

DINÂMICA DAS DESIGUALDADES REGIONAIS NO BRASIL

Alexandre Rands Coelho Barros (UFPE)

Resumo

O artigo critica a hipótese teórica de convergência de renda per capita entre regiões, defendendo a partir de modelo teórico Neoclássico que há tendência a equilíbrio de PIBs per capita relativos com desigualdades regionais, mesmo quando há perfeita arbitragem no mercado de fatores de produção. A partir de críticas aos métodos empíricos com dados longitudinais para testar a convergência de renda per capita entre regiões no Brasil, o artigo utiliza métodos de séries temporais para investigar a existência de tal processo. Conclui que a hipótese de convergência não encontra suporte empírico, apenas tendo havido algumas quebras estruturais no equilíbrio entre 1939 e 2007, período no qual os PIBs per capita relativos do Nordeste e do resto do Brasil estiveram quase sempre em equilíbrio estável.

Abstract

This paper presents a Neoclassical Model of equilibrium between per capita GDP of two regions concluding that there is a tendency to stable regional inequality, differently from what has been argued on the convergence literature. Additionally, it forwards a time series approach to test the hypothesis of convergence in the period between 1939 and 2007 for Northeast and the rest of the country, in contrast to the cross section growth regressions normally used to test this hypothesis. The tests indicate that there was a stable equilibrium, but with some structural breaks within the period, or no convergence of per capita GDP of Northeast and of the rest of Brazil.

1. Introdução

Seguindo os estudos sobre convergência de renda per capita entre países, iniciada por Barro (1991) e estendida para regiões (estados ou municípios) em um mesmo país (Estados Unidos) por Barro e Sala-i-Martin (1992)¹, surgiram alguns estudos no Brasil tentando investigar se havia convergência de PIB per capita entre os estados e até mesmo entre municípios brasileiros.² Esses estudos normalmente concluem que tal convergência pode ser encontrada, mas que ela não é necessariamente convergência absoluta (para a mesma renda per capita), mas sim relativa, pois os diversos países ou regiões poderiam ter níveis de equilíbrio de renda diferentes. Quando a conclusão é para regiões em um mesmo país, ela deveria se aproximar mais da convergência absoluta, pois a mobilidade de fatores de produção e de informação deveria reduzir os limites das disparidades a serem observadas.

As hipóteses de convergência de renda per capita normalmente são obtidas teoricamente a partir de modelos de crescimento econômico de cunho neoclássico, como o conhecido como Modelo de Ramsey.³ Para se obter tal conclusão a partir dele, supõe-se que há apenas um bem na economia e que não haveria fluxo de fatores de produção entre países. Na verdade a conclusão básica do modelo é obtida para uma economia fechada e mostra-se que quanto mais aquém a economia estiver do seu equilíbrio maior tenderá a ser sua taxa de crescimento. Conseqüentemente, quando as economias são mais pobres elas tenderão a estar mais longe de seu equilíbrio de longo prazo e crescerão mais aceleradamente.

Para se obter a conclusão empírica, seja ela geral ou em clubes,⁴ de que há alguma convergência de rendas per capita normalmente se utiliza dados longitudinais para dois ou mais períodos em que as taxas de crescimento das unidades geográficas são regressadas contra vários dos seus atributos individuais, incluindo-se entre eles o PIB per capita do período inicial. Ou seja, a inclusão de mais unidades geográficas aumenta a robustez econométrica dos testes. A inclusão de períodos de tempo mais longos, por sua vez, aumenta a robustez dos testes a partir de sua influência no efeito teórico,⁵ apesar de sob um ponto de vista puramente estatístico não afetar a sua robustez.

Os métodos para se obter as conclusões acima, seja a teórica ou a empírica, apresentam fragilidades importantes. No caso da teórica, quando se utiliza as conclusões para as regiões de um mesmo país, há claramente inadequação das suposições. Nesse caso a migração de fatores de produção tende a acelerar o processo de convergência e a unificar as taxas de retornos deles, principalmente o capital. Conseqüentemente, não haveria tendência a convergência de forma lenta, como é preconizado a partir da utilização de modelo de crescimento para regiões isoladas e sem migração de fatores. As taxas de convergência deveriam ser bem mais rápidas, diferentemente do que tem sido encontrado nos estudos a partir de dados longitudinais com o método de Barro e Sala-i-Martin (1992) e suas variações.⁶

Há também problemas estatísticos associados a esse tipo de estimação com dados longitudinais. Estudo recente de Hauk Jr e Wacziarg (2009) mostrou que os coeficientes estimados nessas regressões tendem a ser viesados, como havia sido apontado por Quah (1996). A intuição porque isso ocorre é trivial. Suponha por exemplo que não haja essa relação de convergência e o verdadeiro valor do coeficiente para o PIB per capita no início do

¹ Ver também Mankiw, Romer e Weil (1992).

² Ver, por exemplo, Azzoni (2000), Azzoni et.al. (2000), Ferreira (1998) e Elléry e Ferreira (1994)..

³ Ver Barro e Sala-i-Martin (1995).

⁴ Para a idéia de convergência em clubes, ver Durlauf e Quah (1999) e Quah (1996) para estudos entre países e Laurini, Andrade e Pereira (2005) para estudo sobre municípios no Brasil.

⁵ Quanto mais longo o período, maior o papel dos ciclos de baixas frequências na variação da variável dependente relevante, que é a taxa média de crescimento do PIB per capita.

⁶ Essa crítica foi previamente apresentada por Glaeser e Gottlieb (2009) ao analisar os estudos de convergência com esse método apresentados para os EUA.

período seja zero. Entretanto, se houver erro de medida, toda vez que o erro for para cima (baixo), a taxa de crescimento será menor (maior). Ou seja, valores erroneamente maiores (menores) no PIB per capita do período inicial levará a taxa de crescimento menor (maior) e a relação de convergência será encontrada econometricamente, mesmo sem existir na realidade.

O presente trabalho introduz abordagens diferentes tanto para a análise de teórica como a empírica. Ele mostra que em um país onde há vários bens e serviços, com tecnologias produtivas que apresentem diferenças nas proporções ótimas de fatores de produção a um mesmo retorno deles, mesmo que haja perfeita migração de fatores de produção entre as regiões, ainda assim não haverá convergência de rendas per capita entre elas. Esse será chamado aqui de Teorema da não-convergência. Além disso, apresenta-se teste empírico da hipótese de convergência utilizando métodos de séries temporais, que evita os problemas existentes nas abordagens com dados longitudinais.

O trabalho está organizado como segue. A próxima seção apresenta o modelo de múltiplos bens e serviços que conclui que não há convergência de rendas per capita entre as regiões. A seção 3 apresenta o método de testar a hipótese de convergência e o aplica a série temporal da proporção do PIB per capita do Nordeste para o do Brasil entre 1939 e 2007. A seção 4 apresenta e discute as principais conclusões do trabalho.

2. Modelo de não convergência

Para concluir pela não convergência inicia-se com um modelo Neoclássico padrão em que há perfeita mobilidade de mercadorias e fatores de produção e arbitragem em todos os mercados de bens e fatores de produção. Por simplicidade assumir-se-á que há duas mercadorias produzidas na economia e que não há comércio internacional.⁷ Além disso, supor-se-á há duas regiões nessa economia. Os dois bens são produzidos de acordo com as seguintes funções de produção.

$$Y_{i1} = A_1 K_{i1}^{\alpha_1} H_{i1}^{\delta_1} L_{i1}^{1-\alpha_1-\delta_1} \quad (1)$$

$$Y_{i2} = A_2 K_{i2}^{\alpha_2} H_{i2}^{\delta_2} L_{i2}^{1-\alpha_2-\delta_2} \quad (2)$$

Onde Y_{i1} e Y_{i2} são as quantidades das mercadorias 1 e 2 produzidas na região i , para $i=N$ ou $i=S$, sendo N e S subscritos para as duas regiões no país, norte e sul, respectivamente. K e H representam os estoques de capital físico e humano, respectivamente e L a quantidade de trabalho utilizada. Os subscritos nesses fatores de produção seguem exatamente a mesma lógica daqueles da produção. O coeficiente de produtividade é representado por A nessas equações. Os seus subscritos 1 e 2 representam as duas mercadorias produzidas, seguindo regra estabelecida para os de Y . As letras gregas são parâmetros, sendo no caso $0 < \alpha < 1$, $0 < \delta < 1$ e $0 < \alpha + \delta < 1$, qualquer que seja o subscrito representando os bens na economia. Também é importante enfatizar que os bens não têm a mesma intensidade de fatores de produção. Assim $\alpha_1 > \alpha_2$ e $\delta_1 > \delta_2$.

Como as funções de produção representadas nas equações (1) e (2) são do tipo Cobb-Douglas, a demonstração a ser realizada será para funções de produção desse tipo, apesar de as conclusões serem válidas para qualquer função de produção $F(.)$ linearmente homogênea

⁷ É possível também se chegar à mesma conclusão em uma economia aberta, mas nesse caso o modelo tem que ser expandido para incluir o comércio internacional. Por simplicidade apenas está se introduzindo a idéia de uma economia fechada para o exterior.

onde $F_i > 0$ (a primeira derivada em relação ao fator de produção i) para qualquer fator de produção i (i=K, H ou L) e $F_{ii} < 0$ (a segunda derivada em relação ao fator de produção i).

A partir dessas funções de produção, é possível se definir as seguintes funções lucros que são maximizadas pelas firmas:

$$\pi_{i1} = P_1 A_1 K_{i1}^{\alpha_1} H_{i1}^{\delta_1} L_{i1}^{1-\alpha_1-\delta_1} - rK_{i1} - wL_{i1} - \rho H_{i1} \quad (3)$$

$$\pi_{i2} = P_2 A_2 K_{i2}^{\alpha_2} H_{i2}^{\delta_2} L_{i2}^{1-\alpha_2-\delta_2} - rK_{i2} - wL_{i2} - \rho H_{i2} \quad (4)$$

Onde P_1 e P_2 são os preços dos bens 1 e 2, respectivamente e w , r e ρ representam os preços dos fatores de produção, trabalho, capital físico e capital humano, respectivamente. Todos esses preços não receberam subscritos de região porque eles são exatamente os mesmos nas duas existentes, o que decorre da arbitragem dos mercados de fatores de produção e a perfeita mobilidade de bens finais entre elas, que implica que os custos de carregamento entre regiões é zero. Assim, qualquer desequilíbrio entre esses preços levaria a fluxo de fatores da região com menor remuneração para a de maior remuneração, tendendo a forçar o retorno do equilíbrio por oferta e demanda. Da mesma forma, se os preços das mercadorias finais não forem os mesmos entre as regiões, alguns indivíduos poderão ganhar um lucro extra arbitrando entre elas. Esse comércio também levaria ao equilíbrio entre os preços por oferta e demanda.

As condições de primeira ordem para a maximização de lucros das firmas acima geram as seguintes equações:

$$\alpha_j P_j A_j \left(\frac{K_{ij}}{L_{ij}} \right)^{\alpha_j-1} \left(\frac{H_{ij}}{L_{ij}} \right)^{\delta_j} = r \quad (5)$$

$$(1 - \alpha_j - \delta_j) P_j A_j \left(\frac{K_{ij}}{L_{ij}} \right)^{\alpha_j} \left(\frac{H_{ij}}{L_{ij}} \right)^{\delta_j} = w \quad (6)$$

$$\delta_j P_j A_j \left(\frac{K_{ij}}{L_{ij}} \right)^{\alpha_j} \left(\frac{H_{ij}}{L_{ij}} \right)^{\delta_j-1} = \rho \quad (7)$$

Cada uma dessas equações existe para cada um dos produtos j em cada uma das regiões i. Assim, são três equações para dois produtos em duas regiões. Ou seja, há doze equações para a condição de primeira ordem da maximização das firmas nas duas regiões somadas. Além dessas doze equações, há também quatro funções de produção, que foram representadas nas equações (1) e (2). Entretanto, o sistema de equilíbrio geral tem doze quantidades de fatores de produção a serem definidas, três fatores em cada um dos dois produtos em cada uma das duas regiões. Além dessas variáveis, há também as quatro quantidades de produtos, Y_{ij} , e cinco preços, P_1 , P_2 , w , r e ρ . Assim, cinco equações mais seriam necessárias para ter uma solução única nesse sistema de equações.

A quantidade de equações adicionais necessárias para identificar o sistema é, contudo, ainda maior, pois as três equações representadas por (5), (6) e (7), quando postas juntas das funções de produção para o produto específico a que se referem, geram uma dependência entre essas elas. Isso decorre do Teorema de Euler. Segundo ele, no caso específico, se $Y = F(K, H, L)$, então $Y = F_K K + F_H H + F_L L$. Por conseqüência, das 16 equações formadas

pelas 12 representadas pelas equações (5), (6) e (7) acima e as quatro funções de produção, representadas pelas equações (1) e (2), apenas 12 são efetivamente independentes.⁸ Assim, seriam necessárias nove equações mais para transformar o sistema de equilíbrio geral formado nessa economia em plenamente identificado.

Três dessas equações adicionais saem da disponibilidade de fatores de produção no país, pois sendo o modelo estático, não há acúmulo desses fatores. São elas:

$$L_{11} + L_{12} + L_{21} + L_{22} = L_T \quad (8)$$

$$K_{11} + K_{12} + K_{21} + K_{22} = K_T \quad (9)$$

$$H_{11} + H_{12} + H_{21} + H_{22} = H_T \quad (10)$$

Onde L_T , K_T e H_T são os totais de trabalho, capital físico e humano, respectivamente. Nas equações acima supõe-se que há um número fixo de firmas engajadas em cada setor e em cada região, sendo esses definidos exogenamente. Esses números, por simplificação foram todos fixados em 1 (um). A suposição de que o número total de firmas é fixo em cada um dos setores não introduz distorção porque as funções de produção individuais têm retorno constante de escala e por tal as empresas podem ter qualquer tamanho, sem que isso altere sua posição de equilíbrio em relação ao sistema de incentivos gerados a partir dos preços prevalecentes. Assim, tanto faria supor que uma firma realiza toda a produção Y_{ij} ou que um número n de firmas o realizam, cada uma delas com uma produção igual a Y_{ij}/n .

As equações (8) a (10) também assumem que os totais de fatores de produção são exógenos nesse modelo, já que sendo ele estático não há previsão para suas acumulações. Conseqüentemente, essas três equações adicionais reduzem os graus de liberdade do sistema para um déficit de apenas seis, de forma que todas as variáveis endógenas sejam identificadas.

Uma equação adicional vem da estrutura da demanda entre os bens 1 e 2, que é definida pelas preferências dos consumidores. Ela pode ser obtida a partir da suposição de que há apenas um consumidor, que se defronta com a seguinte função utilidade:

$$U = C_1^\sigma C_2^{1-\sigma} \quad (11)$$

Onde C_1 e C_2 são o consumo dos bens 1 e 2, respectivamente, U é a utilidade e $0 < \sigma < 1$ é um parâmetro. Essa função utilidade é maximizada considerando uma restrição orçamentária definida como:

$$P_1(Y_{N1} + Y_{S1}) + P_2(Y_{N2} + Y_{S2}) = P_1C_1 + P_2C_2 \quad (12)$$

As condições de primeira ordem para esse problema geram:

$$\frac{\sigma}{1-\sigma} \frac{(Y_{N2} + Y_{S2})}{(Y_{N1} + Y_{S2})} = \frac{P_1}{P_2} \quad (13)$$

⁸ O leitor interessado em uma comprovação mais rigorosa dessa dependência pode tomar logaritmo natural das equações (5) a (7) para um dos bens em uma das regiões e montar um sistema de três equações com os logaritmos naturais de K , H e L . Esse sistema é linear e deveria definir essas três variáveis em função dos logaritmos naturais de P_i , r , w e ρ . Ao calcular o determinante da matriz com os parâmetros para as três variáveis endógenas, poder-se-á se perceber que ele é zero, indicando a dependência das três equações entre si.

Onde se utilizou o fato de que $C_i = (Y_{Ni} + Y_{Si})$ para $i=1$ ou 2 . Ou seja, tudo que é produzido nas duas regiões em cada um dos setores é consumido. Isso ocorre porque o consumidor não poupa já que não existe regra de acúmulo de fatores de produção e não há comércio internacional. A equação (13) é a equação de demanda que será fundamental para definir a proporção dos dois bens que será produzido na economia a partir de seus preços, dadas as preferências dos consumidores.

Uma outra equação é simples e vem da normalização de um dos preços, já que o modelo só é capaz de determinar os preços relativos. A normalização pode ser simples, com digamos $P_1=1$. Ainda assim, o modelo ainda é indeterminado, havendo quatro equações a menos do que variáveis endógenas.

Isso ocorre porque o modelo não é capaz de determinar a distribuição espacial da produção de cada um dos bens entre as regiões. Como não há custo de transporte, há retornos constantes de escala e nenhum fator de produção é fixo, não há diferença em termos de produtividade de qualquer empresa ou dos fatores de produção que ela emprega se a firma da região N for maior ou menor do que a sua contraparte na região S.

Se houver uma redefinição das unidades do bem 2 de forma que em equilíbrio $P_2=1$,⁹ a produção per capita nas duas regiões pode ser definida como:

$$\frac{Y_{i1} + Y_{i2}}{L_{i1} + L_{i2}} = \frac{Y_{i1}}{L_{i1}} \frac{L_{i1}}{L_{i1} + L_{i2}} + \frac{Y_{i2}}{L_{i2}} \frac{L_{i2}}{L_{i1} + L_{i2}} \quad (14)$$

A equação (6) e as funções de produção implicam que:

$$\frac{Y_{N1}}{L_{N1}} = \frac{Y_{S1}}{L_{S1}} = \frac{w}{1 - \alpha_1 - \delta_1} \quad (15')$$

$$\frac{Y_{N2}}{L_{N2}} = \frac{Y_{S2}}{L_{S2}} = \frac{w}{1 - \alpha_2 - \delta_2} \quad (15'')$$

Substituindo essas equações na (14) resulta:

$$\frac{Y_{i1} + Y_{i2}}{L_{i1} + L_{i2}} = \frac{w}{1 - \alpha_1 - \delta_1} \frac{L_{i1}}{L_{i1} + L_{i2}} + \frac{w}{1 - \alpha_2 - \delta_2} \frac{L_{i2}}{L_{i1} + L_{i2}} \quad (14')$$

A equação (14') mostra que se houver a mesma participação da força de trabalho engajada nos dois setores nas duas regiões, a produção per capita será exatamente a mesma. Quando $\alpha_1 = \alpha_2$ e $\delta_1 = \delta_2$, essa equação degenera para o caso em que há apenas uma mercadoria produzida, que implica na igualdade da produção per capita entre as regiões.

Se as hipóteses de que $\alpha_1 > \alpha_2$ e $\delta_1 > \delta_2$ forem mantidas e as duas regiões tiverem proporções diferentes de força de trabalho empregadas nos dois setores, a região que tiver uma maior participação do setor 1 terá também uma renda per capita maior, já que $(1 - \alpha_1 - \delta_1) < (1 - \alpha_2 - \delta_2)$ e conseqüentemente:

$$\frac{w}{1 - \alpha_1 - \delta_1} > \frac{w}{1 - \alpha_2 - \delta_2} \quad (16)$$

⁹ Somente após o equilíbrio essa normalização faz sentido, pois os preços relativos afetam a escolha do consumidor entre os dois bens.

Assim, quando $(1 - \alpha_1 - \delta_1) < (1 - \alpha_2 - \delta_2)$, as duas regiões terão a mesma produção per capita se e somente se:

$$\frac{L_{N1}}{L_{N2}} = \frac{L_{S1}}{L_{S2}} \quad (17)$$

Isso pode ser percebido mais facilmente se a função de produção per capita apresentada na equação (14') for representada graficamente, assim como ela aparece na figura 1. Nela pode se ver que a produção per capita só será igual nas duas regiões quando as tecnologias utilizadas para produção dos dois bens forem as mesmas (α_i e δ_i forem iguais nas duas regiões) e houver arbitragem nos mercados de fatores (w for o mesmo nas duas regiões) se as participações do emprego nos dois setores forem exatamente as mesmas nas duas regiões. Ou seja, as duas regiões terão que estar no mesmo ponto no gráfico da figura 1, o que implica que essas participações serão as mesmas. Somente nesse caso elas terão as mesmas produções per capita, já que a função representada é sempre crescente no intervalo relevante (entre 0 e 1).

Esses primeiros resultados podem ser resumidos como:

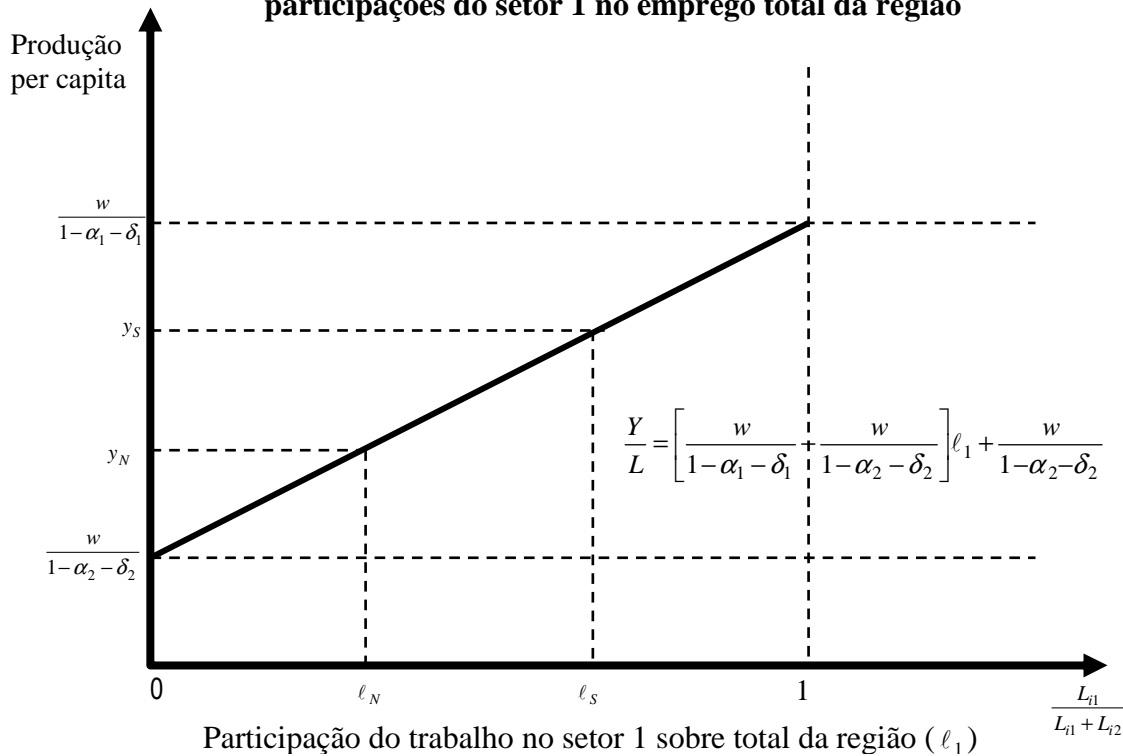
Proposição 1: Em uma economia com duas regiões, que:

- i. produza dois bens;
 - ii. haja perfeita mobilidade de fatores de produção e bens entre elas, com arbitragem perfeita nos seus mercados;
 - iii. todos os bens produzidos sejam comercializáveis;
 - iv. tenha funções de produção tipo Cobb-Douglas para cada um dos bens produzidos;
 - v. haja perfeito fluxo de informações entre elas de forma que as funções de produção para cada um dos produtos seja exatamente igual nas duas regiões;
- As produções per capita nelas serão iguais se e somente se houver a mesma participação do emprego de mão de obra em cada um dos setores no emprego total da região.

A partir desse primeiro resultado, a questão que se coloca em seguida é se existe algum mecanismo intrínseco na economia a partir das forças de mercado que levem essas participações a serem as mesmas. Para averiguar essa possibilidade, vale iniciar em uma economia em que as duas regiões estão em pontos distintos na função representada na figura 1. Ou seja, suponhamos inicialmente que o Norte está no ponto (ℓ_N, y_N) e o Sul no ponto (ℓ_S, y_S) . Ou seja, de acordo com a representação na figura 1, $\ell_S > \ell_N$ e $y_S > y_N$. A diferença nas produções per capita explicam-se por causa das diferenças nas participações do setor 1 no emprego total de cada região. No Sul essa participação é maior do que no Norte nas suposições que nortearam o gráfico apresentado da figura 1.

Apesar das produções per capita distintas, os preços dos fatores de produção, w , r e ρ são os mesmos nas duas regiões. Além disso, os preços dos bens finais também são os mesmos nelas e eles são invariantes a mudanças de proporções na distribuição espacial da produção entre regiões, por haver retorno constante de escala na produção dos dois bens (funções de produção homogêneas de grau 1) e uma determinação da proporção entre eles definida a partir das preferências dos consumidores e dos coeficientes técnicos das funções de produção das duas, que também não se alteram com variações na proporção da produção de cada bem entre as duas regiões.

Figura 1
Representação gráfica da equação (18') para as várias
participações do setor 1 no emprego total da região



Essas igualdades de preços entre as regiões e a perfeita mobilidade de fatores de produção, que podem se deslocar entre elas sem nenhum custo, faz com que um equilíbrio em que a produção per capita entre as regiões seja diferente seja estável e possa perdurar para sempre. Não há nenhuma força econômica gerada endogenamente a partir de arbitragem que venha a alterar esse equilíbrio levando-o a convergir para qualquer outro em que não haja desigualdade da produção per capita regional. Ou seja, quando se põe duas mercadorias com funções de produção distintas, as suposições neoclássicas padrões implicam que a desigualdade regional representa um equilíbrio estável e de longo prazo.

Conseqüentemente, dada uma proporção de demanda dos dois bens definida a partir das preferências dos consumidores, qualquer distribuição entre regiões de cada um dos setores é razoável, desde que a soma da produção das duas regiões assegure que a quantidade ofertada seja igual à quantidade demandada. Dessa forma, pode se pensar, por exemplo, que cada região se especializaria na produção de um dos dois bens. Nesse caso, a desigualdade regional seria a máxima possível nesse modelo. Caso a produção de cada produto em cada uma delas for metade da quantidade determinada pela demanda, haveria igualdade de produção per capita entre as regiões.

Obviamente a conclusão acima pode ser estendida para vários produtos e duas regiões apenas e mesmo para vários produtos e várias regiões, como ocorre no mundo real. Com a diversidade de produtos que existe numa economia real e variações nas funções de produção entre eles a coincidência de produção per capita entre duas regiões qualquer passa a ser uma casualidade de baixa probabilidade de ser encontrada.

Esse modelo mostra que não é preciso violar suposições neoclássicas introduzindo custos de transação e migração, economias de escala e a existência de bens e serviços não transacionáveis, como ocorre em Krugman (1991) e Fujita, Krugman e Venables (1999) ou algumas dessas hipóteses combinadas com a existência de amenidades assimetricamente distribuídas entre regiões, como feito por Glaeser e Gottlieb (2009), Roback (1982) e Rosen (1979), para se explicar a existência de equilíbrio com desigualdades regionais. Esses

modelos, contudo, têm uma grande contribuição a dar à teoria para explicar diferenças em custos de vida e muitas vezes as diferenças nas taxas de crescimento entre as regiões. Ou seja, o que se concluiu aqui não é que eles são irrelevantes teoricamente, mas apenas de que eles não são necessários para explicar as desigualdades regionais.

Esses últimos resultados também podem ser resumidos como:

Proposição 2: Em uma economia dividida em regiões, que

- i. produzam pelo menos dois bens;
 - ii. haja perfeita mobilidade de fatores de produção e bens entre elas, com arbitragem perfeita nos seus mercados;
 - iii. todos os bens produzidos sejam comercializáveis;
 - iv. tenham funções de produção tipo Cobb-Douglas para cada um dos bens produzidos;
 - v. haja perfeito fluxo de informações entre elas de forma que as funções de produção para cada um dos produtos seja exatamente igual nas duas regiões;
- Há um conjunto de possíveis equilíbrios entre as produções per capita de duas regiões, sendo a proporção dessas produções nesse conjunto limitada tanto acima quanto abaixo, mas ainda assim há infinitos pontos que constituem equilíbrio para essa relação. Qualquer um desses equilíbrios é estável no longo prazo, não havendo a geração de nenhum mecanismo intrínseco na economia que leve ao seu deslocamento para outro equilíbrio em que haja maior igualdade das rendas per capita.

2.1. O possível papel do capital humano na convergência das rendas per capita regionais

Na discussão de possíveis fontes endógenas de dinâmica que poderiam emergir a partir de arbitragens entre regiões decorrentes de remunerações diferenciadas para fatores de produção, considerou-se que todos eles teriam suas acumulações definidas pelos seus preços relativos. Entretanto, capital humano, por suas particularidades pode ter outros determinantes de acumulação, além é claro de sua própria remuneração. Ou seja, funções de determinação de sua acumulação contam com outros determinantes que não apenas a remuneração a ser obtida, como o próprio bem estar que pode ser obtido diretamente do status social que ele gera.

A literatura sobre mobilidade social destaca que há uma tendência a pessoas mais educadas terem descendentes também mais educados.¹⁰ Isso decorre de duas possíveis fontes. A primeira é vista como decorrente de relação causal. Os pais podem afetar o desempenho dos filhos por funcionarem domesticamente em padrão de capital humano que motiva os filhos a terem maior refinamento lógico, disciplina e outras características que tendem a facilitar a formação de capital humano. Ainda nesse mesmo conjunto de determinantes causais pode se encontrar também facilitadores genéticos. Uma interpretação alternativa é que a relação emerge de um processo de seleção. Pais com nível de instrução mais elevado tendem a escolher melhores colégios, disponibilizar mais materiais educativos, como livros e jogos, proporcionarem mais atividades que motivem mais estudos, etc. Ou seja, eles não teriam efeito direto na formação da capacidade de aprendizado e desempenho que proporcionaria a maior formação de capital humano, mas apenas proporcionariam um

¹⁰ Ver, por exemplo, Hertz, Jayasundera, Piraino, Selcuk, Smith e Verashchagina (2007).

ambiente de mais oportunidades. Ambas hipóteses podem ter um papel a desempenhar e a maioria dos estudos tende a aceitar que elas estão presentes na realidade.¹¹

Essas fontes de impacto inter-gerações do capital humano gera uma persistência das desigualdades de níveis de educação entre indivíduos, segundo a literatura sobre mobilidade social. Ou seja, ela termina por reforçar a não convergência das rendas per capita regionais, pois se uma região possui uma proporção maior de pessoas mais bem educadas, ela tenderá a perpetuar tal diferença a partir dos comportamentos individuais em relação a acumulação de capital humano. Conseqüentemente, os determinantes não puramente de mercado que contribuem para o acúmulo de capital humano na verdade também tendem a perpetuar as desigualdades regionais e assim reforçam a não convergência de rendas per capita entre regiões, conforme apresentado acima.

3. Testes de convergência a partir de séries temporais

Os testes de convergência de PIB per capita pioneiramente introduzidos por Barro e Sala-i-Martin (1992) para regiões tem como base um estudo econométrico com dados longitudinais (*cross section*) em que cada região representa uma amostra e todas elas têm uma taxa de crescimento definida para um período específico. Conseqüentemente, apenas os valores extremos para o PIB per capita no período são levados em consideração. As informações para todos os anos no meio do período são abandonadas. Nesses estudos, a taxa média de crescimento no período é regressada contra o PIB per capita no período inicial, entre outras variáveis. Se o coeficiente para o PIB per capita no ano inicial for negativa e significativamente diferente de zero, o teste estaria indicando que há convergência de renda per capita. Sofisticações desse teste estenderam a análise empírica a painéis, em que subperíodos são incluídos, em vez de apenas a variação para o período completo e introduziram a idéia de convergência em clubes ou subgrupos dentro do universo total. A essência do método, contudo, continua a mesma.

Os testes desenvolvidos aqui têm uma inversão do papel da quantidade de unidades espaciais e da extensão do prazo nas conclusões a serem geradas. Neles, criar uma estatística para um universo grande de unidades espaciais torna os resultados menos precisos estatisticamente. Entretanto, aumentar a extensão do período para os quais se utiliza dados eleva o poder estatístico do teste. Essa segunda conclusão advém do fato de que cada ano representa uma amostra, não havendo apenas inclusão de extremos, como no caso do teste de Barro e Sala-i-Martin (1992).

Mais precisamente, utilizou-se uma série temporal da proporção do PIB do Nordeste sobre o do Sudeste. Essa série temporal foi submetida a testes de raiz unitária, que serão mais detalhados abaixo, assim como suas relevâncias para o problema em análise. Primeiramente, vale introduzir a equação que serviu de base para os testes, que é:

$$\Delta Y_t = \alpha_0 + \beta t + \alpha_1 Y_{t-1} + \sum_{i=1}^k \delta_i \Delta Y_{t-i} + \sum_{i=1}^m \phi_i D_i + e_t \quad (18)$$

Onde Y_t é a proporção do PIB per capita do Nordeste em relação ao do Sudeste, ambos no ano t . Ou seja,

$$Y_t = \frac{\text{Pib per capita do Nordeste no ano } t}{\text{Pib per capita do Brasil no ano } t}$$

¹¹ Ver, por exemplo, Hertz, Jayasundera, Piraino, Selcuk, Smith e Verashchagina (2007).

As letras gregas representam parâmetros, com exceção de Δ que é o operador de diferenciação. D_i e t representam *dummies* para quebras estruturais em nível e uma tendência determinística, respectivamente. Permitir-se-á nas estimações a seguir a existência de mais de uma quebra estrutural, havendo variação de estimação para estimação. O último termo e_t é um componente aleatório com média zero e variância finita.

A lógica por trás da utilização da equação (18) é simples. Se houver tendência à convergência absoluta de PIB per capita, Y_t tenderá a um (1,0). Isso ocorrerá porque o numerador e o denominador utilizados na sua definição tenderão ao mesmo valor. Se a convergência de renda per capita existir, mas ela não for absoluta, mas apenas relativa, ainda assim Y_t terá uma tendência de longo prazo, que o retirará de um patamar atual inferior para um outro, superior, implicando que o PIB per capita do Nordeste terá uma participação mais elevada em relação à média nacional do que tem hoje, mesmo que não venha a ser igual, ou seja mesmo que Y de equilíbrio seja menor do que um (1,0).

Obviamente tal convergência com Y_t crescendo ocorre porque o Nordeste é a região de menor PIB per capita do país. A hipótese de divergência representaria uma queda de longo prazo em Y_t , algo que normalmente não tem sido cogitado pela literatura e por tal estaria fora de nossas preocupações. Entretanto, também nesse caso a série histórica representada por Y_t também teria uma tendência de longo prazo.

Em todos esses casos de convergência ou divergência, a série representada por Y_t terá uma tendência de longo prazo, podendo ela ser determinística ou estocástica. Em ambos os casos, o valor esperado da variável Y_t estará se movendo no tempo. Somente no caso em que a relação entre o PIB per capita do Nordeste e do Brasil já estejam em equilíbrio, a variável Y_t será estacionária, ou terá média constante ao longo do tempo.

Matematicamente isso significa que se a relação entre os PIBs per capita estiver em equilíbrio $\beta = 0$ e $\alpha_1 < 0$ na equação (18).¹² Além disso, todas as raízes do polinômio gerado na solução de longo prazo pelos termos em primeira diferença estão fora do círculo unitário. Isso implicará que qualquer erro, representado pela seqüência e_t , para $t=0, 1, \dots, n$, onde n é um ano qualquer, será plenamente revertido pelas variações futuras de Y_t ou do próprio e_t .¹³ Além disso, o valor esperado $E(Y_{t+j})=E(Y_{t+i})$ para $i \neq j$, sendo ambos pertencentes ao conjunto dos números naturais e suficientemente grandes para que todos os efeitos temporários de e_t sejam eliminados no futuro. A ausência de uma tendência determinística também é essencial para que essa relação entre Y_{t+j} e Y_{t+i} exista.

Contrariamente, se Y_t não estiver em equilíbrio, $\alpha_1 = 0$ e/ou $\beta \neq 0$. No caso, se houver convergência de PIBs per capita teremos que $\alpha_1 \geq 0$ e/ou $\beta > 0$. Assim, o teste da hipótese de convergência de PIBs per capita, seja ela absoluta ou relativa, a partir de uma análise da série temporal para Y_t é um teste da hipótese conjunta $\beta = 0$ e $\alpha_1 < 0$. Realizar tal teste é o objetivo desse trabalho.

3.1. Estratégia de teste

Três são os problemas para a realização do teste da hipótese de não convergência dos PIBs per capita como ela foi representada na equação (18), incluindo-se as hipóteses de que $\beta = 0$ e $\alpha_1 < 0$. O primeiro aparece na definição mais rigorosa da hipótese nula H_0 , pois a hipótese nula contém uma desigualdade nela. Para que um teste tenha força, normalmente

¹² Dickey e Fuller (1981) e Phillips e Perron (1988) trazem demonstrações mais rigorosas da relação dessas suposições com a hipótese de que não há tendência de longo prazo, seja ela estocástica ou determinística.

¹³ Note que não se supôs que e_t seja inter-temporalmente independente.

busca-se definir como hipótese nula aquela que se quer refutar. No caso em questão, então, a hipótese nula seria $H_0: \alpha_1 = 0$ e/ou $\beta > 0$. Ou seja, a primeira hipótese é pontual e pode ser bem definida, mas a segunda é uma hipótese de desigualdade, o que restringe a possibilidade de montar um teste para ela e mesmo que isso seja feito, a sua potência é substancialmente reduzida.¹⁴

O segundo problema é aquele tratado por Dickey e Fuller (1981) e que se tornou clássico em Macroeconomia.¹⁵ Sob a hipótese nula, $\alpha_1 = 0$, a estatística t-Student que normalmente é utilizada para realizar testes de hipóteses como essa não se distribui como uma distribuição t-Student. Ela tem uma distribuição especial, cujos alguns valores críticos foram computados por aqueles autores. Somente se essa hipótese nula não for verdadeira é que a distribuição para um teste da hipótese $\beta = 0$ terá uma distribuição t-Student padrão.

O terceiro problema é que a hipótese de que $\beta = 0$ e $\alpha_1 < 0$ não implica que α_0 seja constante ao longo do período observado, entre 1939 e 2007. Estudos como o de Barros (2010, cap. 3) sugerem que pode ter havido algumas mudanças estruturais do patamar de Y ao longo da história, mas que essas quebras estruturais na série histórica não implicam em nenhuma convergência de renda per capita, seja ela relativa ou absoluta. Após essas quebras ocorrerem haveria um processo rápido de ajuste, por causa da migração de fatores de produção, e que o equilíbrio passou a girar em torno do novo patamar. Daí a inclusão das *dummies* D_i na equação (18) acima.

A existência desses três problemas demanda uma estratégia especial de teste da hipótese que se deseja, não podendo ser um teste único e direto a partir de estimação da equação (18). Antes de descrever essa estratégia, vale lembrar que a representação da série temporal da proporção do PIB per capita do Nordeste em relação ao do Sudeste que foi apresentada na equação (18) não necessariamente implica que ela seja a verdadeira especificação do processo de geração de dados. A representação correta é de fato desconhecida, podendo ser alguma variação daquela que aparece na equação (18) ou mesmo comportar algumas não-linearidades não contempladas naquela equação que façam com que ela seja apenas uma aproximação da verdade. Esse comentário é pertinente porque na estratégia de testes a ser apresentada abaixo, às vezes algumas restrições à equação (18) serão incluídas de forma arbitrária, desde que possam facilitar os testes e a identificação da melhor representação para a dinâmica temporal de Y_t . O recurso a tal instrumento implica que estimações da equação (18) não serão realizadas sob a suposição de que ela é uma equação estrutural.

Com esses problemas em mente e a incerteza quanto ao verdadeiro processo de geração dos dados (DGP), a estratégia a ser adotada para o teste da hipótese que se pretende contará com uma sucessão de testes que envolverão a estimação de mais de uma especificação da equação (18). A seção seguinte procederá à implementação de tal estratégia, junto com a discussão dos resultados encontrados em cada um dos passos. Não se optou por uma descrição prévia da estratégia de forma genérica nessa seção porque cada passo a ser tomado depende dos resultados encontrados nos passos anteriores. Assim, fazer a discussão de forma genérica, implicaria em discutir muitos possíveis passos que não serão de fato adotados por causa dos resultados previamente encontrados.

3.2. Resultados encontrados

¹⁴ Kwiatkowski, Phillips, Schmidt e Shin (1992) é um dos poucos exemplos de trabalho onde se transformou uma desigualdade em hipótese nula.

¹⁵ Ver por exemplo, Christiano e Eichenbaum (1989) e Campbell e Perron (1991).

Iniciar-se-á todo o processo de teste com a estimação de uma forma simplificada da equação (18), supondo-se que $\beta = 0$ e $\phi_i = 0$ para todos os i 's. Ou seja, iniciar-se-á supondo-se que o processo de geração de dados pode ser representado da forma mais simples possível, não havendo tendência determinística e nenhuma quebra estrutural ao longo do período coberto pelos dados, que é entre 1939 e 2007. Esse passo será utilizado para a realização de um teste da hipótese nula $\alpha_1 = 0$, que poderá identificar se não há uma raiz unitária na variável Y_t sob suposições bem restritivas.

Os resultados dessa primeira estimação aparecem na tabela 1. Utilizou-se o método de Akaike para selecionar o número ótimo de defasagens de ΔY_t , tendo como limite um máximo de sete defasagens. O número selecionado foi 5 defasagens e essa é a equação estimada para a qual os resultado do ADF aparece na tabela 1. A estimativa indica que a hipótese $H_0: \alpha_1 = 0$ é rejeitada a 5%. Isso seria uma indicação de que não há raiz unitária, se esse modelo estimado for o correto e conseqüentemente não haveria tendência estocástica de longo prazo na série, que possa estar indicando a existência de convergência dos PIBs per capita do Nordeste e do Brasil.

Alguns problemas podem aparecer nessa estimação. O primeiro deles é que a determinação do número de defasagens pode ser restritiva, pois, sendo ela desconhecida, a seleção pode não levar a erros e_t que satisfaçam as condições necessárias para que a estatística t-Student estimada tenham a distribuição prevista por Dickey e Fuller (1981).¹⁶

¹⁶ Essas condições são basicamente que os erros sejam independentes intertemporalmente e identicamente distribuídos.

Tabela 1
 Teste de Dickey e Fuller Aumentado e Phillips e Perron para modelo com cinco defasagens e sem tendência determinística

	Estatística estimada	Valores críticos		
		1%	5%	10%
Estatística T para ADF	-3,30258	-3,51	-2,89	-2,58
Phillips e Perron	-2,51	-3,53	-2,90	-2,59

Fonte: Calculo próprio.

Para resolver esse problema, Phillips e Perron (1988) desenvolveram alguns métodos de correção dessa estatística para que haja maior robustez na sua distribuição. O teste com a correção devida também aparece na tabela 1 para o mesmo número de defasagens. A estimativa endógena de alguns parâmetros para correção da estatística T a partir do comportamento dos erros adiciona robustez aos resultados, mas, como seria de se esperar, perde-se em precisão do estimador. Assim a hipótese nula de raiz unitária, $\alpha_1 = 0$, não é mais rejeitada dentro dos valores críticos normalmente utilizados na literatura. Com essa maior robustez não se pode concluir que não há uma tendência estocástica na série, que poderia estar levando à convergência de PIB per capita.

Um outro problema nesse primeiro teste é que se fez uma suposição forte de que alguns dos parâmetros da equação (18) eram zero, o que não necessariamente é verdade. A suposição de que $\beta = 0$ torna o nosso teste para $H_0: \alpha_1 = 0$ mais robusto, pois se ela não for verdadeira os testes apresentados tendem a não rejeitar a hipótese, mesmo que ela seja verdadeira.¹⁷ Da mesma forma, a outra suposição, que $\phi_i = 0$, também reduz a precisão dos testes apresentados, pois as mudanças estruturais levam à não rejeição da hipótese de que há uma raiz unitária, mesmo que isso não seja verdadeiro.¹⁸

Como o teste de Phillips e Perron (1988), ao aumentar a robustez, pôs os resultados originários do ADF sob suspeita, por não ter rejeitado a hipótese de raiz unitária, procedeu-se a um passo mais, que foi incluir a possibilidade de que haja quebras estruturais no comportamento da série temporal. Ou seja, resolveu se relaxar a hipótese de que $\phi_i = 0$. O teste utilizado nesse caso foi sugerido por Lee and Strazicich (2003 e 2004), que permite que haja mais de uma quebra estrutural, mas com quantidade definida *a priori*. Esses testes são apontados na literatura como os mais adequados por terem suas estatísticas não variantes sob as hipóteses nula e alternativa.¹⁹

Os resultados para os testes da hipótese $H_0: \alpha_1 = 0$ sob a hipótese de que há quebras estruturais aparecem na tabela 2. Como pode se ver, os resultados parecem rejeitar a hipótese nula de que $\alpha_1 = 0$, ou que há uma raiz unitária, a 1% no caso de duas e três quebras estruturais²⁰ e a quase 5% quando se admite apenas uma quebra estrutural. Ou seja, os resultados dão suporte à hipótese de que a série é estacionária com algumas mudanças estruturais.

¹⁷ Para uma demonstração dessa consequência, ver Park e Phillips (1988).

¹⁸ Ver Perron (1989 e 2006).

¹⁹ Ver por exemplo Eksi (2009) e Perron (2006).

²⁰ Apesar dos autores Lee and Strazicich (2003 e 2004) não apresentarem valores críticos para o teste com 3 quebras estruturais, a tendência dos valores críticos mostra que provavelmente em cada um dos percentuais mostrados na tabela 2, eles seriam menores do que aqueles para duas quebras. Consequentemente, pode se inferir que nesse caso a rejeição será a valores menores do que 1%.

Tabela 2

Teste da hipótese nula de que há raiz unitária ($\alpha_1 = 0$) quando há quebras estruturais no nível da série.

Número de quebras estruturais	Estatística estimada	Valores críticos		
		1%	5%	10%
1	-3,5376	-4,24	-3,566	-3,211
2	-3,9488	-3,610	-3,047	-2,763
3	-4,1372			

Note: Valores críticos foram estimados para amostra com T=100 e foram extraídos de Lee and Strazicich (2003 e 2004).

Diante das evidências de que não há raiz unitária ou tendência estocástica nos dados, procedeu-se a utilização de métodos tradicionais de regressão para investigar a possibilidade de existência de uma tendência determinística. Para tal mudou-se a equação a ser estimada, que assumiu a forma a seguir:

$$Y_t = \alpha_0 + \beta t + \sum_{i=1}^k \rho_i Y_{t-i} + \sum_{i=1}^m \phi_i D_i + e_t \quad (19)$$

Onde as letras gregas continuam sendo parâmetros e e_t é erro, com média zero e variância finita. Essa equação é a mesma anterior, somente com algumas transformações algébricas. Assim, $\rho_1 = (1 + \alpha_1 + \delta_1)$ e os demais ρ_i são transformações a partir dos δ_i . A única modificação efetiva na especificação da equação (19) para a (18) foi que, estando a variável Y_t em nível, extraiu-se seu logaritmo natural antes de se proceder às estimações.

Livre dos vieses que a existência de uma raiz unitária poderia trazer para a estimação dessa equação por mínimos quadrados ordinários, procedeu-se à estimação das várias versões a serem apresentadas abaixo a partir desse método, mas com correção para heteroscedasticidade e autocorrelação pelo método de Newey e West (1987). Várias defasagens possíveis para esse método foram adotadas com vistas a aumentar a robustez das conclusões.

As primeiras estimações foram sob a suposição de que $\rho_i = 0$ e $\phi_i = 0$ para todo i . Esse modelo simplificado rendeu os resultados apresentados na tabela 3, onde eles aparecem especificados para várias especificações das defasagens possíveis na correção de Newey e West. Em todos os modelos a hipótese nula $\beta_1 = 0$, que implica que não há tendência determinística, não é rejeitada dentro dos níveis de significância tradicionais. A estatística de Durbin-Watson indica que há autocorrelação dos erros, mas as estatísticas t-Student estimadas incorporam a correção delas através do método de Newey e West. Vários limites máximos de defasagens para a representação de médias móveis dos erros foram permitidos, sendo entre 1 e 5 esses limites. Com todas as configurações os resultados são qualitativamente os mesmos. Eles apontam para a não existência de tendência determinística na proporção do PIB do Nordeste em relação ao do Brasil.

Tabela 3
Testes para $\beta_1 = 0$ na equação (19)

Modelo (Defasagens)	Estatística T	Significância
1	1,35220	0,17630999
2	1,14242	0,25328093
3	1,02201	0,30677544
4	0,94661	0,34384009
5	0,89610	0,37019797

Nota: Modelo também inclui constante. Número de defasagens é o permitido na média móvel na correção para heteroscedasticidade e autocorrelação dos erros pelo método de Newey e West (1987). β_1 estimado = 0,000686989. $R^2=0,029$.

A forma específica em que o teste para a existência de uma tendência determinística foi conduzido acima o imputa muito pouca robustez, pela sua sustentação a partir de hipótese de inexistência de erro tipo II, em vez de erro tipo I, que é muito menor. Assim, introduziram-se novas extensões ao modelo com vistas a tentar elevar a robustez dos resultados. A primeira foi a introdução de defasagens para Y_t como especificado na equação (2). Utilizou-se os métodos de Akaike e Schwarz para a seleção do número de defasagens adequado. Ambos os testes indicam que o modelo com uma defasagem é o mais adequado.

O modelo resultante dessa extensão também foi estimado utilizando-se correção pelo método de Newey e West para auto-correlação e heteroscedasticidade. Os resultados encontrados para o mesmo conjunto de defasagens possíveis aparecem especificados na tabela 4. Com essa modificação, todos os testes rejeitam a hipótese nula de que não há uma tendência determinística nos dados. Assim, com essa especificação a hipótese de convergência aparentemente ressurgiu com robustez.

Tabela 4
Testes para $\beta_1 = 0$ na equação (19) com uma defasagem de Y_t incluída ($\rho_1 \neq 0$)

Modelo (Defasagens)	Estatística T	Significância
1	1,91474	0,05552520
2	1,91961	0,05490772
3	1,82643	0,06778574
4	1,81625	0,06933156
5	1,80035	0,07180487

Nota: Modelo também inclui constante e defasagem para Y_t . Número de defasagens é o permitido na média móvel na correção para heteroscedasticidade e autocorrelação dos erros pelo método de Newey e West (1987). β_1 estimado = 0,000280644. $R^2=0,794048$.

O modelo preconizado como adequado nesse trabalho, contudo, não tem especificação como aquela usada para gerar os resultados da tabela 4. Na verdade o que se defende é que é possível identificar algumas quebras estruturais na série Y_t de forma que a sua representação resultante seja tal que não haja uma tendência de longo prazo nela que possa implicar em convergência de PIB per capita.

Uma observação detalhada dessa série identifica algumas quebras estruturais óbvias. Uma primeira e importante é sua mudança de comportamento após 1983. Nos anos

subseqüentes ela ficou estável em torno de um patamar elevado, sendo ele um dos mais elevados desde 1939. Assim, a primeira dummy colocada para identificar uma quebra estrutural tem valor um nos anos subseqüentes a 1983. A inclusão dessa dummy no modelo levou a novo conjunto de estimações, seguindo a lógica anterior. Ou seja, manteve-se uma defasagem de Y_t e utilizou-se o método de Newey e West para corrigir para heteroscedasticidade e auto-correlação dos erros quando se calculou a estatística t-Student. Os resultados aparecem na tabela 5, também para cinco especificações diferentes, uma para cada máximo de defasagem permitida na representação por médias móveis dos erros.

Os resultados indicam que a simples introdução de uma *dummy* para controlar para essa quebra estrutural já é suficiente para eliminar a significância estatística da existência de uma tendência determinística. As estatísticas t-Student para a tendência determinística caem bastante, ficando fora dos limites normalmente aceitos como razoáveis para a rejeição da hipótese nula.

Tabela 5

Testes para $\beta_1 = 0$ na equação (19) com uma defasagem de Y_t incluída ($\rho_1 \neq 0$) e uma dummy para quebra estrutural a partir de 1983.

Modelo (Defasagens)	Estatística T	Significância
1	0,00456	0,99635843
2	0,00435	0,99652916
3	0,00434	0,99653880
4	0,00438	0,99650360
5	0,00442	0,99647693

Nota: Modelo também inclui constante, defasagem para Y_t e dummy com 1,0 a partir de 1983. Número de defasagens é o permitido na média móvel na correção para heteroscedasticidade e autocorrelação dos erros pelo método de Newey e West (1987). β_1 estimado = 0,000002047. $R^2=0,795750$.

A princípio esses resultados poderiam já justificar a hipótese desse trabalho, pois a introdução da primeira quebra estrutural já eliminou a significância da rejeição da hipótese nula de que não há uma tendência determinística de longo prazo que possa implicar em convergência de PIB per capita. Entretanto, os resultados da tabela 5 não são muito melhores do que aqueles da tabela 4, particularmente no que se refere à estatística Durbin-Watson. Essa estatística não é mais relevante para indicar autocorrelação de primeira ordem no modelo estimado, dada sua especificação e mesmo assim já há correção para ela. Contudo, essa estatística pode estar mostrando que há problemas de especificação, que estariam levando a flutuações dos erros que podem estar inclusive levando a um comportamento que não seja integrado de ordem zero, $I(0)$. Assim, optou-se por prosseguir ainda com a melhor especificação do modelo.

Duas novas *dummies* para quebra estrutural foram introduzidas. Uma para o período pós SUDENE efetiva, que começou a ter atuação maior a partir de 1968, e uma outra para o período pós 1954, período que as disparidades regionais se acentuaram e passou a haver o clamor nacional por uma política de desenvolvimento regional. Assim, ambas as *dummies* têm valor 1,0 para períodos posteriores a esses anos, 1954 e 1968, e zero para períodos anteriores. Os resultados aparecem na tabela 6 e continuam não rejeitando a hipótese nula de que não há uma tendência de longo prazo, trazendo mais suporte a nossa hipótese, embora a forma como o teste é especificado resulta em teste com baixa robustez.

Tabela 6

Testes para $\beta_1 = 0$ na equação (19) com uma defasagem de Y_t incluída ($\rho_1 \neq 0$) e três dummies para quebras estruturais a partir de 1954, 1968 e 1983.

Modelo (Defasagens)	Estatística T	Significância
1	0,81646	0,41423506
2	0,83224	0,40527455
3	0,87075	0,38389099
4	0,95533	0,33941050
5	1,02639	0,30470933

Nota: Modelo também inclui constante, defasagem para Y_t e três dummies com 1,0 a partir de 1954, 1968 e 1983, respectivamente. Número de defasagens é o permitido na média móvel na correção para heteroscedasticidade e autocorrelação dos erros pelo método de Newey e West (1987). β_1 estimado = 0,000634799. $R^2=0,825416$.

Uma outra bateria de testes semelhantes aos que foram apresentados para modelo em logaritmo natural de Y_t com uma tendência determinística foi utilizada, apenas substituindo-se a tendência por o logaritmo natural dela. Como Y_t está em logaritmo natural pode ser razoável que a tendência também, o esteja. Por problemas de espaço esses testes não são apresentados. Porém, eles mostram que nesse caso a hipótese desse trabalho seria ainda mais robusta, pois em algumas estimações o coeficiente estimado para o logaritmo da tendência é negativo. Ou seja, a hipótese de que não há tendência determinística é ainda mais forte nessa especificação.

4. Conclusões

As duas principais conclusões desse trabalho desafiam duas hipóteses que foram sub-repeticamente introduzidas na análise da questão regional no Brasil a partir da importação e adaptação inadequada de resultados obtidos para economias espacialmente separadas que operam sob condições bem diferentes do que prevalece na relação entre as regiões brasileiras. A primeira dessas hipóteses é que o equilíbrio hoje existente entre as rendas per capita regionais não estariam em seu equilíbrio teórico e por tal haveria uma tendência de longo prazo à convergência dos PIBs per capita regionais. A segunda conclusão é empírica e defende que essa convergência existe e pode ser encontrada nos dados.

A conclusão teórica tem como base modelo de crescimento econômico neoclássico com apenas um bem produzido na economia e sem haver migração inter-regional de fatores de produção. Obteve-se aqui a conclusão oposta de que as disparidades regionais podem ser elevadas e que tenderiam a cada momento a estar próximas ao seu equilíbrio, desde que a economia funcione de forma que possa ser mais adequadamente representada por um modelo também neoclássico, com mais de um bem produzido na economia e com perfeita mobilidade de fatores de produção, algo que também é distorcivo, mas que seguramente é uma aproximação maior da realidade do que a alternativa de não migração de fatores de produção.

A hipótese empírica de que há convergência de PIBs per capita normalmente é obtida a partir de testes com dados longitudinais em que apenas pontos extremos de PIB per capita são levados em consideração. Esses testes apresentam fragilidades estatísticas já comprovadas em outros trabalhos e sobrevalorizam os períodos extremos, que podem fortuitamente estar em posição de dar suporte à convergência. Aqui se recorreu a métodos de séries temporais em que se busca encontrar suporte à hipótese de convergência no comportamento de série temporal de proporção do PIBs do Nordeste e do Brasil ano a ano. Nesses testes admitiu-se a possibilidade de haver quebras estruturais na série e tanto uma tendência estocástica como uma determinística. A hipótese de que há tendência estocástica foi rejeitada a nível estatístico

padrão e a de que há uma tendência determinística também não recebeu suporte empírico. Essa última, contudo, não tem robustez estatística porque na verdade consiste em uma não rejeição de hipótese de que ela não existe.

Os resultados encontrados nesse trabalho, tanto teoricamente como empiricamente, nos levam a concluir que a proporção de PIBs per capita entre o Nordeste e o Sudeste normalmente encontra-se muito próxima ao equilíbrio e que as disparidades regionais no Brasil verificadas nos últimos 70 anos representam um equilíbrio estrutural, sem que haja nenhuma tendência imposta pelas forças de mercado para que elas sejam paulatinamente eliminadas, mesmo que apenas parcialmente.

Referências

- Azzoni, C. R., “Geografia e convergência de renda entre os estados brasileiros.” In R. Henriques (Org.), *Desigualdade e Pobreza*, Rio de Janeiro: IPEA, 2000.
- Azzoni, C., N. Menezes-Filho, Menezes e Silveira Neto, “Geography and Income Convergence among Brazilian States,” *Research network Working paper*, 3096, Inter-American Development Bank, May, 2000.
- Barro, R., “Economic Growth in a Cross Section of Countries,” *Quarterly Journal of Economics*, 106(2): 407-443, 1991.
- Barro, R. e X. Sala-i-Martin, “Convergence,” *Journal of Political Economy*, 100(2): 223-51, 1992.
- Barro, R. e X. Sala-i-Martin, *Economic Growth*, London: McGraw Hill, 1995.
- Campbell, J. e P. Perron, “Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Root”, in O. Blanchard e S. Fischer (eds.), *NBER Macroeconomics Annual 1991*, Vol. 6, 141-220, Cambridge, MA: National Bureau of Economic Research, 1991.
- Christiano, L e M Eichenbaum, “Unit Roots in Real GNP: do we know and do we care?” *NBER Working Paper*, # 3130, Cambridge MA, Outubro, 1989.
- Dickey, D. A., and W. A. Fuller, “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root” *Econometrica*, 49(4): 1057-1072, 1981.
- Durlauf, S. e D. Quah, “The New Empirics of Economic Growth,” em J. Taylor e M. Woodford, *Handbook of Macroeconomics*, Cap. 4, pp. 235-308, Vol. 1, Part A, Amsterdã: Elsevier, 1999.
- Eksi, O., “Structural Break Estimation: A Survey”, *Working paper*, Universitat Pompeu Fabra, Outubro, 2009.
- Elléry Jr., R. e P. C. Ferreira, “Crescimento econômico e convergência entre as rendas dos estados brasileiros,” *Anais do XVIº Encontro Brasileiro de Econometria (SBE)*, pp. 264-286, 1994.

- Ferreira, A., “Convergence in Brazil: Past and Future”, *Texto para Discussão # 119*, Belo Horizonte: CEDEPLAR, 1998.
- Fujita, M., P. Krugman e A. Venables, *The Spatial Economy*, Cambridge (Mass.): MIT Press, 1999.
- Glaeser, E. and J. Gottlieb, “The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States,” *Journal of Economic Literature*, 47(4): 983-1028, 2009.
- Hauk Jr., W. and R. Wacziarg, “A Monte Carlo study of growth regressions”, *Journal of Economic Growth*, 14(2): 103-147, 2009.
- Hertz, T., T. Jayasundera, P. Piraino, S. Selcuk, N. Smith e A. Verashchagina, “The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends,” *The B.E. Journal of Economic Analysis & Policy*, 7(2): 1-46, 2007.
- Krugman, P., “Increasing Returns and Economic Geography”, *Journal of Political Economy*, 99(3): 483-99, 1991.
- Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin, Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics* 54, 159–178, 1992.
- Laurini, M, E. Andrade e P. Pereira, “Income convergence clubs for Brazilian Municipalities: a non-parametric analysis,” *Applied Economics*, 37, 2099-2118, 2005.
- Lee, J. and M. Strazicich, "Minimum LM Unit Root Test with One Structural Break", Appalachian State University Working Paper, 2004.
- Mankiw, G., D. Romer e D. Weil, “A Contribution to the Empirics of Economic Growth,” *Quarterly Journal of Economics*, 107(2): 407-437, 1992.
- Newey, W. e K. West, “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 55(3): 703-08, 1987.
- Park, J. e P. Phillips, “Statistical Inference in Regressions with Integrated Processes: Part 1”, *Econometric Theory*, 4, 468-497, 1988.
- Perron, P., “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis.” *Econometrica* 57(6): 1361-1401, 1989.
- Perron, P., “Dealing with Structural Breaks,” in T. Mills e K. Patterson, (eds.), *Palgrave Handbook of Econometrics: Econometric Theory*, Vol. 1, London: MacMillan, 2006.
- Phillips, P. and P. Perron, “Testing for a Unit Root in Time Series Regression,” *Biometrika*, 75(2): 335-346, 1988.
- Quah, D., “The Empirics for Economic growth and Convergence”, *European Economic Review*, 40(6): 1353-1375, 1996.

Roback, J., "Wages, Rents and the Quality of Life," *Journal of Political Economy*, 90(6): 1257-78, 1982.

Rosen, S., "Wage-Based Indexes of Urban Quality of Life," in P. Mieszkowski and M. Straszheim (eds.), *Current Issues in Urban Economics*, Baltimore: John Hopkins University Press, 1979.