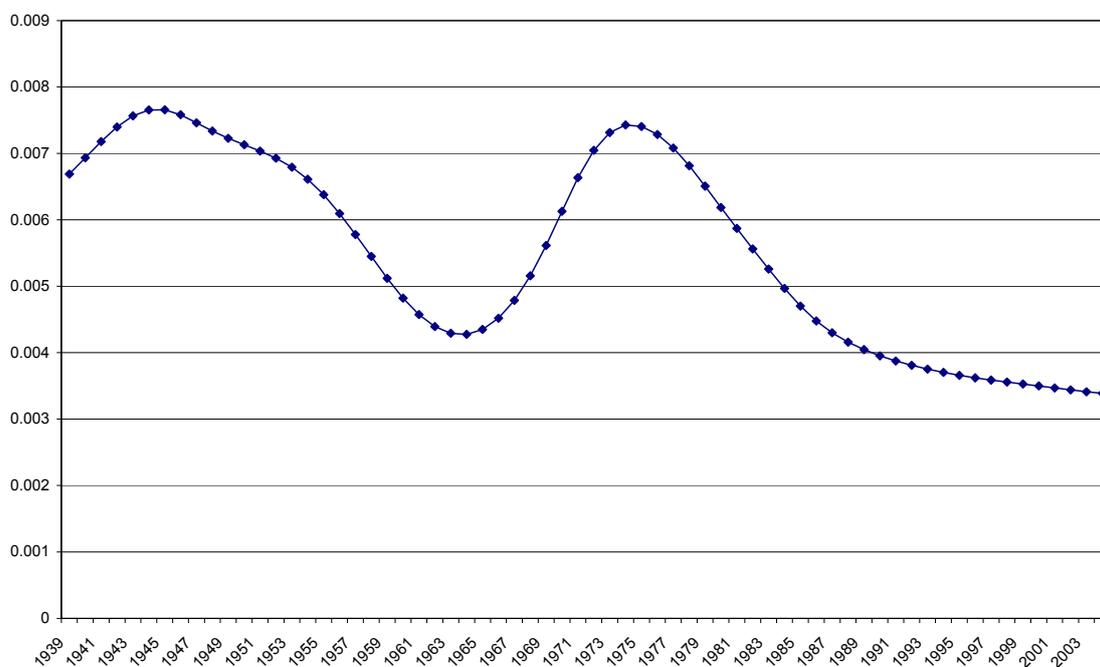
Os dados utilizados para esta seção são os do PIB *per capita* dos estados brasileiros referente aos anos de 1939 a 2004. Tendo em vista o interesse pelo comportamento de longo prazo, foi removido da série o componente de ciclos reais (*business cycle*) de X_{it} ao se empregar o filtro de Hodrick-Prescott (1997). O valor adotado no filtro para o parâmetro de suavização λ foi 100, que corresponde ao valor recomendado para dados anualizados.

Apesar de não se exibir neste trabalho, foram realizados os testes subsequentes sem a utilização do filtro Hodrick-Prescott. No entanto, a desconsideração do filtro não modificaria as análises obtidas nesta seção.

Antes de se mostrar o resultado das regressões realizadas, é interessante observar algumas particularidades de algumas variáveis. Em primeiro lugar, observa-se no Gráfico 4.1 o comportamento da variável H_t nos anos da amostra coletada.

Gráfico 4.1

Evolução de H_t (Variância do Corte Transversal de h_{it}) – 1939 a 2004



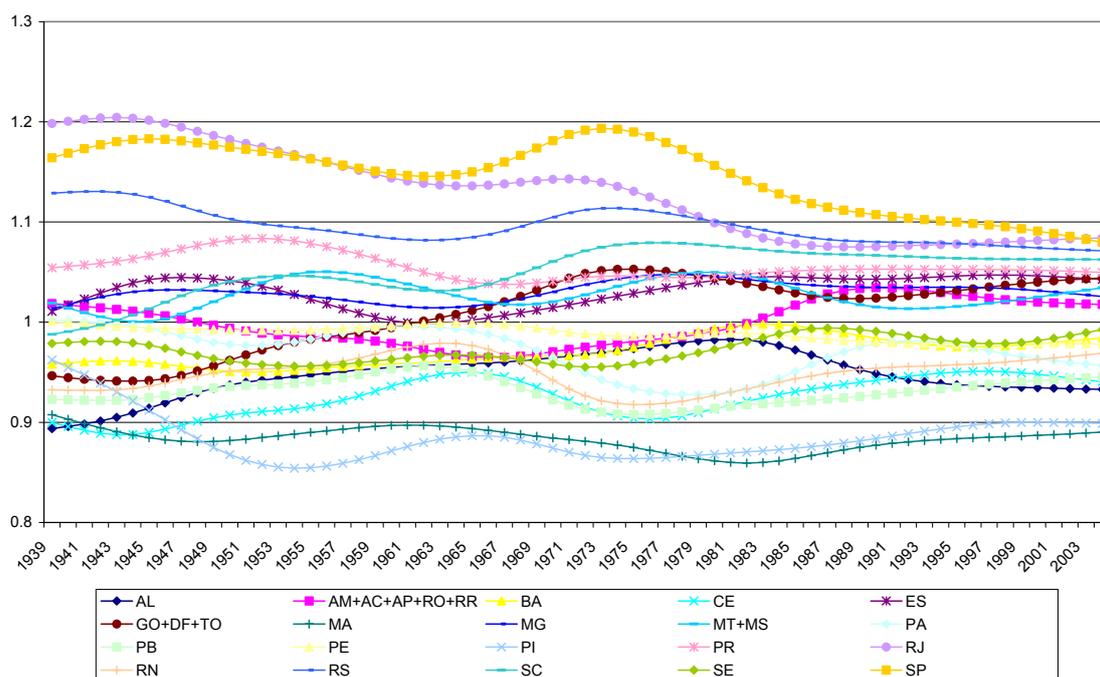
A variável H_t , ou variância do corte transversal de h_{it} , representada pela equação (12) reproduz uma medida de desigualdade tendo em vista que, em um contexto de convergência absoluta, seu valor tende a zero quando t vai para o infinito.

Analisando o Gráfico 4.1, notam-se dois períodos de redução de desigualdade, um antes do período do milagre econômico (1969-1973) e outro após esse período. Verifica-se, portanto, que apesar do período do milagre econômico resultar em aumentos significativos dos níveis dos PIBs *per capita* dos estados brasileiros, a contrapartida

disso foi a elevação considerável da desigualdade de renda dos estados. Após o período do milagre e a partir do ano de 1985, os níveis de desigualdade já se encontram em um patamar inferior ao período pré-milagre, e apresentam sinais de queda sustentada desde então.

O Gráfico 4.2 mostra as trajetórias relativas de transição (equação 8) do PIB *per capita* de cada estado. Em outras palavras, observa-se nesse gráfico o caminho de crescimento de cada estado em relação à média da amostra. Os pontos do eixo das ordenadas acima de um representam os estados com a renda *per capita* acima da média e vice-versa. Em um contexto de convergência absoluta, as trajetórias dos estados, h_{it} , deveriam tender assintoticamente para um (equação 8).

Gráfico 4.2
Trajetórias Relativas de Transição, h_{it} – 1939 a 2004



A análise do Gráfico 4.2 pode ser simplificada quando se observa o Gráfico 4.1 uma vez que se observa igualmente nesse último um aumento na dispersão das rendas *per capita* no período do milagre (1969-1973), e uma redução gradativa da dispersão após o mencionado período. O aumento ou redução na dispersão pode ser verificado de acordo com a variação da amplitude vertical das trajetórias relativas de transição. Desse modo,

quanto mais as trajetórias tenderem para o valor 1 (eixo das ordenadas), menor a dispersão.

Na Tabela 4.1, são apresentados os resultados da metodologia de Phillips e Sul (2007) para os PIB *per capita* estaduais. As estimativas dos erros padrões foram calculadas de acordo o estimador consistente heterocedástico e autocorrelativo de Newey e West (1987).

Tabela 4.1
Resultado das Regressões para Períodos Seleccionados

Período	Coeficiente \hat{b}	Desvio-padrão	Estatística t	Probabilidade
1939 a 2004	-0,14	0,18	-0,77	0,44
1939 a 1968	-1,94	0,96	-2,02	0,06
1973 a 2004	-3,56	0,25	-14,06	0,00

Nota-se que a hipótese nula de convergência não foi rejeitada para o período de 1939 a 2004. A rejeição da hipótese nula ocorreria para valores $t_{\hat{b}} < -1,65$ ao nível de 5% de significância (coluna 4 da Tabela 4.1). Não obstante, Apergis, Panopoulou e Tsoumas (2008) salientam para o fato de que um teste mais conservador para avaliar a convergência absoluta seria assumir $t_{\hat{b}} > 0$.

Se fizermos uma divisão entre os períodos pré e pós milagre (linhas 3 e 4 da Tabela 4.1) verifica-se a rejeição da hipótese de convergência absoluta. Atenta-se para o fato que a estatística $t_{\hat{b}}$ do período pós milagre (linha 4, coluna 4) apresenta uma distância relativamente grande do nível crítico de significância 5%, $t_{\hat{b}} = -1,65$. Ou seja, rejeita-se fortemente a hipótese de convergência absoluta dos PIB *per capita* dos estados brasileiros para o período pós milagre.

Uma vez que não se constatou convergência absoluta para os períodos pré e pós milagre, busca-se verificar seja a existência de clubes de convergência ou a divergência

no comportamento da renda. Na Tabela 4.2, são exibidos os resultados das regressões de acordo com o algoritmo proposto por Phillips e Sul (2007).

Tabela 4.2
Resultado das Regressões – Pré e Pós Milagre

Período	Grupo de países ⁷	Coefficiente \hat{b}	Desvio-padrão	Estatística t	Probabilidade
1939 a 1968	1º grupo (A)	2,30	1,43	1,60	0,12
	2º grupo (A)	0,08	1,33	0,06	0,95
1974 a 2004	1º grupo (B)	3,68	1,25	2,93	0,01
	2º grupo (B)	2,39	0,80	2,96	0,01

Antes do milagre

1º grupo (A): Goiás+Distrito Federal + Tocantins, Minas Gerais, Mato Grosso + Mato Grosso do Sul, Paraná, Rio de Janeiro, Rio Grande do Sul, Santa Catarina e São Paulo

2º grupo (A): Alagoas, Amazonas + Acre + Amapá + Rondônia + Roraima, Bahia, Ceará, Espírito Santo, Maranhão, Pará, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe.

Após o milagre

1º grupo (B): Amazonas + Acre + Amapá + Rondônia + Roraima, Espírito Santo, Goiás+Distrito Federal + Tocantins, Minas Gerais, Mato Grosso + Mato Grosso do Sul, Paraná, Rio de Janeiro, Rio Grande do Sul, Santa Catarina e São Paulo

2º grupo (B): Alagoas, Bahia, Ceará, Maranhão, Pará, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe.

É interessante notar que não há comportamento divergente do PIB *per capita* dos estados seja antes ou depois do período do milagre, pois nenhuma das estatísticas t

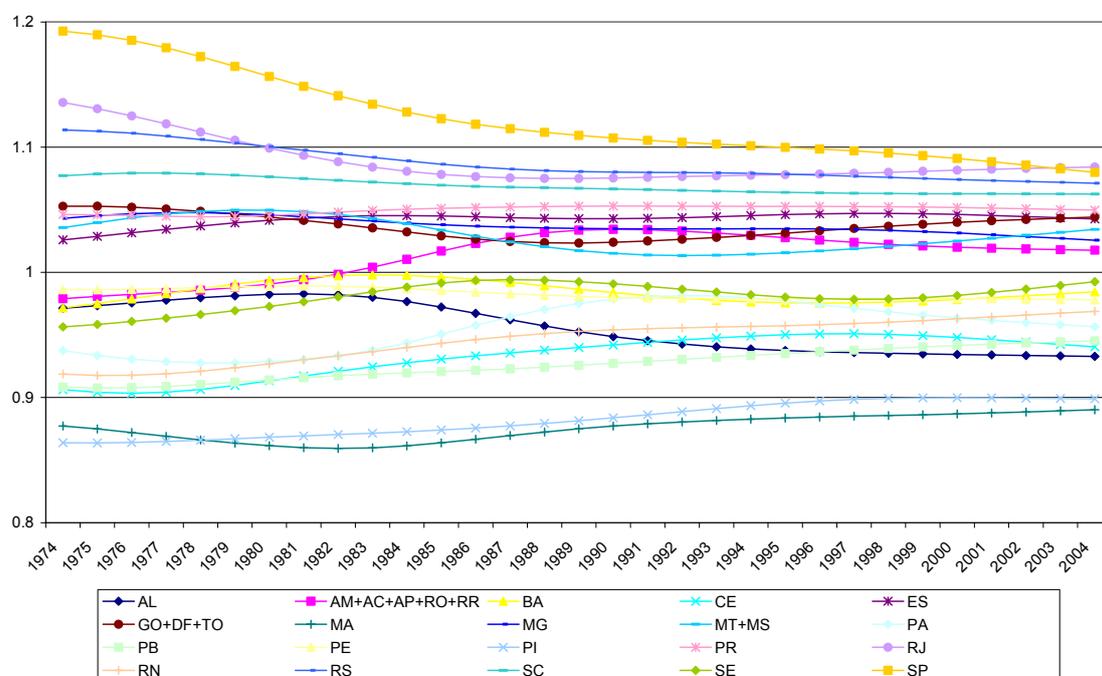
⁷ Os estados indicados pelo sinal + são tratados de forma aglomerada tendo em vista que alguns desses estados foram criados ao longo do período da amostra coletada.

apresenta sinal negativo na 5ª coluna da Tabela 4.2. Assim, tanto antes como depois do milagre houve a presença de dois grupos de convergência de renda. Os clubes são denominados por (A) antes e (B) depois do milagre.

A diferenciação entre os grupos A e B é a transferência dos estados do Espírito Santo e o conjunto de Amazonas, Acre, Amapá, Rondônia e Roraima para o clube de convergência de maior PIB *per capita* no período pós milagre. A entrada do Espírito Santo para o clube de convergência de renda mais alta pode ser conferida por ser um estado da região sudeste com a proximidade dos estados mais ricos do Brasil (mostrou-se no Capítulo 2 a importância da questão da dependência espacial/regional dos dados estaduais brasileiros). Por outro lado, a inserção do grupo Amazonas, Acre, Amapá, Rondônia e Roraima no clube de renda mais alta pode ser atribuída à transformação dos territórios em estados mediante a Constituição Federal de 1988, gerando maiores aportes de recursos do governo federal, e isso pode ser visualizado conforme a trajetória relativa de transição desse grupo no Gráfico 4.3.

Gráfico 4.3

Trajetórias Relativas de Transição, h_{it} – 1974 a 2004



O Gráfico 4.3 é similar ao Gráfico 4.2, mas com enfoque no período pós milagre econômico. Vale notar que no ano de 2004, os estados cujos pontos do eixo das ordenadas estejam acima do valor 1 representam o primeiro grupo de estados e os abaixo desse valor, o segundo grupo.

4.4. Conclusão

Entre os benefícios encontrados na metodologia de Phillips e Sul (2007) são: (1) permitir que o comportamento sistemático idiossincrático de economias heterogêneas varie no tempo; (2) teste de convergência que não supõe comportamento específico de estacionariedade nas variáveis de interesse; (3) utilização de um modelo não-linear que possibilite uma maior amplitude em termos de trajetórias; (4) no caso de as variáveis de interesse sejam estocasticamente não-estacionárias, os testes de convergência são distintos aos convencionais testes de cointegração. Isso permite, por exemplo, uma melhor adequação no caso de comportamento transitório idiossincrático de cada economia.

O melhor ajuste para a série dos PIBs *per capita* dos estados brasileiros foi detecção de dois clubes de convergência para o período pós milagre econômico. Utilizando-se de outras técnicas Azzoni e Silveira Neto (2006) apontam para a formação de dois clubes de convergência para o período de 1985 a 2001, um composto de estados de mais baixa renda *per capita* das regiões Norte e Nordeste, e outro composto de estados mais ricos das regiões Sul e Sudeste.

Gondim *et alii* (2007), Hewings *et alii* (2005), Mossi *et alii* (2003) e Maciel *et alii* (2008) encontram igualmente evidências na formação de clubes de convergência em diversos períodos analisados apesar de não usarem a metodologia de dados de painel dinâmicos.

CONCLUSÃO GERAL

Procurou-se nesta Tese avaliar as diferentes metodologias de análise de convergência para os estados brasileiros, bem como foram realizados testes econométricos a fim de se estabelecer confronto entre as diversas óticas.

Entende-se que a seqüência de capítulos nesta Tese mostra a evolução das metodologias para aferir convergência. Isto é, cada capítulo subsequente apresenta vantagens técnicas em relação ao anterior.

No Capítulo 1, os estudos de corte transversal (*cross-section*) mostraram um processo de convergência de renda *per capita* dos estados brasileiros para o período pós milagre econômico sejam pelas análises de β -convergência (p. 14) quanto σ -convergência (p. 19).

Ferreira e Diniz (1995) e Ferreira e Ellery Jr. (1996) aceitaram igualmente a hipótese de β -convergência e σ -convergência para o período de 1970 a 1985. Por outro lado, Azzoni (2001) encontra a presença de convergências regionais mediante os testes de β -convergência e σ -convergência para o período de 1939 a 1995.

Verificou-se ainda no Capítulo 1 um aumento significativo na desigualdade da renda *per capita* estadual no período do milagre seja pelo coeficiente positivo da regressão para aferir β -convergência (p. 14) seja pelas elevações dos indicadores de coeficiente de variação de Williamson e índice de Theil no período entre 1969 a 1973 (p.18). Portanto, torna-se fundamental no estudo de convergência fazer menção ao período do milagre econômico quando se faz uma análise histórica do caso brasileiro.

A crítica de Bernard e Durlauf (1996) em relação às abordagens de corte transversal (*cross-section*) são as inconsistências nas avaliações de β -convergência e sugere a alternativa proposta por Quah (1993a,b) na adoção de cadeias de Markov para avaliar convergência.

Ferreira (1998) faz a análise de Markov para o caso da renda *per capita* dos estados brasileiros entre os anos de 1970 e 1995, mas não faz avaliação dos pressupostos de

estabilidade da matriz de transição. Isso é feito no Capítulo 2, e constatou-se que as hipóteses para adoção de cadeias de Markov para o caso dos estados brasileiros foram violados, tais como a autoregressividade da renda (p. 33) e homogeneidade espacial (p. 30).

A técnica de decomposição da desigualdade proposta por Jenkins e Van Kerm (2003) apresenta um avanço em relação ao estudo de β -convergência e σ -convergência. Torna-se possível diagnosticar o efeito ultrapassagem de tal maneira a alterar o ordenamento (*ranking*) das rendas.

Quando se observa as Figuras 3.1 e 3.2 do Capítulo 3 (p. 41), que representam duas economias onde ocorrem β -convergência e σ -convergência, nota-se a importância do efeito ultrapassagem uma vez que no longo prazo ($t \rightarrow \infty$) as economias da Figura 3.1 apresentam uma tendência de convergência absoluta de renda *per capita* e as da Figura 3.2 mostram um caminho de divergência. Apesar disso, as avaliações feitas sobre o efeito ultrapassagem no caso dos estados brasileiros revelaram que o mesmo possui um papel diminuto na avaliação de convergência (p. 50).

As técnicas utilizadas nos Capítulos 1 a 3 não permitem diagnosticar sub-grupos de trajetórias temporais múltiplas da evolução da renda *per capita*, tais como algumas economias convergindo e outras com comportamento divergente. Dessa forma, torna-se factível no Capítulo 4 verificar a formação de clubes de convergência, onde existe a possibilidade de predominar dois ou mais *steady-states*.

Assim, aplica-se no Capítulo 4 a metodologia de Phillips e Sul (2007) ao caso dos estados brasileiros onde é possível testar convergência em economias que transitam ao estado estacionário e considerando a heterogeneidade de cada economia variando no tempo.

Nesse capítulo, foi diagnosticada a presença de dois clubes de convergência para o caso dos estados brasileiros para o período pós milagre econômico (p. 64) – um baseado no Norte e Nordeste e outro no Centro-Sul.

Diversos trabalhos, tal como Azzoni e Silveira Neto (2006), Gondim *et alii* (2007), Hewings *et alii* (2005), Mossi *et alii* (2003) e Maciel *et alii* (2008) encontram resultados semelhantes de clubes de convergência, porém empregando técnicas diferentes.

Destaca-se Maciel *et alii* (2008) que buscou avaliar o processo de convergência das regiões brasileiras entre 1985 e 2000 a partir de uma adaptação do modelo Hall e Jones (1999) que incorpora mobilidade de trabalho e capital. Apesar de metodologia diferente, Maciel *et alii* (2008) chegam a resultados similares aos encontrados no Capítulo 4. As simulações indicam um processo de concentração espacial da atividade econômica nas regiões Sudeste, Centro-Oeste e Sul, enquanto as regiões Norte e Nordeste permanecem em um patamar com taxas de crescimento mais baixas. A razão justificada pelos autores para esse processo de concentração de renda nas regiões mais ricas é a superioridade dos produtos marginais do capital físico e da mão-de-obra apesar de apresentarem estoques de capital maiores.

Em relação à desigualdade regional observada no período pós milagre econômico, Barros *et alii* (2000) argumentam que as taxas salariais são na média iguais para as mesmas qualificações dos trabalhadores em qualquer região do país. Portanto, a desigualdade que se encontra nas regiões relativamente mais pobres, tais como Norte e Nordeste, é resultado do perfil das distribuições das qualificações e, não, dos diferenciais salariais entre as regiões.

Um aspecto curioso que merece atenção para um trabalho futuro é saber quais as causas e a maneira que se desenvolveu para que ocorresse um aumento na desigualdade da renda *per capita* dos estados brasileiros no período do milagre, exibidos nos Capítulos 1, 3 e 4. Atenta-se ao fato que esse fenômeno de concentração de renda estadual foi exclusivo ao período do milagre econômico (1968-1973).

Outras sugestões de trabalhos futuros podem ser feitas no sentido de analisar a evolução da decomposição das rendas dos estados por meio dos fatores de produção, capital e trabalho. Paralelamente, seria interessante avaliar a magnitude que a migração dos fatores de produção responde na desigualdade regional.

REFERÊNCIAS BIBLIOGRÁFICAS

- Apergis, N., Panopoulou, E. e Tsouman, C. (2008). “Old Wine in a New Bottle: Growth Convergence Dynamics in the EU”. Universidade de Piraeus, Grécia. Working Paper de 29 de fevereiro de 2008.
- Azzoni, C.R. (1997). “Concentração Regional e Dispersão das Rendas *per capita* Estaduais: Análise a partir de Séries Históricas Estaduais de PIB, 1939-1995”. Estudos Econômicos, v. 27, n. 3.
- _____. (2001). “Economic Growth and Regional Income Inequality in Brazil”. The Annals of Regional Science, n. 35, p. 133-152.
- Azzoni, C.R. e Silveira Neto, R. (2006). “Location and Regional Income Disparity Dynamics: The Brazilian Case”. Papers in Regional Science, v. 85, n. 4, p. 599-613.
- Barro, R.J. (1991). “Economic Growth in a *Cross-section* of Countries”. The Quarterly Journal of Economics, v.106, n. 2, p. 407-443.
- Barro, R.J. e Sala-i-Martin, X. (1990). “Economic Growth and Convergence Across the United States”. NBER, Working Paper, n. 3419.
- _____. (1991). “Convergence across States and Regions”. Brookings Papers on Economic Activity, v. 22, n. 1, p. 107-182.
- _____. (1992). “Convergence”. Journal of Political Economy, v. 100, n. 2, p. 223-251.
- Barros, R.P., Henriques, R. e Mendonça, R. (2000). “Education and Equitable Economic Development”. Economia, v.1, n.1, p. 111-144.

- Baumol, William J. (1986). "Productivity Growth, Convergence, and Welfare". *American Economic Review*, v. 76, n. 5, p. 1072-85.
- Bernard, A.B. e Durlauf, S.N. (1995). "Convergence in International Output". *Journal of Applied Econometrics*, v. 10, n. 2, p. 97-108.
- _____. (1996). "Interpreting Tests of the Convergence Hypothesis". *Journal of Econometrics*, v. 71, n. 1-2, p. 161-173.
- Bickenbach, Frank e Bode, Eckhardt (2001). "Markov or Not Markov – This Should Be a Question". Kiel Institute of World Economics, Kiel Working Paper n. 1086.
- Durlauf, S.N. e Quah, D.T. (1999). *The New Empirics of Economic Growth*. Handbook of Macroeconomics, editora North-Holland.
- Essama-Nssah, B. (2002). "Assessing the Distributional Impact of Public Policy". Banco Mundial, Policy Research Working Paper, n. 2883.
- Evans, P. (1998). "Using Panel Data to Evaluate Growth Theories". *International Economic Review*, v. 39, n. 2, p. 295-306.
- Evans, P. e Karras, G. (1996). "Convergence Revisited". *Journal of Monetary Economics*, v. 37, n. 2, p. 249-265.
- Ferreira, A. (1996). "A Distribuição Interestadual da Renda no Brasil (1950-85)". *Revista Brasileira de Economia*, v. 50, n. 4, p. 469-485.
- Ferreira, A.H. (1998). "Concentração Regional e Dispersão das Rendas *per capita* Estaduais: Um Comentário". Cedeplar – UFMG, Texto para Discussão, n. 121.
- Ferreira, A.H. e Diniz, C.C. (1995). "Convergência entre Rendas *per capita* Estaduais no Brasil". *Revista de Economia Política*, v. 15, n. 4, p. 38-56.

- Ferreira, P.C.G. e Ellery Jr., R.G. (1996). “Convergência entre a Renda *per capita* dos Estados Brasileiros”. *Revista de Econometria*, Rio de Janeiro, v. 16, n. 1, p. 83-104.
- Fleissig, A. e Strauss, J. (2001). “Panel Unit-Root Tests of OECD Stochastic Convergence”. *Review of International Economics*, v. 9, n. 1, p. 153-162.
- Friedman, M. (1992). “Do Old Fallacies Ever Die?”. *Journal of Economic Literature*, v. 30, n. 4, p. 2129-2132.
- Giambiagi, F. (1988). “A Economia Brasileira nos Anos do “Milagre”: 1968/1973.” FEA/UFRJ. Texto Didático n. 34.
- Gondim, J.L.B., Barreto, F.A. e Carvalho, J.R. (2007). “Condicionantes de Clubes de Convergência no Brasil”. *Estudos Econômicos*, São Paulo, v. 37, n. 1, p. 71-100.
- Hall, R. e Jones, C. (1999). “Why Do Some Countries Produce So Much More Output Per Worker Than Others?”. *The Quarterly Journal of Economics*, v. 114, n. 1, p. 83-116.
- Hewings, A., Magalhães, A. e Azzoni, C. (2005). “Spatial Dependence and Regional Inequality in Brazil”. *Investigaciones Regionales*, v.6.
- Hodrick, R. e Prescott, E. (1997). “Post War Business Cycles: An Empirical Investigation”. *Journal of Money, Credit and Banking*, v. 29, n. 1, p. 1-16.
- Islam, N. (1995). “Growth Empirics: A Panel Data Approach”. *The Quarterly Journal of Economics*, v. 110, n. 4, p. 1127-1170.
- Jenkins, S. e Van Kerm, P. (2003). “Trends in Income Inequality, Pro-Poor Income Growth and Income Mobility”. CEPS/INSTEAD, Differdange, G.D, Luxemburgo, IRISS Working Paper, n. 2033-11.

- Kakwani, N. C. (1977). “Applications of Lorenz Curves in Economic Analysis”. *Econometrica*, v. 45, n. 3, p. 719-727.
- King, R.G. e Rebelo, S.T. (1989) “Transitional Dynamics and Economic Growth in the Neoclassical Model”. NBER, Working Paper n. 3185.
- Kocenda, E. (2001). “Macroeconomic Convergence in Transition Economies”. *Journal of Comparative Economics*, v. 29, n. 1, p. 1-23.
- Kutan, A.M. e Yigit, T.M. (2004). “Nominal and Real Stochastic Convergence within the Transition Economies: Panel Evidence”. *Journal of Comparative Economics*, v. 32, p. 23-36.
- Lago, L.A. (1990). “A Retomada do Crescimento e as Distorções do “Milagre”: 1967-1973”. In Abreu, M. (org.), *A Ordem do Progresso*. Editora Campus, Rio de Janeiro.
- Lee, H.A., Kim, K. e Azali, M. (2005). “Income Disparity between Japan and Asia – 5 Economies: Converge, Catching-up or Diverge?”. *Economics Bulletin*, v. 6, n. 13, p. 1-20.
- Lee, K., Pesaran, H. and Smith, R. (1997). “Growth and Convergence in Multi-country Empirical Stochastic Solow Model”. *Journal of Applied Econometrics*, v. 12, n. 4, p. 357-392.
- Li, Q. e Papell, D. (1999). “Convergence of International Output: Time Series Evidence for 16 OECD countries”. *International Review of Economics and Finance*, v. 8, n. 3, p. 267-280.
- Maciel, P.J., Andrade, J. e Teles, V.K. (2008). “Convergência Regional Brasileira Revisitada”. *Ipea. Pesquisa e Planejamento Econômico*, v. 38, n. 1, p. 37-65.

- Mankiw, N. G., Romer, D. e Weil, D. N. (1992). “A Contribution to the Empirics of Economic Growth”. *The Quarterly Journal of Economics*, v. 107, n. 2, p. 407-437.
- Mossi, M.B., Aroca, P., Fernandez, I.J. e Azzoni, C.R. (2003). “Growth Dynamics and Space in Brazil”. *International Regional Science Review*, v. 26, n. 3, p. 393-418.
- Newey, Whitney e West, Kenneth (1987). “A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. *Econometrica*, v. 55, n.3, p. 703-708.
- O’Neill, D. e Van Kerm, P. (2004) “A New Approach for Analysing Income Convergence across Countries”. CEPS/INSTEAD, Differdange, G.D., Luxemburgo, IRISS Working paper, n. 2004-03.
- Oxley, L. e Greasley, D. (1995). “A Time-series Perspective on Convergence: Australia, UK and USA since 1870”. *Economic Record*, v. 71, n. 214, p. 259-270.
- Phillips, P.C.B e Sul, D. (2005). “Economic Transition and Growth”. Universidade de Yale, EUA. Cowles Foundation Discussion Paper, n. 1514.
- _____. (2007). “Transition Modeling and Econometric Convergence Tests”. *Econometrica*, v. 75, n. 6, p. 1771-1855.
- Quah, D.T. (1993a) “Galton’s Fallacy and Tests of the Convergence Hypothesis”. *The Scandinavian Journal of Economics*, v. 95, n. 4, p. 427-443.
- _____. (1993b) “Empirical Cross-Section Dynamics in Economic Growth”. *European Economic Review*, v. 37, n. 2, p. 426-434.
- _____. (1996). “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics”. *The Economic Journal*, v. 106, n. 437, p. 1045-1055.

- _____. (1997). “Empirics for Growth and Distribution: Stratification, Polarization and Convergence Clubs”. *Journal of Economic Growth*, v. 2, n. 1, p. 27-59.
- Rey, S. (2001a). “Spatial Dependence in the Evolution of Regional Income Distributions”. Palm Springs, CA, EUA. Artigo apresentado no 40º encontro de WRSA.
- Rey, S. (2001b). “Spatial Empirics for Economic Growth and Convergence”. *Geographical Analysis*, v. 33, n. 3, p. 195-214.
- Rey, S. e Montouri, B.D. (1999). “US Regional Income Convergence: A Spatial Econometric Perspective”. *Regional Studies*, v. 33, n. 2, p. 143-156.
- Sala-i-Martin, X. (1996). “The Classical Approach to Convergence Analysis”. *Economic Journal*, v. 106, n. 437, p. 1019-1036.
- Schechtman, E. e Yitzhaki, S. (2005). “Calculating the Extended Gini Coefficient from Grouped Data – A Covariance Presentation”. Working Paper Series. Disponível em: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=890041.
- Silva, A.B.O. e Medina, M.H. (1999). “Produto Interno Bruto por Unidade da Federação – 1985-1998”. IPEA, Texto para Discussão n. 677.
- Williamson, J.G. (1965). “Regional Inequality and the Process of National Development: A Description of the Patterns”. *Economic Development and Cultural Change*, v. 13, n. 4, p. 1-84.
- Young, A.T., Higgins, M.J., Levy, D. (2007). “Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data”. EconWPA, Macroeconomics, n. 0505008. Disponível em: <http://ideas.repec.org/p/wpa/wuwpma/0505008.html>.

APÊNDICE

Construção da Base de Dados

O período inicial da série anual começa no ano de 1939 uma vez que foi o primeiro ano disponibilizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) de renda interna por unidade de federação. Os dados dos anos de 1939 e de 1947 a 1967 se referem à renda interna (soma de todos os rendimentos gerados na economia no período tais como salários, ordenados, renda da terra, juros, lucros e impostos), que se equivalem contabilmente aos valores de PIB (soma do valor de todos os bens e serviços produzidos pela economia em um dado período).

Entre os anos de 1939 a 1966, o IBGE disponibiliza os valores dos PIBs de apenas 20 estados. Isto é, diversos estados foram criados após esse período seja por transformação de território em estado ou por divisão. Houve a necessidade de se efetuar agrupamentos a fim de se manter constante o número de estados em toda a série. Assim, a série estimada se compõe de 20 estados de 1939 a 2004, cujo tratamento se identifica com Azzoni (1997).

Os dados dos estados de Rondônia, Roraima, Amapá e Acre foram agrupados para o estado do Amazonas. Os dados de Mato Grosso do Sul foram incluídos no estado de Mato Grosso. Os dados dos estados do Distrito Federal e de Tocantins foram introduzidos no estado do Goiás. Os de Fernando de Noronha, foram inseridos no estado de Pernambuco, Serra Aimorés, no estado de Minas Gerais, e Guanabara, no estado do Rio de Janeiro.

Uma restrição dos dados disponibilizados pelo IBGE é a inexistência da contabilidade dos PIBs por unidade de federação de alguns anos. Não há divulgação desses dados para os anos de 1940 a 1946, de 1971 a 1974, de 1976 a 1979 e de 1981 a 1984. Desse modo, foi realizada uma estimação de acordo com indicadores que possuam uma alta correlação com a evolução dos PIBs estaduais. Os indicadores foram as taxas de crescimento da receita orçamentária estadual e da receita tributária estadual.

Para os anos de 1940 a 1946, foram utilizados os dados da receita orçamentária estadual disponibilizados no Anuário Estatístico do IBGE tendo em vista que não há disponibilidade dos dados da receita tributária estadual para esses anos. E, para os anos de 1971 a 1974, de 1976 a 1979 e de 1981 a 1984, estimou-se de acordo com a evolução da receita tributária estadual uma vez que esse indicador apresenta uma correlação ainda maior ao PIB.

Após as estimativas, os valores dos PIBs foram corrigidos pelo deflator implícito do PIB, elaborado pelo IBGE, para se ajustarem às oscilações dos níveis de preço. Finalmente, foram feitos os agrupamentos das populações dos estados, que estão disponíveis no Anuário Estatístico do IBGE, a fim de se obter os valores dos PIBs *per capita* estaduais para os anos de 1939 a 2004. Vide o Gráfico A1 que representa a evolução dos PIBs *per capita* de 20 estados brasileiros para os anos compreendidos entre 1939 a 2004.

