

# Testes de Robustez: Uma Aplicação para os Determinantes do Crescimento Econômico Estadual Brasileiro Entre 1960 e 2000\*

## RESUMO

---

Este trabalho tem por objetivo determinar quais variáveis possuem uma correlação robusta com as variações do Produto Interno Bruto (PIB) *per capita* dos estados brasileiros entre 1960 e 2000. Com esse intuito, executa dois testes de robustez sugeridos pela literatura. A primeira abordagem é proposta por Levine e Renelt (1992), que usaram o teste chamado *Extreme Bounds Analysis* (EBA) para identificarem variáveis robustas relacionadas com o crescimento econômico. Um enfoque alternativo ao anterior foi proposto por Sala-i-Martin (1997). Em resumo, com base nos testes efetuados, pode-se afirmar que urbanização, mortalidade infantil, fecundidade, pluviometria, carga tributária e migração têm uma correlação robusta com as taxas de crescimento do PIB *per capita* dos estados brasileiros. Além disso, de acordo com os testes, confirma-se a ocorrência de convergência condicional dos PIBs *per capita* estaduais.

## PALAVRAS-CHAVE:

---

Teste de Robustez. Crescimento Econômico. Convergência.

### Guilherme Mendes Resende

- Ph.D. candidate em Geografia Econômica pela *London School of Economics* (LSE);
- Mestre em Economia pelo Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR), da Universidade Federal de Minas Gerais (UFMG);
- Pesquisador do Instituto de Pesquisa Econômica Aplicada (IPEA).

### Lízia de Figueirêdo

- Ph.D. em Economia;
- Professora de Economia do CEDEPLAR/UFMG.

---

\* Este artigo é parte da dissertação de mestrado defendida por Guilherme Mendes Resende para a obtenção do título pela UFMG/CEDEPLAR e orientada por Lízia de Figueirêdo. Em 2005, esta dissertação de mestrado obteve a segunda colocação no IX Prêmio Banco do Nordeste de Economia Regional patrocinado pelo Banco do Nordeste (BNB) e Associação Nacional dos Centros de Pós-Graduação em Economia (ANPEC).

## 1 – INTRODUÇÃO

O tema principal deste estudo é o crescimento econômico, mais especificamente o crescimento do Produto Interno Bruto (PIB) *per capita* dos estados brasileiros. No *mainstream* da teoria econômica, o modelo de crescimento de Solow, de 1956, é o ponto de partida para a discussão dos possíveis fatores que afetam o crescimento do produto *per capita* no longo prazo. Posteriormente, procurou-se uma alternativa para a teoria neoclássica de crescimento – dois artigos seminais são os de Romer (1986) e Lucas (1988) –, incorporando retornos não-decrescentes para o capital e dando explicações endógenas para o crescimento tecnológico. Esses são os chamados modelos de crescimento endógeno.

A partir dos avanços da teoria de crescimento econômico, seguiu-se uma vasta literatura empírica. Um problema enfrentado pelos pesquisadores que fazem trabalhos empíricos sobre o crescimento é determinar de forma precisa quais variáveis pertencem à regressão que realmente explica as taxas de crescimento do PIB *per capita*. Os trabalhos empíricos incluem variáveis que se acredita serem importantes para o crescimento, com base na literatura teórica. Nessa tentativa de encontrar variáveis significativas, chega-se a um dilema. Muitas vezes, verifica-se que a variável  $x_1$  é significativa quando as variáveis  $x_2$  e  $x_3$  estão incluídas no modelo; entretanto, ela se torna insignificante quando  $x_4$  é incluída. Como não se sabe *a priori* as variáveis realmente robustas a serem incluídas, tem-se o problema: quais variáveis possuem uma correlação robusta com o crescimento? Mais especificamente, este estudo procura responder a seguinte questão: quais variáveis possuem uma correlação robusta com as taxas de crescimento do PIB *per capita* dos estados brasileiros?

Para responder a essa pergunta, serão seguidas duas abordagens. A primeira é proposta por Levine e Renelt (1992), que usaram o teste chamado *Extreme Bounds Analysis* (EBA) para identificarem variáveis robustas relacionadas com o crescimento econômico. Um enfoque alternativo ao anterior foi proposto por Sala-i-Martin (1997). O autor argumenta que, em vez de analisar os extremos das estimativas dos

coeficientes de uma variável específica, é necessário fazer a análise de toda a distribuição desses coeficientes.

Vale destacar, neste ponto, a importância dos testes de robustez para a discussão dos determinantes do crescimento econômico<sup>1</sup>. A estrutura básica utilizada em tais testes tem o intuito de lidar com a incerteza de modelos (*model uncertainty*)<sup>2</sup>. Brock e Durlauf (2001), salientam o desafio de se estudar os determinantes do crescimento econômico visto que, em adição à complexidade e à heterogeneidade do objeto de estudo, uma característica chave da literatura empírica sobre o tema encontra-se em sua “*open-endedness*”, isto é, em um nível conceitual, teorias alternativas sugerem determinantes para o crescimento sem, necessariamente, excluir os determinantes propostos pelas outras teorias. As regressões de crescimento entre países são exemplos extremos desse fato, como bem salientam Sala-i-Martin; Doppelhofer e Miller (2004): o número de regressores propostos é maior que o número de países no mundo, sendo impossível, dessa forma, estimar uma regressão com todas as variáveis. Por essa razão, alguns estudos simplesmente “tentam” combinações de variáveis que podem ser importantes para determinar o crescimento econômico e reportam os resultados da especificação preferida. Este tipo de “*data-mining*” pode levar a inferências espúrias. (SALA-I-MARTIN; DOPPELHOFER; MILLER, 2004). É no intuito de tratar desse problema que os testes de robustez se inserem.

Para tratar da fragilidade da inferência econométrica em relação à escolha de modelos em geral, Leamer (1978, 1983, 1985) propõe uma análise de sensibilidade, mais especificamente, o *extreme bounds analysis*, para identificar relações empíricas robustas. No que tange à discussão de crescimento econômico, as análises de robustez de seus determinantes aparecem em três trabalhos seminais publicados na *American Economic Review*.

1 Um dos principais determinantes investigados é a renda no início do período, visto que com tal investigação é possível verificar a ocorrência de convergência/divergência de renda entre regiões.

2 Uma discussão mais específica das várias abordagens que tratam da incerteza de modelos, no contexto das regressões de crescimento, pode ser encontrada em Brock e Durlauf (2001); Temple (2000) e Brock; Durlauf e West (2003).

(LEVINE; RENELT, 1992; SALA-I-MARTIN, 1997; SALA-I-MARTIN; DOPPELHOFFER; MILLER, 2004)<sup>3</sup>. Levine e Renelt (1992) verificam a robustez dos determinantes do crescimento de países aplicando uma versão do chamado *extreme bounds analysis* proposta por Leamer (1983). Os autores concluem que muito poucos regressores passam nos testes. Em resposta a esse resultado pessimista, Sala-i-Martin (1997) emprega um teste menos severo nas variáveis explicativas das regressões de crescimento analisando toda a distribuição das variáveis e verificando as respectivas significâncias estatísticas. Com base em seu método, Sala-i-Martin (1997) identifica um número relativamente grande de variáveis estatisticamente significativas com o crescimento econômico. Em trabalho posterior, Sala-i-Martin; Doppelhofer e Miller (2004) mostraram que a abordagem proposta por Sala-i-Martin (1997) é um caso particular da “média de modelos Bayesianos”, e propuseram, assim, uma técnica de média de modelos, a qual eles chamam de “média Bayesiana de estimativas clássicas” Bayesian Averaging of Classical Estimates (BACE), para determinar a importância de uma variável na regressão de crescimento econômico. Os resultados encontrados utilizando a abordagem bayesiana confirmam aqueles verificados em Sala-i-Martin (1997)<sup>4</sup>. Vale salientar que, Durlauf; Johnson e Temple (2004), ao fazerem um *survey* das técnicas econométricas que são usadas para estudar o crescimento econômico, salientam que uma das razões para ser otimista com a linha de pesquisa empírica sobre crescimento é a potencial contribuição que os desenvolvimentos recentes dos métodos que tratam da incerteza de modelos podem trazer para as discussões de questões sobre crescimento.

3 Outros artigos, sobre testes de robustez aplicados às regressões de crescimento, têm sido publicados recentemente, como por exemplo: Fernández; Ley e Steel (2001); Brock; Durlauf e West (2003); Crespo Cuaresma e Doppelhofer (2007); Ley e Steel (2007) e Eicher; Papageorgiou e Roehn (2007).

4 De fato, os métodos são bastante semelhantes. Uma vantagem do método proposto por Sala-i-Martin; Doppelhofer e Miller (2004) é a possibilidade de se considerar modelos de todos os tamanhos e nenhuma variável é mantida “fixa” e, portanto, “não testada”. Em outras palavras, pode-se dizer que na abordagem de Sala-i-Martin (1997), mantém-se algumas variáveis fixas, que apresentam probabilidade *a priori* igual a 1. Em trabalhos futuros, os autores têm a intenção de executar o teste de robustez para o caso dos estados brasileiros utilizando a abordagem proposta por Sala-i-Martin; Doppelhofer e Miller (2004).

Ademais, é importante esclarecer que os métodos que testam a robustez de variáveis explicativas têm o objetivo principal de identificar tais variáveis. Não têm, portanto, o objetivo de estimar uma equação final com as variáveis classificadas como robustas. Observa-se que, estimando uma única equação (ou um pequeno conjunto de equações) com as variáveis robustas poder-se-ia produzir um resultado incompleto sobre a incerteza dos parâmetros (*parameter uncertainty*) visto que, falhar-se-ia ao não reconhecer que uma única especificação em si mesma é incerta. Bosworth e Collins (2003) são criticados da seguinte forma por Durlauf e Frankel (2003, p. 188) ao tentarem estimar poucas regressões com algumas variáveis consideradas robustas em estudos anteriores:

*Therefore I see no reason to conclude that the choice of variables by Bosworth and Collins makes any progress in addressing the problems of theory uncertainty and specification uncertainty; rather, theirs simply represents one more cross-country growth regression specification to add to the many hundreds that have appeared.*<sup>5</sup>

Na literatura ainda não foi feito nenhum teste de robustez para o caso dos estados brasileiros. Os trabalhos empíricos sobre os determinantes das taxas de crescimento da renda *per capita* dos estados brasileiros mostram uma série de possíveis variáveis explicativas. Por exemplo, pode-se citar densidade populacional, mortalidade infantil, carga tributária e participação do setor industrial no produto de cada estado. Todas essas variáveis em pelo menos uma das regressões estimadas foram significativas. Entretanto, não se sabe como elas se comportam quando é realizado um teste sistemático, fazendo a combinação de uma variável específica com um conjunto de outras variáveis. É possível que muitos dos coeficientes das variáveis testadas tornem-se estatisticamente iguais a zero. Por outro lado, podem existir variáveis que, de acordo com a literatura empírica,

5 Tradução feita pelos autores: “Portanto, eu não vejo nenhuma razão para concluir que as escolhas das variáveis feitas por Bosworth e Collins fazem algum progresso em lidar com os problemas de incerteza de teorias e incerteza de especificação; ao contrário, elas simplesmente representam uma especificação de regressão de crescimento a mais para se adicionar a muitas centenas que têm aparecido.” (DURLAUF; FRANKEL, 2003, p. 188).

não são correlacionadas com as taxas de crescimento do PIB *per capita* e que, após os testes, possa-se verificar que são importantes na sua determinação. Assim, o objetivo deste trabalho é determinar quais variáveis possuem uma correlação robusta com as variações do PIB *per capita* dos estados brasileiros entre 1960 e 2000. Em resumo, com base nos testes efetuados, pode-se afirmar que urbanização, mortalidade infantil, fecundidade, pluviometria, carga tributária e migração têm uma correlação robusta com as taxas de crescimento do PIB *per capita* dos estados brasileiros. Além disso, de acordo com os testes, confirmou-se a ocorrência de convergência condicional dos PIBs *per capita* estaduais.

Este artigo está dividido em cinco seções. Além desta introdução, a segunda seção busca fazer uma revisão da literatura teórica sobre os modelos de crescimento, bem como da literatura empírica. Na terceira, é mostrada a definição do modelo econométrico, a descrição dos dados e o método usado para os testes de robustez. A quarta seção é reservada para a análise dos resultados empíricos verificados a partir deste trabalho; e na última, são feitas as conclusões.

## 2 – REVISÃO DA LITERATURA

Nesta seção, será apresentado o referencial teórico em que se baseia a discussão dos determinantes do crescimento econômico. Para tanto, é analisada uma série de modelos, a começar pelo modelo de Solow (1956). Em seguida, é feita uma revisão da literatura empírica sobre os determinantes do crescimento econômico dos estados brasileiros. Nesta parte do trabalho, buscou-se coletar todas as variáveis explicativas sugeridas pela literatura empírica para os posteriores testes de robustez (os resultados dos testes estão na seção 4).

### 2.1 – Revisão da Literatura Teórica<sup>6</sup>

O modelo de maior importância no debate sobre crescimento econômico no pós-guerra é, certamente, o de Solow (1956). Tal modelo é construído em torno de duas equações, uma função de produção

e uma equação de acumulação de capital. Na sua forma neoclássica, a função de produção apresenta retornos constantes de escala e retornos decrescentes para cada insumo (trabalho e capital). Por sua vez, a equação de acumulação de capital *per capita* descreve a variação no estoque de capital *per capita*, sendo essa variação igual ao investimento<sup>7</sup> bruto *per capita* menos o montante da depreciação *per capita* que ocorre durante o processo produtivo e descontado o crescimento populacional. Portanto, variações no estoque de capital *per capita* resultam da combinação de duas forças opostas: poupança versus depreciação e crescimento populacional. Assim, o modelo sugere que há um único nível de estoque de capital *per capita* no qual a quantidade do investimento iguala o montante da depreciação (mais o crescimento populacional). Se tal estoque existe na economia, o capital *per capita* será apenas alterado pelo crescimento da tecnologia, pois as três forças que atuam para modificá-lo – investimento, depreciação e crescimento populacional – equilibram-se no nível de estoques. Denomina-se este nível de estado estacionário do capital. Encontrando-se a economia em estado estacionário, o investimento tem duas finalidades: uma parcela substitui o capital depreciado e o restante provê os novos habitantes com o volume de capital em estado estacionário.

Conclui-se que, no estado estacionário, as variáveis capital, produto e consumo *per capita* crescem à taxa exógena do progresso tecnológico. E os níveis das variáveis – capital, produto e consumo – crescem à taxa da soma do crescimento populacional e progresso tecnológico. Observe-se que deslocamentos na taxa de poupança, do crescimento populacional ou no nível da função de produção afetam os níveis de longo prazo das variáveis, mas não as taxas de crescimento no estado estacionário. Esses tipos de distúrbios influenciam apenas as taxas de crescimento durante a transição de sua posição inicial para o valor de estado estacionário (efeito de nível). Assim, o nível da renda *per capita* de equilíbrio é determinado positivamente pela taxa de poupança, pelo nível da tecnologia e

6. Esta seção baseia-se em Barro e Sala-i-Martin (2003), Fujita; Krugman e Venables. (1999) e Jones (2000), além de artigos selecionados.

7. A economia é fechada, de modo que a poupança é igual ao investimento, e a única utilização do investimento na economia é a acumulação de capital. Assim, o investimento líquido é função da renda, ou seja, é uma proporção constante da renda descontada a depreciação e o necessário para a manutenção do estoque de capital *per capita* (em decorrência do crescimento populacional).

negativamente pela taxa de depreciação e pela taxa de crescimento populacional.

Logo, se um país aumenta sua taxa de poupança, durante algum tempo (que pode ser longo), o investimento supera a depreciação, gerando taxas de crescimento acima da taxa de progresso tecnológico. O valor dessa taxa será tanto maior quanto menor o estoque de capital desse país. Dito isso, parece que o modelo diz algo sobre a possibilidade de regiões mais pobres alcançarem rapidamente a renda *per capita* das regiões mais ricas. Esse é um caso bem particular de convergência, a convergência beta absoluta<sup>8</sup>. Mais realisticamente, tem-se que permitir uma heterogeneidade entre as economias, ou seja, deve-se relaxar a hipótese de que todas as regiões têm parâmetros idênticos e, portanto, as mesmas posições de estado estacionário. Se o estado estacionário for diferente, então deve-se modificar a análise para introduzir o conceito de convergência beta condicional. A ideia principal é que uma economia crescerá mais rapidamente para seu próprio estado estacionário. Logo, o modelo não prevê convergência em todas as circunstâncias; um país pobre pode crescer mais devagar que um país rico. Nesse sentido, Barro e Sala-i-Martin (2003, p. 48) afirmam: “The neoclassical model does predict that each economy converges to its own steady state and that speed of this convergence relates inversely to the distance from the steady state”.<sup>9</sup>

Tendo analisado o modelo padrão de Solow (1956), podem-se tirar algumas proposições de política a fim de elevar as taxas de crescimento das economias (ainda que temporariamente). Elevação da taxa de poupança, do nível da tecnologia e redução do crescimento populacional seriam intervenções óbvias. Entretanto, tais recomendações não param por

8. Considerando um grupo de regiões (ou países) que são estruturalmente similares (tenham parâmetros idênticos: taxa de poupança, crescimento populacional, depreciação e a mesma função de produção), elas terão os mesmos valores de estado estacionário para o capital ( $k$ ) e produto *per capita* ( $y$ ). Sendo que a única diferença entre as regiões é a quantidade inicial de capital *per capita*, o modelo então sugere que economias menos avançadas – com menores valores de  $k$  e  $y$  – têm maiores taxas de crescimento de  $k$  e  $y$ .

9. Tradução feita pelos autores: “O modelo neoclássico prediz que cada economia converge para seu próprio estado estacionário e que a velocidade desta convergência se relaciona inversamente com a distância do estado estacionário.” (BARRO; SALA-I-MARTIN, 2003, p. 48).

aí, tendo em vista que esse modelo foi “ampliado”, propiciando outras formas de atuação para alavancar o crescimento econômico. Mankiw; Romer e Weil (1992) propuseram um modelo de Solow “ampliado” pela inclusão de uma variável que representa a acumulação de capital humano<sup>10</sup>. Assim, a variável capital, presente no modelo original, poderia ser desmembrada em duas: capital físico e capital humano. Reconhece-se então, que a mão-de-obra de diferentes regiões tem diferentes níveis de instrução e qualificação. A inclusão dessa variável no modelo eleva o grau de explicação dos diferenciais de renda *per capita* entre os países de 59% para 78% no período de 1960-1985. Portanto, mais uma proposição de política é considerada. Uma elevação da escolaridade da população pode elevar (transitoriamente) as taxas de crescimento da renda *per capita* e (permanentemente) o nível da renda *per capita*.

Outra maneira de ampliar o modelo de Solow é analisar a relação entre o estado de saúde<sup>11</sup> médio da economia e o estoque de capital humano. A forma mais tradicional de se incorporar o estado de saúde em modelos de crescimento é considerar o estado de saúde parte do estoque de capital humano, alterando diretamente a capacidade produtiva dos indivíduos. Conclui-se que regiões com maiores níveis de saúde teriam maiores taxas de crescimento da renda *per capita* na transição e maiores níveis de renda *per capita* permanentemente<sup>12</sup>.

O último modelo de Solow (1956) “ampliado” a ser abordado, é o modelo com migração. (BARRO; SALA-I-MARTIN, 2003)<sup>13</sup>. A migração de pessoas é um dos mecanismos para alterações na população e na oferta de trabalho. A diferença entre a mobilidade de capital e a de trabalho é que, enquanto o capital tende a fluir de regiões com baixas taxas de retorno para aquelas com altas taxas de retorno, os trabalhadores tendem a sair de regiões com baixos salários (ou outras características desfavoráveis) para lugares onde o

10. No artigo, os autores usam como *proxy* para capital humano a porcentagem da População Economicamente Ativa matriculada no ensino secundário.

11. Geralmente, como *proxy* para o estado de saúde, usa-se, nos trabalhos empíricos, taxa de mortalidade infantil e expectativa de vida.

12. Ver, por exemplo, Bloom; Canning e Sevilla (2001).

13. O modelo apenas postula uma forma funcional para a função de migração.

salário é alto (ou onde há outros elementos favoráveis). Sabe-se que a mobilidade do capital acelera o processo de convergência das economias em direção ao seu estado estacionário. Conclui-se, também, que a mobilidade da mão-de-obra atua na mesma direção. Primeiramente, é importante ressaltar que a migração difere em alguns aspectos de mudanças no crescimento natural da população, isto é, a diferença entre nascimentos e óbitos, visto que, no caso da migração, ganhos de população para a economia de destino representam perdas para a economia de onde saiu essa população e, diferentemente de indivíduos recém-nascidos, os migrantes podem ter capital humano acumulado.

Pode-se, então, analisar o modelo da seguinte forma: caso se pense no exemplo brasileiro (migração Nordeste–Sudeste, por exemplo), pessoas que migram do Nordeste para o Sudeste levam consigo menos capital humano do que encontram no lugar de destino<sup>14</sup>. Assim, um aumento na quantidade de migrantes para o Sudeste faz que o estado estacionário do capital *per capita* desta região seja menor. Isso representa taxa de crescimento do capital *per capita* negativa, *ceteris paribus*, e, por conseguinte, taxa de crescimento do produto *per capita* também negativa. Por sua vez, está ocorrendo uma emigração de pessoas menos qualificadas do Nordeste, ocasionando uma elevação do capital por trabalhador e da renda *per capita* de equilíbrio, e um aumento na taxa de crescimento de transição do produto *per capita* da região. Logo, pode-se ver que o processo de convergência das rendas *per capita* se acelera com a inclusão de migração no modelo de Solow (1956).

Em resumo, todos os modelos de Solow (1956) “ampliados” que foram discutidos até aqui preveem a convergência  $\beta$  (beta)-condicional, isto é, se as regiões diferem em seus parâmetros que determinam os seus estados estacionários (tais como características estruturais da economia: taxas de poupança, escolaridade, infraestrutura, etc.), cada região deve convergir em direção ao seu próprio nível de estado estacionário de renda *per capita* e não para um nível

comum (como no caso da convergência  $\beta$ -absoluta que assume que as economias são estruturalmente similares). A similaridade em ambos os casos (condicional e absoluta) é que, em equilíbrio, existe convergência nas taxas de crescimento. Entretanto, apenas a convergência  $\beta$ -condicional é compatível com a persistência de grandes diferenças nos níveis de renda *per capita* entre regiões.

Vale ressaltar que, ao menos teoricamente, é possível verificar regiões pobres inicialmente crescerem mais rapidamente que regiões mais ricas sem se observar uma diminuição da dispersão das rendas *per capita* ao longo do tempo. (SALA-I- MARTIN, 1996). Em outras palavras, é possível observar convergência  $\beta$ -condicional sem encontrar evidências de convergência sigma ( $\sigma$ ), isto é, uma diminuição da dispersão dos níveis de renda *per capita* das economias ao longo do tempo. Islam (2003) salienta que apesar das limitações da convergência  $\beta$ , os pesquisadores têm mantido seu interesse neste conceito, em parte, porque os métodos associados à convergência  $\beta$  também dão informações relacionadas aos parâmetros estruturais dos modelos de crescimento econômico, enquanto que os estudos que focam na dispersão da renda (convergência sigma,  $\sigma$ ) geralmente não são capazes de dar tais informações.

Como visto, o modelo de Solow não explica a taxa de crescimento do progresso tecnológico, por isso é chamado de modelo de crescimento exógeno. Porém, a partir da década de 1980, iniciam-se os estudos para dar uma explicação teórica para o crescimento de longo prazo do progresso tecnológico, e, por conseguinte, da taxa de crescimento de longo prazo do produto *per capita*. Tais modelos são chamados de modelos de crescimento endógeno e fazem parte do *hall* das novas teorias do crescimento econômico.

Os artigos seminais que endogenizam o progresso tecnológico foram os de Romer (1986); Lucas (1988) e Romer (1990). Na análise de Romer (1986), os retornos decrescentes para o capital são cessados assumindo-se que a criação de conhecimento é um produto do investimento, ou seja, um aumento no estoque de capital da firma leva a igual aumento no estoque de conhecimento. O modelo baseia-se no fato de que o conhecimento de cada firma é um bem

14. Pode-se verificar tal fato analisando os anos de escolaridade das pessoas que vivem na região Nordeste e daquelas que vivem na região Sudeste. Ver no Anexo B os gráficos das *proxies* para capital humano.

público, podendo outras firmas ter acesso a esse conhecimento sem nenhum custo (dessa forma, o conhecimento transborda, ou seja, ocorre *spillover*, por toda a economia). Assim, uma firma que aumenta seu capital físico aprende simultaneamente como produzir mais eficientemente. Esse efeito positivo na produtividade é chamado de aprender-fazendo (*learning-by-doing*) ou, nesse caso, *learning-by-investing*. É nesse sentido que o capital continua apresentando retornos decrescentes, mas estes são totalmente compensados pela maior produtividade que decorre do avanço tecnológico da economia.

Por sua vez, o modelo de Lucas (1988) acrescenta capital humano ao modelo de crescimento neoclássico (Solow), mas de forma distinta de Mankiw; Romer e Weil (1992). Por capital humano ele quer dizer nível de qualificação dos trabalhadores. A teoria do capital humano trabalha com o fato de que um indivíduo aloca seu tempo em várias atividades no período corrente que afeta sua produtividade no período futuro. No modelo de Lucas, uma política que conduz a um aumento permanente no tempo que as pessoas despendem obtendo qualificações gera um aumento permanente no crescimento do produto por trabalhador. Quando o indivíduo aumenta sua qualificação, há a geração de externalidades para a economia, cujos efeitos sobre a produtividade novamente compensam o declínio da produtividade do capital oriundo dos rendimentos decrescentes. Uma conclusão desse modelo é que uma economia com baixos níveis de capital humano e capital físico ficará permanentemente abaixo de uma economia inicialmente mais bem dotada.

Em outro modelo de Romer (1990), o progresso tecnológico é movido pela Pesquisa e Desenvolvimento (P&D) nos países avançados. A função de produção no modelo de Romer (1990) descreve como o estoque de capital e o trabalho se combinam para gerar o produto usando o estoque de ideias. Assim, quando se admite que as ideias também são um insumo da produção, a função apresenta retornos crescentes. Logo, o desenvolvimento da tecnologia depende das atividades de pesquisa que visam à remuneração do pesquisador pelo seu esforço, conferindo-lhe até mesmo um poder monopolista durante determinado período sobre o fruto de sua criação. Havendo continuamente a possibilidade

de novas criações, a taxa de crescimento pode manter-se para sempre.

Da análise dos três modelos descritos anteriormente, pode-se notar que há duas maneiras básicas de tratar os retornos crescentes à escala que são exigidos caso se deseje tornar endógena a acumulação do conhecimento (progresso tecnológico): concorrência imperfeita ou externalidade. Pode-se abandonar a hipótese da concorrência perfeita e modelar a acumulação de conhecimento como resultado de esforços intencionais de pesquisadores que buscam novas ideias. (ROMER, 1990). Aqui a fonte do crescimento é diretamente o progresso tecnológico – ou pode-se manter a concorrência perfeita e supor que a acumulação de conhecimento é um subproduto acidental, tal como a acumulação de capital. (ROMER, 1986; LUCAS, 1988). Nesse caso, o crescimento volta a ser fruto do investimento na economia.

Outros modelos endógenos dizem respeito aos efeitos da política fiscal, infraestrutura e desigualdade de renda sobre o crescimento econômico<sup>15</sup>. Na literatura de crescimento endógeno existem dois enfoques para se tratar de tais relações. O primeiro enfoque considera a política fiscal como determinada exogenamente. Rebelo (1991) estuda o efeito de longo prazo sobre o crescimento provocado pela introdução de um imposto proporcional sobre a renda. Formula um modelo que estabelece relação de longo prazo linear e negativa entre política tributária e crescimento, uma vez que a taxa reduz o retorno do investimento e a renda dos fatores. Nesse modelo, a função de produção é linear no seu único fator, que pode ser definido como uma combinação de capital físico e humano, sendo que a presença de retornos constantes de escala na acumulação desse fator é o mecanismo capaz de gerar o crescimento de forma endógena. Barro (1990), por sua vez, construiu um modelo em que gastos do governo financiados por taxação entram na função de produção da economia. Assim, de forma análoga ao modelo anterior, o crescimento é gerado endogenamente por meio da presença de retornos constantes de escala na acumulação conjunta de

15. Dentre a vasta variedade de modelos endógenos, estes foram selecionados por terem uma contrapartida no trabalho econométrico aqui realizado.

capital e gastos do governo, fatores produtivos na função de produção. O autor introduz capital público financiado via taxaço sobre a renda derivando uma relação não-linear entre política fiscal e crescimento. O efeito negativo da taxaço sobre a renda dos fatores seria compensado por um efeito positivo de gastos públicos em infraestrutura – financiados pela receita tributária – sobre os mesmos fatores, gerando a não-linearidade da relação.

Enquanto que nos exemplos anteriores a política fiscal é determinada exogenamente, no enfoque discutido a seguir a política fiscal a ser implementada pelo governo é derivada endogenamente no modelo. O grau de concentração de renda é um fator determinante para escolha dessa política. Alesina e Rodrick (1994) constroem um modelo de crescimento e política fiscal endógenos em que a carga tributária é escolhida pelos indivíduos através de voto direto. A política fiscal adotada pelo governo é formalizada a partir do teorema do eleitor mediano. Esse teorema estabelece que, em democracias, quando se escolhe por voto majoritário uma única questão, a opção escolhida pela sociedade será a opção mediana, ou seja, a opção capaz de maximizar a utilidade do eleitor mediano. Assim, quanto mais pobre em capital for o eleitor mediano, maior será a alíquota resultante. Logo, quanto mais concentrada for a distribuição de renda, mais elevada a carga tributária escolhida democraticamente, uma vez que essa incidiria majoritariamente sobre o capital. Distribuição de renda e crescimento apresentam, conseqüentemente, uma relação de longo prazo linear e negativa.

Um campo da economia que desde o início da década de 1990 tem acrescentado elementos à discussão do crescimento econômico e feito vários estudos teóricos e empíricos sobre as estruturas regionais e urbanas é a Nova Geografia Econômica (NGE). Uma das principais contribuições da NGE é mostrar como as interações entre os retornos crescentes no nível da empresa individual, os custos de transporte e a mobilidade de fatores podem fazer a estrutura espacial econômica surgir e se modificar. (FUJITA; KRUGMAN, P.; VENABLES, 1999).

Portanto, é com base nessa variedade de modelos que os pesquisadores fazem seus trabalhos empíricos. Na próxima subseção, ver-se-á o que os

trabalhos empíricos dizem a respeito das variações do Produto Interno Bruto (PIB) *per capita* dos estados brasileiros.

## 2.2 – Revisão da Literatura Empírica

Conforme mencionado na introdução deste estudo, na década de 1990, houve uma diversidade de artigos que buscavam mostrar quais os fatores que determinavam as diferenças de taxas de crescimento *per capita* entre países ou regiões. Entre alguns estudos, pode-se citar Barro (1991); Barro (1996) e Hall e Jones (1996). Todos esses trabalhos, com base em uma *cross-section* de cerca de cem países, identificaram um número grande de variáveis que são correlacionadas com as taxas de crescimento econômico ou com o nível da renda. O método básico consiste em fazer regressões em *cross-section* na forma:

$$\Delta y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon, \quad (1)$$

onde  $\Delta y$  é o vetor das taxas de crescimento econômico e  $x_1, x_2, \dots, x_n$  são os vetores das variáveis explicativas, os quais variam entre pesquisadores e entre os artigos. É bom observar que na literatura podem-se encontrar mais de sessenta variáveis que foram significativamente correlacionadas com as taxas crescimento *per capita* entre os países. (SALA-I-MARTIN, 1997).

A partir da década de 1990, pesquisas que usam tal método começaram a ser desenvolvidas para estudar o crescimento econômico dos estados brasileiros. Segundo Ferreira e Diniz (1994), a partir de meados da década de 1950, começou a ganhar importância, no âmbito dos estudos sobre tendências espaciais ou regionais do desenvolvimento econômico, o debate sobre a relação entre etapas do desenvolvimento e a tendência divergente ou convergente das rendas absoluta e *per capita* de diferentes regiões. No caso brasileiro, como se encontra analisado em ampla literatura, o crescimento econômico, desde meados dos séculos XIX e até recentemente, se fez acompanhar de uma grande concentração das atividades produtivas em um número restrito de estados e regiões, especialmente no Rio de Janeiro e em São Paulo. Embora esse processo tenha provocado forte corrente migratória das regiões de ocupação antiga e de menor dinamismo (do Nordeste brasileiro e de Minas Gerais) para as regiões

dinâmicas, este movimento migratório não foi suficiente para compensar as diferenças nas taxas de crescimento econômico, gerando um processo de divergência interestadual e inter-regional das rendas absolutas e *per capita* e uma marcante desigualdade regional no Brasil.

Na década de 1970 e na primeira metade de 1980, no entanto, observou-se significativa alteração no padrão regional brasileiro, com a unificação do mercado nacional, uma relativa desconcentração geográfica da produção e o melhor desempenho relativo das regiões com menor nível de renda. (FERREIRA; DINIZ, 1994). Como consequência, ocorreu uma inversão da tendência histórica de divergência interestadual e inter-regional das rendas *per capita*. É nesse contexto que os autores apresentam os indicadores quantitativos da evolução da distribuição interestadual e inter-regional da renda no Brasil nesse último período, bem como as principais explicações para esta inversão de tendência.

Analisando o período 1970-1985, os autores mostram uma tendência à convergência sigma<sup>16</sup>. Segundo eles, tal convergência decorreu de um conjunto de fatores relacionados com a ação da política econômica e com a lógica econômica da competição e localização. Entre esses, cabe ressaltar, conforme Ferreira e Diniz (1994): o desenvolvimento e a ampliação da infraestrutura básica; o movimento das fronteiras agrícola e mineral; a ação direta do Estado em termos de investimentos e concessão de subsídios e incentivos fiscais; a crise econômica e política do Rio de Janeiro; a reversão da polarização industrial da área metropolitana de São Paulo; e os movimentos migratórios e as alterações na distribuição regional da população. Por fim, a análise desenvolvida pelos autores sobre o impacto regional das mudanças estruturais que vêm ocorrendo na economia brasileira sugere uma tendência à continuação da convergência entre as rendas *per capita* estaduais.

Ferreira e Ellery Junior (1996), ao analisar o período de 1970-1990, salientaram a existência de um processo de convergência sigma entre o PIB *per capita* dos diversos estados brasileiros. Observaram também a ocorrência de convergência beta, ou seja, estados mais

pobres cresceram mais rapidamente que os mais ricos.

Azzoni (1994), na análise das desigualdades regionais de renda no Brasil, utilizando os dados de 1939 até 1990, calcula indicadores de desigualdade, mostrando sua evolução no tempo e comparando o crescimento da renda *per capita* com os níveis iniciais de renda *per capita* de estados e regiões. Iniciando-se pela convergência beta, os resultados não permitem concluir pela existência da convergência absoluta de rendas *per capita* entre as regiões brasileiras, embora a rejeição dessa ideia tenha ocorrido por decisão situada próximo do limite estatístico de rejeição e aceitação (significância ao nível de 8%). Quanto à convergência sigma, observa-se uma oscilação ao longo do tempo no indicador geral de desigualdades de renda *per capita* entre as regiões no país. Após 1970, as indicações são de diminuição na desigualdade, mas períodos como esses já ocorreram no passado no Brasil, havendo reversão posterior. Sendo assim, é temerário, segundo o autor, aceitar os resultados presentes como indicadores da tendência futura.

Em outro estudo do caso dos estados brasileiros, Ferreira (1995) tenta esclarecer como as variáveis sugeridas pela literatura de crescimento econômico se relacionariam com os níveis de renda *per capita* estaduais. Utilizando os dados de 1980, a renda *per capita* dos estados foi regredida contra as variáveis: taxa estadual de fecundidade, escolaridade da população estadual, parcela da população estadual residente em grandes áreas urbanas, densidade rodoviária, participação da renda do setor agrícola no total da renda interna estadual e participação da renda do setor industrial no total da renda interna estadual. A fim de captar a influência de fatores regionais específicos, foram ainda incluídas *dummies* regionais na regressão.

A equação estimada explicou 88% da variação observada na variável dependente. A hipótese de que todos os coeficientes são, em conjunto, iguais a zero é rejeitada pelo teste F, ao nível de significância de 1%. Apenas duas variáveis explicativas apresentam coeficientes não-significativos estatisticamente: a variável *proxy* para as economias de aglomeração – urbanização (porcentagem da população do estado residente em áreas metropolitanas, aglomerações urbanas ou municípios de mais de 100 mil

16. A redução da dispersão das rendas *per capita* é conhecida na literatura como convergência sigma. (BARRO; SALA-I-MARTIN, 2003).

habitantes) – e a variável *dummy* relativa à região Centro-Oeste. Todas as demais variáveis explicativas apresentaram coeficientes com os sinais esperados e significativamente diferentes de zero, ao nível de significância de 10% ou menor.

O estudo feito por Llédó e Ferreira (1997) investigou as relações de longo prazo existentes entre crescimento econômico, distribuição de renda e política fiscal, obtidas a partir de modelos de crescimento endógeno. Fazendo uma análise *cross-section* para os estados brasileiros entre 1970 e 1990, obteve-se uma relação de longo prazo não-linear em forma de “U-invertido” entre carga tributária global incidente em cada estado e o crescimento da renda *per capita*. A relação linear entre taxa e crescimento não foi aceita. Além disso, o grau de concentração de renda estadual não apresentou um efeito de longo prazo significativo sobre o crescimento.

Azzoni et al. (2000) utilizam os dados das Pesquisas Nacionais de Amostra por Domicílio (Pnads) entre 1981 a 1996, sendo considerados apenas dezoito estados brasileiros devido às limitações de dados. Em todas as especificações, exceto em uma, o coeficiente da renda defasada é negativo e significativo, revelando que os estados com menores níveis de renda inicial tendem a apresentar crescimento mais rápido da renda. Entretanto, o que se verifica é convergência beta condicional, isto é, cada estado converge para sua própria situação de estado estacionário, visto que as diferentes provisões dos estados para educação, infraestrutura e variáveis geográficas impedem a equalização dos níveis de renda entre os estados. Os principais resultados revelam que as variáveis geográficas (clima, latitude, índice pluviométrico) são importantes para a explicação das diferenças nos níveis e no crescimento de renda dos estados brasileiros. Ademais, variáveis de capital humano (nível de escolaridade do chefe da família), bem como de infraestrutura (esgoto, abastecimento de água e coleta de lixo), mortalidade infantil e densidade populacional aparecem de modo geral como significativas.

Bleaney e Figueirêdo (2002) fazem uma estimação (em painel de tempo fixo) das taxas de crescimento da renda total, da população e da taxa de crescimento da renda *per capita* para os estados brasileiros, no período de 1950 até 1995. Os autores testaram se as ideias

da Nova Geografia Econômica ajudam a explicar as desigualdades regionais no Brasil. Encontraram que os determinantes do crescimento da renda *per capita* foram: nível da renda no início do período (convergência beta), participação da indústria e do setor serviços na renda do estado (*proxies* para economias de escala) e custos de transporte.

Figueirêdo; Noronha e Andrade (2003) fizeram um estudo sobre os impactos da saúde sobre o crescimento durante os anos 1990. Concluíram que os fatores que afetaram positivamente o crescimento do produto dos estados brasileiros foram: o maior nível de escolaridade, menor mortalidade infantil, maior densidade demográfica, proximidade a grandes mercados, maior uso de energia elétrica e maior participação da indústria no PIB.

Desde já, vale ressaltar que para se executar os testes de robustez discutidos a seguir é necessária a escolha de variáveis fixas (e, portanto não serão testadas), isto é, variáveis que aparecerão em todas as regressões a fim de minimizar o problema de má especificação do modelo (omissão de variável relevante). A revisão da literatura teórica e empírica nos guiou na escolha do vetor de variáveis consensuais na literatura e, portanto, mantidas fixas. Como discutido nesta seção variáveis de capital humano e variáveis demográficas parecem ser importantes para o crescimento econômico tanto em nível teórico quanto empírico para os estados brasileiros. Neste sentido, foram escolhidas as variáveis taxa de matrícula no ensino primário (*proxy* para capital humano) e crescimento populacional (modelo básico de Solow) para serem mantidas fixas em todos os modelos testados.

Após essa resenha de alguns trabalhos empíricos sobre os determinantes do crescimento dos estados brasileiros, tem-se uma grande diversidade de possíveis variáveis que podem estar afetando o crescimento da renda absoluta e renda *per capita* dos estados. Entretanto, como dito anteriormente, um problema enfrentado por aqueles que fazem trabalhos empíricos sobre o crescimento é que muitas vezes verifica-se, na equação (1), que a variável  $x_1$  é significativa quando as variáveis  $x_2$  e  $x_3$  estão incluídas no modelo; entretanto, ela se torna insignificante quando  $x_4$  é incluída. Como não se sabe *a priori* as variáveis realmente importantes a serem incluídas

no modelo, é difícil determinar quais as variáveis são correlacionadas, de forma robusta, com o crescimento.

Assim, apesar de existirem trabalhos que buscam entender os fatores que influenciam as taxas de crescimento da renda *per capita* dos estados brasileiros, nenhum deles faz testes para verificar se as variáveis explicativas encontradas são estatisticamente robustas, ou seja, são realmente importantes para explicar as variações nas taxas de crescimento da renda *per capita*.

Portanto, a sugestão, pela literatura empírica brasileira, de uma série de variáveis para a explicação das taxas de crescimento estaduais, juntamente com a existência de testes de robustez, torna este trabalho pertinente. Logo, o estudo proposto tentou clarificar a questão de quais variáveis são robustas na determinação das taxas de crescimento da renda *per capita* dos estados brasileiros.

### 3 – MÉTODO

Foram aqui investigados os determinantes das taxas de crescimento do PIB *per capita* dos estados brasileiros. O objetivo foi determinar, dentre uma série de possíveis variáveis, aquelas que realmente influenciaram as taxas de crescimento dos estados brasileiros no período de 1960 a 2000, à luz das teorias aqui expostas. Para tentar resolver essa questão, realizaram-se dois tipos de testes nas variáveis que buscam explicar os diferenciais nas taxas de crescimento do PIB *per capita* dos estados brasileiros. O método do primeiro teste, que será discutido na seção 3.3, é proposta por Levine e Renelt (1992), que usaram o chamado Extreme Bounds Analysis (EBA) para identificar variáveis “robustas” relacionadas com o crescimento econômico. O segundo teste foi proposto por Sala-i-Martin (1997). O autor critica a severidade do teste feito por Levine e Renelt (1992) e argumenta que em vez de se analisar os extremos das estimativas de uma variável específica, é necessário fazer a análise de toda a distribuição da variável explicativa, calculando-se, assim, a média dos coeficientes e dos desvios-padrão. Esse teste será discutido na subseção 3.4.

#### 3.1 – O Modelo Básico

O método básico consistiu em fazer regressões em painel de tempo fixo (tendo em vista que foram controladas a especificação por variáveis de tempo) na forma:

$$y_{it} = \alpha_0 + \gamma_1 + \gamma_2 + \dots + \gamma_{t-1} + \beta_1 x_{1it} + \beta_2 x_{2it} + \dots + \beta_n x_{nit} + \varepsilon_{it}, (2)$$

onde  $y$  é o vetor das taxas de crescimento do PIB *per capita* de cada estado brasileiro em cada período (o subscrito  $i$  se refere a unidade federativa e  $t$  ao ano),  $\gamma_1, \gamma_2, \dots, \gamma_t$  são os vetores das variáveis *dummies* de tempo (começando na década 1960, e com periodicidade de 10 em 10 anos, até a década de 1990) e  $x_{1it}, x_{2it}, \dots, x_{nit}$  são os vetores das variáveis explicativas. Na forma matricial pode-se escrever o modelo desta maneira:

$$y = X\beta + D\gamma + \varepsilon (3)$$

Este modelo é chamado de modelo de Mínimos Quadrados de Variável *Dummy* (MQVD), embora a parte do nome “mínimos quadrados” se refira à técnica geralmente usada para a estimação, e não ao modelo propriamente dito. (GREENE, 2003). Esse modelo é o de regressão clássico, portanto, as suas propriedades e resultados continuam valendo. No caso deste trabalho, a variável *dummy* incluída foi o tempo. É importante observar que se devem incluir  $t-1$  variáveis *dummies*. Assim, um dos efeitos de tempo deve ser retirado para evitar perfeita colinearidade – quando os efeitos de tempo somam um. Ao se incluir variáveis *dummies* como proposto (três colunas de 1's para as décadas de 1970, 1980 e 1990), estas *dummies* vão captar os efeitos marginais de cada década em relação à década de 1960. Visto que tal década não tem uma *dummy* específica, o efeito para ela será captado pelo termo constante.

É importante salientar que o uso de dados em painel em estudos de crescimento econômico permite fazer um controle para a omissão de variáveis que estão presentes ao longo do tempo. Segundo observa Temple (1999), a princípio, se utilizados dados em *cross-section*, para um modelo de crescimento estar completo o nível da tecnologia deveria ser incluído nas regressões. Entretanto, aqui esta variável não pode ser observável e tem de ser omitida. Existindo um problema de omissão de variável (no caso, nível da tecnologia),

os outros parâmetros estimados são viesados se um ou mais regressores são correlacionados com o nível da tecnologia.<sup>17</sup> Na prática, regiões (ou países) que são relativamente menos eficientes provavelmente têm níveis de renda menores, e assim pode-se facilmente pensar em outras correlações com variáveis explicativas. Em resumo, Temple (1999, p. 123) afirma:

*In the absence of a suitable proxy for the level of technology, the only way to obtain consistent estimates of conditional convergence regression is to use panel data methods. Since initial efficiency is an omitted variable that is constant over time, it can be treated as a fixed effect, and the time dimension of a panel used to eliminate its influence.*<sup>18</sup>

Assim, utilizando um modelo econométrico em painel de tempo fixo, o estudo consistiu em testar a “robustez” das variáveis explicativas implementando duas abordagens: *Extreme Bounds Analysis* e o teste da distribuição inteira das variáveis explicativas. Antes de comentar o método de cada teste, será determinada a amostra e descrita a base de dados.

### 3.2 – Dados

A amostra consistiu em 18 estados do Brasil (são excluídos os estados do Norte e o Distrito Federal, e agrega-se Mato Grosso com Mato Grosso do Sul e Goiás com Tocantins, em razão da inexistência de dados ou de sua precariedade), para os períodos de 1960-1970, 1970-1980, 1980-1990 e 1990-2000<sup>19</sup>. No Quadro 1 são apresentadas as variáveis dependente e explicativas, bem como a fonte de dados, sua descrição, o sinal do coeficiente previsto pela teoria e o referencial teórico e/ou empírico<sup>20</sup> em que este trabalho se baseou para incluí-la no modelo

econométrico.

A escolha das variáveis explicativas foi baseada na literatura empírica (e justificada pelos modelos teóricos) discutida na seção 2; entretanto, foi restrita pela disponibilidade de dados. Essa restrição não afetou a obtenção da variável em si, mas a sua periodicidade. Isso quer dizer que, por exemplo, em algum estudo empírico, os dados são quinquenais ou anuais, mas para a aplicação dos testes de robustez foi necessário uniformizar os dados em uma periodicidade decenal. Assim, a periodicidade escolhida levou em conta dois fatores, um prático e outro teórico. O fator prático, como já explicado, refere-se à disponibilidade de dados, visto que para um número enorme de variáveis que este trabalho se propõe a trabalhar, apenas dados decenais estão disponíveis para todas elas. Por sua vez, a questão teórica diz respeito à repercussão dos efeitos das variáveis explicativas sobre o crescimento do PIB estadual *per capita* ao longo do tempo. Um problema que surge quando se utiliza uma periodicidade menor (por exemplo, anual), é que inevitavelmente as variações nas taxas de crescimento vão ser dominadas pelos efeitos dos ciclos econômicos, e não pelas mudanças nas perspectivas de crescimento de longo-prazo: “*Only long time averages of growth rates, compared in the cross-section or using a panel, can allow us to address the determinants of long-run growth with any degree of confidence.*” (TEMPLE, 1999, p. 133)<sup>21</sup>. Nesse sentido, este trabalho utiliza uma periodicidade decenal. Supõe-se, assim, que as variáveis levam até dez anos para influenciarem nas taxas de crescimento estaduais. Esta é uma hipótese um tanto plausível, tendo em vista, por exemplo, que os efeitos da educação não são imediatos.

Ademais, vale observar que apesar de existirem dados disponíveis dos PIB estaduais para o ano de 2006, optou-se por restringir a análise até o ano 2000 basicamente por dois motivos. Primeiramente, o estudo utiliza uma base de dados que abrange um período de 40 anos que, a nosso ver, tem a capacidade de apreender os determinantes de longo prazo discutidos

17. Em um modelo geral do tipo  $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + e_i$ ,

a omissão da variável explicativa relevante,  $X_2$ , para explicar a variável dependente,  $Y$ , torna a estimativa dos coeficientes  $\beta_0$  e  $\beta_1$  tendenciosa e inconsistente, se somente se a variável omitida  $X_2$ , for correlacionada com a variável incluída  $X_1$ .

18. Tradução feita pelos autores: “Na ausência de uma *proxy* conveniente para o nível da tecnologia, a única maneira de se conseguir estimativas consistentes da regressão de convergência condicional é utilizar métodos de dados em painel. Visto que a eficiência no início do período é uma variável omitida que é constante ao longo do tempo, ela pode ser tratada como um efeito fixo, e a dimensão de tempo do painel é utilizada para eliminar sua influência” (TEMPLE, 1999, p. 123).

19. Sendo os dados decenais, a mostra contou com 72 observações.

20. Vale ressaltar que as variáveis citadas não foram necessariamente significativas nos trabalhos relacionados na coluna de referencial empírico (E), apenas foram incluídas nos modelos estimados por cada autor.

21. Tradução feita pelos autores: “Apenas longos períodos de taxas de crescimento, confrontados em dados em corte ou usando-se painel de dados, são capazes de nos permitir tratar dos determinantes de crescimento de longo prazo com algum grau de confiança.” (TEMPLE, 1999, p. 133).

Código	Variável	Fontes dos dados	Descrição	Sinal esperado	Referencial teórico (T) e empírico (E)
Variável dependente	Taxa de crescimento do PIB <i>per capita</i>	Ipeadata	Produto Interno Bruto (PIB) a preços constantes de 2000. Taxas anuais médias de crescimento para os períodos 1960/1970, 1970/1980, 1980/1990 e 1990/2000.		
X1, X2 e X3	<i>Dummies</i> de tempo		<i>Dummies</i> de tempo para as décadas de 1970, 1980 e 1990 respectivamente. X1 é um vetor de com 1's para a década de 1970, X2 é um vetor de com 1's para a década de 1980 e X3 é um vetor de com 1's para a década de 1990.		
X4	Ln (PIB <i>per capita</i> )	Ipeadata	Logaritmo neperiano do PIB a preços constantes. Unidade: R\$ mil do ano 2000. PIB no início do período (1960, 1970, 1980 e 1990).	-	T: (SOLOW, 1956) E: (AZZONI, 1994), (FERREIRA; ELLERY JUNIOR, 1996; AZZONI et al., 2000)
X5	Taxa de crescimento da população	Ipeadata	Taxas anuais médias de crescimento para os períodos 1960/1970, 1970/1980, 1980/1990 e 1990/2000.	-	T: (SOLOW, 1956) E: (BLEANEY; FIGUEIRÉDO, 2002)
X6	Participação do setor industrial no PIB	Ipeadata	Parcela do PIB industrial estadual em relação ao PIB total de cada Estado. PIB no início do período (1960, 1970, 1980 e 1990).	+	T: (FUJITA; KRUGMAN, P.; VENABLES, 1999) E: (FERREIRA, 1995; LLÉDO; FERREIRA, 1997; FIGUEIRÉDO; NORONHA; ANDRADE, 2003)
X7	Participação do setor agropecuário no PIB	Ipeadata	Parcela do PIB agropecuário estadual em relação ao PIB total de cada Estado. PIB no início do período (1960, 1970, 1980 e 1990).	-	T: (FUJITA; KRUGMAN; VENABLES, 1999) E: (FERREIRA, 1995)
X8	Participação do setor de comércio no PIB	Ipeadata	Parcela do PIB do setor de comércio em relação ao PIB total de cada Estado. PIB no início do período (1960, 1970, 1980 e 1990).	+	T: (FUJITA; KRUGMAN; VENABLES, 1999)
Código	Variável	Fontes dos dados	Descrição	Sinal esperado	Referencial teórico (T) e empírico (E)
X9	Participação do setor de serviços no PIB	Ipeadata	Parcela do PIB do setor de serviços estadual em relação ao PIB total de cada Estado. PIB no início do período (1960, 1970, 1980 e 1990).	+	T: (FUJITA; KRUGMAN; VENABLES, 1999) E: (BLEANEY; FIGUEIREDO, 2002)
X10	Densidade populacional ( <i>Proxy</i> para efeitos de congestão)	Ipeadata	Densidade populacional = pop. Total de cada estado/área total do estado. Densidade populacional no início do período (1960, 1970, 1980 e 1990).	-	T: (FUJITA; KRUGMAN; VENABLES, 1999) E: (FIGUEIRÉDO; NORONHA; ANDRADE, 2003; AZZONI et al., 2000)
X11	Taxa de urbanização ( <i>Proxy</i> para economias de aglomeração)	Ipeadata	Taxa de urbanização = pop. Urbana estadual/pop. Total estadual. Taxa de urbanização no início do período (1960, 1970, 1980 e 1990).	+	T: (FUJITA; KRUGMAN; VENABLES, 1999) E: (FERREIRA, 1995)
X12	<i>Proxy</i> 1 para capital humano	Anuário Estatístico do Brasil (AEB)	Taxa de matrícula no ensino primário (número total de matrículas dividido pela população total) no início do período. Para o ano de 1960, o cálculo foi feito por extrapolação.	+	T: (LUCAS, 1988; MANKIW; ROMER; WEIL, 1992) E: (FERREIRA, 1995; AZZONI et al., 2000)
X13	<i>Proxy</i> 2 para capital humano	AEB	Taxa de matrícula no ensino secundário (número total de matrículas dividido pela população total) no início do período. Para o ano de 1960 o cálculo foi feito por extrapolação.	+	T: (LUCAS, 1988; MANKIW; ROMER; WEIL, 1992) E: (FERREIRA, 1995; AZZONI et al., 2000)
X14	<i>Proxy</i> 3 para capital humano	Ipeadata	Número médio de anos de estudo das pessoas de 25 ou + anos de idade, no início do período (1960, 1970, 1980 e 1990). Unidade: Ano. Comentário: Razão entre o somatório do número de anos de estudo completos das pessoas nessa faixa etária e o total dessas pessoas. Para o ano de 1960 foi calculado por meio de extrapolação.	+	T: (LUCAS, 1988; MANKIW; ROMER; WEIL, 1992) E: (FERREIRA, 1995; AZZONI et al., 2000)
X15	<i>Proxy</i> 1 para benefício de transporte (Densidade rodoviária)	AEB	Densidade rodoviária = extensão das estradas dentro do estado dividido por unidade da área daquele estado (x1000), no início do período (1960, 1970, 1980 e 1990).	+	T: (FUJITA; KRUGMAN; VENABLES, 1999) E: (FERREIRA, 1995)

### Quadro 1 – Descrição das Variáveis

Fonte: Elaboração Própria dos autores.

Observação: \*Não existe uma teoria que dê suporte ao sinal esperado do coeficiente.

Código	Variável	Fontes dos dados	Descrição	Sinal esperado	Referencial teórico (T) e empírico (E)
X16	Proxy 2 para benefício de transporte	AEB	Inverso da média ponderada das distâncias entre as capitais dos estados, sendo que o peso é o produto da região de destino com relação ao produto brasileiro no início do período (1960, 1970, 1980 e 1990).	+	T: (FUJITA; KRUGMAN; VENABLES, 1999) E: (BLEANEY; FIGUEIRÉDO, 2002)
X17	Taxa de mortalidade infantil (Proxy1 para o estado de saúde)	Ipeadata	Número de crianças que não irão sobreviver ao primeiro ano de vida em cada mil crianças nascidas vivas [início do período (1960, 1970, 1980 e 1990)]. Unidade: para mil nascidos vivos. Para o ano de 1960, o cálculo foi feito por extrapolação.	-	T: (BLOOM; CANNING; SEVILHA, 2001) E: (FIGUEIRÉDO; NORONHA; ANDARADE, 2003; AZZONI et al., 2000)
X18	Taxa de fecundidade	Ipeadata	Número médio de filhos que uma mulher teria ao terminar o período reprodutivo. Para os anos de 1960, 1970 e 1980 o cálculo foi feito por extrapolação.	-	T: (SOLOW, 1956) E: (FIGUEIRÉDO; NORONHA; ANDARADE, 2003; FERREIRA, 1995)
X19	Expectativa de vida (Proxy2 para o estado de saúde)	Ipeadata	Número de anos de vida que uma pessoa nascida hoje esperaria viver, se todas as taxas de mortalidade por idade se mantivessem idênticas ao que são hoje. Dado para o início das décadas (1960, 1970, 1980 e 1990). Para o ano de 1960 o cálculo foi feito por extrapolação.	+	T: (BLOOM; CANNING; SEVILHA, 2001) E: (AZZONI et al., 2000)
X20	Temperatura	AEB	Temperaturas (°C) médias das capitais para os períodos 1960/1970, 1970/1980, 1980/1990 e 1990/2000.	?*	E: (AZZONI et al., 2000)
Código	Variável	Fontes dos dados	Descrição	Sinal esperado	Referencial teórico (T) e empírico (E)
X21	Latitude	AEB	Foi informado o grau (x°) da latitude das capitais de cada estado.	?*	E: (AZZONI et al., 2000)
X22	Índice pluviométrico	AEB	Altura total (mm) (soma dos doze meses) média (das capitais) para os períodos 1960/1970, 1970/1980, 1980/1990 e 1990/2000.	?*	E: (AZZONI et al., 2000)
X23	Consumo de energia elétrica (Proxy para estoque de capital físico)	AEB	Consumo de energia elétrica (GWh) estadual no início de cada década (1960, 1970, 1980 e 1990).	+	T: (SOLOW, 1956) E: (FIGUEIRÉDO; NORONHA; ANDRADE, 2003)
X24	Proxy1 para infra-estrutura (Esgoto)	Ipeadata	Porcentagem de domicílios com instalações sanitárias rede geral. Dado para o início das décadas (1960, 1970, 1980 e 1991).	+	T: (BARRO, 1990) E: (AZZONI et al., 2000)
X25	Proxy2 para infra-estrutura (Abastecimento de água)	Ipeadata	Porcentagem de domicílios com água canalizada rede geral. Dado para o início das décadas (1960, 1970, 1980 e 1991).	+	T: (BARRO, 1990) E: (AZZONI et al., 2000)
X26	Proxy3 para infra-estrutura (Iluminação elétrica)	Ipeadata	Porcentagem de domicílios com iluminação elétrica. Dado para o início das décadas (1960, 1970, 1980 e 1991).	+	T: (BARRO, 1990)
X27	Proxy para desigualdade na distribuição de renda (Índice L de Theil)	Ipeadata	Índice L de Theil. Comentário: Mede a desigualdade na distribuição de indivíduos segundo a renda domiciliar <i>per capita</i> . É o logaritmo da razão entre as médias aritmética e geométrica das rendas individuais, sendo nulo quando não existir desigualdade de renda entre os indivíduos e tendente ao infinito quando a desigualdade tender ao máximo. Para seu cálculo, excluem-se do universo os indivíduos com renda domiciliar <i>per capita</i> nula. Dado para o início das décadas (1970, 1980 e 1991).	-	T: (ALESINA; RODRICK, 1994) E: (LLÉDO; FERREIRA, 1997)
X28	Carga tributária global	Blanco e Reis (1996)	A carga tributária global é calculada baseada na arrecadação tributária real média do triênio centrada nos anos censitários em relação ao PIB desses anos. É o somatório da arrecadação federal, estadual e municipal em cada estado da Federação. Unidade: %. Dado para o início das décadas (1970, 1980 e 1990).	-	T: (REBELO, 1991), (ALESINA; RODRICK, 1994) E: (LLÉDO; FERREIRA, 1997)
X29	(Carga tributária global) <sup>2</sup>	Blanco e Reis (1996)	A carga tributária global elevada ao quadrado é calculada baseada na arrecadação tributária real média do triênio centrada nos anos censitários em relação ao PIB desses anos. Unidade: %. Dado para o início das décadas (1970, 1980 e 1990).	-	T: (BARRO, 1990) E: (LLÉDO; FERREIRA, 1997)
X30	Taxa líquida de migração	CEDEPLAR	Taxa líquida de migração = (saldo migratório/ população esperada)*100. Saldo migratório = Imigrantes - Emigrantes. Imigrante = indivíduo que há cinco anos não residia na UF em que reside no ano da pesquisa. Emigrante = residia, cinco anos antes, em UF diferente daquela onde atualmente mora. Dado para as décadas de 1960, 1970 e 1980.	+ (NGE) ou - (Solow)	T: (FUJITA; KRUGMAN; VENABLES, 1999) Solow com Migração E: (FERREIRA; DINIZ, 1994)

## Quadro 1 – Descrição das Variáveis

Fonte: Elaboração Própria dos autores.

Observação: \*Não existe uma teoria que dê suporte ao sinal esperado do coeficiente.

nos modelos teóricos de crescimento econômico. A adição de mais seis anos não traria uma modificação substancial nos resultados, que são robustos para 40 anos, isto é, focamos no período 1960 a 2000. Por fim, o artigo faz a análise utilizando dados a cada 10 anos. Assim, a análise econométrica seria bastante prejudicada se fosse incluído um período com um número de anos diferente de 10. Trabalhos futuros poderão atualizar o presente artigo utilizando dados dos PIBs estaduais de 2010 quando estes estiverem disponíveis.

Uma questão importante a ser ressaltada é que a relação entre crescimento econômico e muitas das variáveis explicativas é endógena, ou seja, determinada simultaneamente pelo modelo. Uma forma de controlar a endogeneidade é pela utilização das variáveis explicativas no início de cada década, em vez de utilizar uma média para a década. Assim, isto foi feito quando necessário e explicitado no Quadro 1.

A variável dependente do modelo é a taxa de crescimento médio anual de cada década do Produto Interno Bruto (PIB) *per capita* medido a preços constantes (em mil reais) de 2000. Para esse cálculo, foram usados os dados referentes ao PIB dos estados e de suas respectivas populações, obtidos no Ipeadata. Vale observar as mudanças de metodologias no cálculo do PIB ao longo dos anos. As estimativas feitas pela Fundação Getúlio Vargas (FGV) relacionadas ao período pré-1970 correspondem ao conceito de produto interno líquido a custo de fatores, ou seja, não incluem a depreciação do capital fixo. Por sua vez, as estimativas para o período 1970-1985 referem-se ao conceito de produto interno bruto (PIB) a custo de fatores e foram elaboradas pelo IBGE. A partir de 1986, o conceito do PIB utilizado é o de preços de mercado também estimado pelo IBGE.

A população total estadual é usada para o cálculo das seguintes variáveis explicativas: PIB *per capita* no início do período (para testar a existência de convergência beta condicional do crescimento econômico) e crescimento populacional. Também no Ipeadata são encontradas as participações do setor industrial, agropecuário, de comércio e de serviços no produto de cada estado; a taxa de urbanização

(é a parcela da população urbana em relação à população total); o número médio de anos de estudo das pessoas de 25 ou mais anos de idade (*proxy* 3 para capital humano); a densidade populacional, a taxa de mortalidade infantil; a taxa de fecundidade; a expectativa de vida; e as variáveis relativas à infraestrutura (esgoto, abastecimento de água e iluminação elétrica).

Para a variável benefício de transporte (que é o oposto de custo de transporte, sugerida pela teoria) foram usadas duas *proxies*: extensão das estradas dentro do estado por unidade da área daquele estado (densidade rodoviária) e o inverso da média ponderada das distâncias entre as capitais dos estados, sendo que o peso é o produto da região de destino com relação ao produto brasileiro no início do período. A extensão das estradas e as distâncias entre as capitais dos estados são encontradas no Anuário Estatístico do Brasil (AEB).

As taxas de matrícula no ensino primário e secundário (*proxies* 1 e 2 para capital humano) foram calculadas a partir dos dados de matrícula que constam no AEB. Outras variáveis explicativas encontradas no AEB são: consumo de energia elétrica e variáveis geográficas (temperatura, latitude e índice pluviométrico).

A *proxy* usada para desigualdade de renda foi o índice *L* de Theil, disponível no Ipeadata apenas para as décadas de 1970, 1980 e 1990. O mesmo período foi usado para a carga tributária global (Blanco e Reis, 1996) e carga tributária global ao quadrado. Essa variável é definida como a razão entre a receita advinda da arrecadação de todos os impostos federais, estaduais e municipais e o PIB, em cada estado. A carga tributária ao quadrado tenta captar a relação não-linear entre tributação e crescimento econômico. Logo, os testes de robustez para essas variáveis foram feitos para o período de 1970 a 2000.

A variável taxa líquida de migração é calculada pelo Centro de Desenvolvimento e Planejamento Regional da Universidade Federal de Minas Gerais (Cedeplar/UFMG). Entretanto, para o período 1991-2000 a variável não foi encontrada, fazendo que os testes de robustez para essa variável específica sejam realizados apenas para o

período de 1960 a 1990.

### 3.3 – Teste de Robustez 1: *Extreme Bounds Analysis* (EBA)

Será detalhado a seguir o método utilizado para testar a robustez de cada variável explicativa. O primeiro teste empregado, proposto por Levine e Renelt (1992), é chamado *Extreme Bounds Analysis* (EBA). Embora seja comum falar variável robusta, rigorosamente, no caso, trata-se de uma discussão de estimador robusto. Judge et al (1988, p. 887) definem um estimador robusto como o menos sensível a erros na especificação do modelo: “[...] a *robust estimator is one that produces estimates that are ‘good’ (in some sense) under a wide variety of possible data-generating processes*”<sup>22</sup>. Na literatura empírica de crescimento, na qual o uso de modelos lineares é predominante, a busca por estimadores robustos (em detrimento dos estimadores de MQO) decorre do usual não cumprimento das hipóteses clássicas deste modelo (JUDGE et al., 1988; THOMAS, 1997).

O método aqui utilizado é baseado na proposta de Leamer (1983, 1985). Leamer (1983) advoga a impossibilidade de serem feitas inferências a partir da estimação não-experimental, mas não nega a relevância da mesma desde que o pesquisador submeta sua pesquisa a uma “análise de sensibilidade” (“*sensitivity analyses*”), Leamer (1983, p. 38). A questão central para o autor é que, no tratamento empírico, muitas das hipóteses utilizadas devem ser assumidas como “opiniões” (LEAMER, 1983, p. 37) do pesquisador, as quais devem ser sustentadas (as hipóteses sobre a distribuição amostral e sobre a distribuição dos parâmetros). Leamer (1985) defende um método de análise de sensibilidade que estabeleça um espaço o mais completo possível dos modelos a serem testados, estabelecendo “*priors*” para os mesmos. Quanto maior o espaço e quanto menor o intervalo de resultados gerados por suas unidades (por exemplo, quanto maior o número de combinações lineares das variáveis independentes, numa

regressão linear, e quanto menor a variabilidade do valor do estimador de uma variável), mais confiável seria o estimador em questão. Em Leamer (1985) expõem-se as limitações do próprio método “*extreme bounds analysis*”, dada a restrição implícita das “*priors*”. Prefere o que chama de “*Global sensitivity analysis*”, que expandiria o espaço dos modelos a serem testados e que incluiria também testes sobre variações possíveis nas “*priors*” utilizadas, numa perspectiva puramente Bayesiana.

Levine e Reneult (1992) desenvolvem o teste aqui apresentado tendo por base esta proposta de Leamer (1983). Sala-i-Martin (1997) também o utiliza, mas caminha na direção do tratamento bayesiano em Sala-i-Martin; Doppelhofer e Miller (2004).

Em resumo, o teste EBA funciona da seguinte forma: imagine que se tem um *pool* de  $K$  variáveis, previamente identificadas como relacionadas com as variações das taxas de crescimento do produto sendo o interesse saber se a variável  $z$  é robusta. Então, deve-se estimar a seguinte regressão:

$$\Delta y = \alpha_i + \beta_{wi} w + \beta_{zi} z + \beta_{xi} x_i + \varepsilon, \quad (4)$$

em que  $w$  é um vetor de variáveis que ficam fixas, ou seja, aparecem em todas as regressões,  $z$  é a variável investigada e  $x_i$  é um vetor de três variáveis não-fixas entre o *pool* de  $K$  variáveis disponíveis. Então, é necessário estimar esse modelo para todas as combinações possíveis de  $x_i$ . É definido o limite do extremo superior como o maior valor de  $\beta_{zi}$  somado a dois desvios-padrão; e definido como limite do extremo inferior o menor valor de  $\beta_{zi}$  subtraído de dois desvios-padrão. Feito isso, o teste EBA para a variável  $z$  diz que se o limite do extremo inferior é negativo e o limite do extremo superior é positivo, então a variável  $z$  é “frágil”. Para ser considerada robusta, a variável  $z$  deve ter o mesmo sinal nos dois limites e ainda permanecer estatisticamente significativa em todas as regressões.

No presente trabalho foram feitas regressões em painel de tempo fixo. Neste caso, a regressão (4) do teste EBA passou a ter a seguinte forma:

$$y_{it} = \alpha_0 + \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \beta_w w_{it} + \beta_z z_{it} + \beta_x x_{it} + \varepsilon_{it}, \quad (5)$$

22 Tradução feita pelos autores: “[...] um estimador robusto é aquele que produz estimativas que são ‘boas’ (em algum sentido) sobre uma ampla variedade de possíveis processos de geração de dados.”

em que as *dummies* de tempo e  $w$ , são os vetores de variáveis que ficam fixas<sup>23</sup>, ou seja, aparecem em todas as regressões,  $z$  é a variável investigada e  $x$  é um vetor de três variáveis não-fixas entre um *pool* das  $K$  variáveis disponíveis, citadas anteriormente. Então, é necessário estimar esse modelo para todas as combinações possíveis de  $x$ , e identificar o maior e o menor valor para o coeficiente da variável investigada,  $\beta_z$ , o qual não pode ser rejeitado ao nível de significância de 5%. É definido o limite do extremo superior como o maior valor de  $\beta_z$  somado a dois desvios-padrão, e definido como limite do extremo inferior o menor valor de  $\beta_z$  subtraído de dois desvios-padrão. Assim, se  $\beta_z$  permanecer significativo e com o mesmo sinal nos limites extremos, a variável estudada será robusta. E se o coeficiente não permanecer significativo em todas as regressões ou se o coeficiente mudar de sinal nos limites extremos, então o resultado para a variável estudada será frágil. Tendo em vista os problemas de multicolinearidade entre as variáveis, foram feitas duas restrições ao teste EBA. Além de limitar o vetor  $x$  a combinações de três variáveis, também foram excluídas do *pool* de variáveis aquelas que podiam estar medindo o mesmo fenômeno da variável de teste e que tinham uma correlação acima de 0,70<sup>24</sup>. A matriz de correlação encontra-se no Anexo A.

### 3.4 – Teste de Robustez 2: Teste da Distribuição Inteira das Variáveis Explicativas

O segundo método foi proposto por Sala-i-Martin (1997). O teste consistiu em fazer regressões da mesma forma que em (5). Porém, em vez de verificar os limites extremos, analisamos a distribuição dos coeficientes como um todo. Em seu artigo de 1997, Sala-i-Martin trabalha com duas hipóteses. Primeiramente, supõe que os parâmetros  $\beta_z$  são distribuídos normalmente. Segundo, supõe que os parâmetros não são distribuídos normalmente. No caso

deste trabalho, usou-se hipótese de normalidade, tendo em vista que em seu artigo o autor conclui que o coeficiente de correlação entre os dois casos (normal e não normal) foi de 0,98. Assim, pelo fato de tais casos serem bastante parecidos, adotou-se apenas a hipótese de que os parâmetros são distribuídos normalmente<sup>25</sup>. Foi necessário, então, calcular a média e o desvio-padrão dessa distribuição. Calculou-se um  $\beta_z$  ( $z$  é a variável a ser testada) para todas as combinações possíveis de  $x$  (não foram feitas restrições, como no teste EBA), ou seja, foram estimados  $M$  modelos<sup>26</sup>. Com todos os coeficientes  $\beta_z$ , construiu-se a estimativa média,  $\hat{\beta}_z$ , como sendo:

$$\hat{\beta}_z = \frac{\sum_{j=1}^M \beta_z}{M} \quad (6)$$

Além disso, calculou-se a variância média com base nas  $M$  variâncias estimadas:

$$\hat{\sigma}_z^2 = \frac{\sum_{j=1}^M \sigma_z^2}{M} \quad (7)$$

E, com base na tabela da Distribuição *t* de Student, utilizando um nível de 5% de significância, pode-se analisar a variável testada. Se a estatística *t* calculada para determinada variável explicativa foi maior que o valor crítico de *t* (ao nível de 5% de significância), então a variável foi considerada robusta. Caso contrário, foi denominada como sendo não robusta.

Em síntese, é com base nesses dois testes de robustez (teste EBA e análise da distribuição inteira) que foram determinadas as variáveis explicativas correlacionadas de forma robusta com as taxas de crescimento do PIB *per capita* dos estados brasileiros. Esses resultados estão na próxima seção.

23. No caso deste estudo, o vetor  $w$  é composto por duas variáveis: crescimento populacional e taxa de matrícula no ensino primário (*proxy*1 para capital humano). Essas variáveis são mantidas fixas e, portanto, não testadas, visto que são consensuais na literatura (tanto teórica quanto empírica).

24. Este número de corte foi proposto pelos autores, tendo como objetivo minimizar a ocorrência de multicolinearidade. Sabe-se que no caso de uma forte colinearidade entre  $X_{1i}$  e  $X_{2i}$  (em um modelo geral do tipo  $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + e_i$ ), tende-se a aceitar a hipótese nula de que  $\beta_2 = 0$ , pois a estatística *t* é subestimada.

25. The correlation between these two columns (normal and not normal) is 0,98, which can be interpreted as an indication that the density function of the estimates of  $\beta_z$  is fairly close to normal. (SALA-I-MARTIN, 1997, p. 10).

26. Por exemplo, se se tem um pool de 24 variáveis, serão estimados  $C_3^{24} = \frac{24!}{(24-3)!} = 2024$  modelos para cada variável testada, totalizando 48.576 regressões.

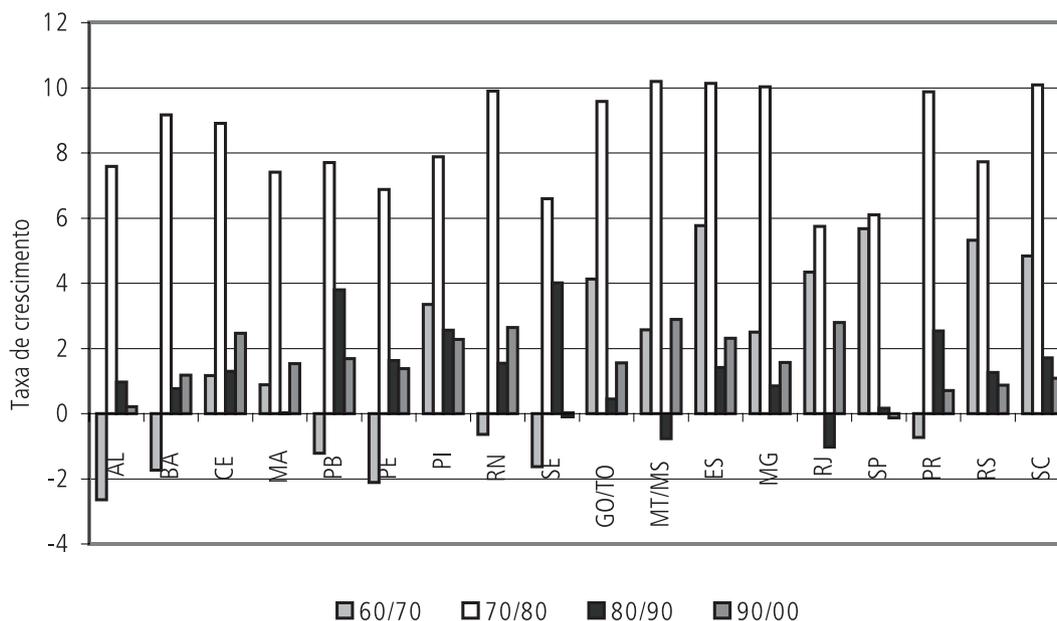
## 4 – ANÁLISES E RESULTADOS

O Brasil caracteriza-se por ser um país em que as disparidades econômicas e sociais entre seus estados são enormes. Entre essas disparidades, também se incluem as diferentes performances de crescimento observadas ao longo das quatro últimas décadas. Da análise do Gráfico 1, pode-se observar uma dispersão muito grande nas taxas de crescimento econômico entre as décadas e entre os estados brasileiros. Na década de 1960 observa-se um crescimento maior dos estados das regiões Sul e Sudeste. Por outro lado, vê-se um péssimo desempenho da região Nordeste. Entretanto, a partir da década de 1970 até a década de 1980, vê-se um melhor desempenho relativo das regiões com menor nível de renda (Nordeste e Centro-Oeste). Já na década de 1990, tem-se um baixo e uniforme crescimento das regiões. É justamente a explicação desses diferenciais de crescimento econômico dos estados ao longo das décadas que este estudo tentou clarificar.

Uma das primeiras decisões para estimar as regressões e fazer os testes de robustez foi escolher as duas variáveis que ficariam fixas em

todas as estimativas. Para isso, necessitava-se de que tais variáveis atendessem a duas propriedades. Primeiro, as variáveis deveriam ter uma justificativa teórica. Segundo, deveriam ser de alguma forma consideradas robustas, de maneira que fossem sistematicamente correlacionadas com a variável dependente em pesquisas anteriores. Cabe mencionar que, além de atenderem a essas duas propriedades, as variáveis escolhidas apresentaram correlações baixas com aquelas do *pool* de variáveis que foram testadas. Este ponto é importante, visto que, assim, minimiza-se o problema de multicolinearidade. É com esses argumentos que a taxa de crescimento populacional (X5) e a taxa de matrícula no ensino primário (X12) sempre estiveram presentes nas regressões, juntamente com as variáveis *dummies* de tempo (X1, X2 e X3). Estas últimas também sempre fizeram parte dos modelos estimados, tendo em vista o tipo de modelo usado, qual seja, painel de efeitos fixos.

Em resumo, tem-se um total de 30 variáveis. Foram usadas cinco delas como variáveis fixas, como explicitado anteriormente. Então, restou um *pool* de 25 variáveis a testar. De acordo com o teste 2 (Sala-i-Martin, 1997), para cada uma, de um



**Gráfico 1 – Taxa Anual Média de Crescimento do PIB Per capita Estadual – 1960/2000 (Em %)**

Fonte: Elaboração Própria dos Autores.

**Tabela 1 – Resultados dos Testes de Robustez – (Variável Dependente = Taxa de Crescimento do PIB *Per capita* dos Estados)**

Continua

Código	Variável	Teste 1					Frágil/ Robusta	Teste 2			Não Robusta/ Robusta
			$\beta$	Desvio- padrão	Teste t	R <sup>2</sup>		$\beta$ (médio)	Desvio- padrão (médio)	Teste t (% t >  2 )*	
X4	Ln (PIB <i>per capita</i> )	Superior:	0,544	0,427	1,273	0,849	Frágil	-1,476	0,736	-2,005 (57%)	Robusta
		Inferior:	-3,332	0,682	-4,888	0,880					
X6	Participação do setor industrial no PIB	Superior:	2,230	2,630	0,848	0,847	Frágil	-2,532	3,145	-0,832 (6%)	Não Robusta
		Inferior:	-4,89	2,808	-1,741	0,831					
X7	Participação do setor agropecuário no PIB	Superior:	4,441	2,551	1,741	0,838	Frágil	-3,308	2,968	-1,115 (25%)	Não Robusta
		Inferior:	-4,623	2,644	-1,748	0,858					
X8	Participação do setor de comércio no PIB	Superior:	8,468	7,639	1,174	0,869	Frágil	-0,331	8,514	-0,039 (0,1%)	Não Robusta
		Inferior:	-19,852	9,154	-2,169	0,853					
X9	Participação do setor de serviços no PIB	Superior:	15,071	3,049	4,943	0,886	Frágil	5,588	2,832	1,973 (47%)	Não Robusta
		Inferior:	-1,689	2,580	-0,655	0,861					
X10	Densidade populacional	Superior:	0,003	0,005	0,638	0,857	Frágil	-0,007	0,005	-1,416 (23%)	Não Robusta
		Inferior:	-0,021	0,006	-3,752	0,858					
X11	Taxa de urbanização	Superior:	-2,482	1,758	-1,412	0,860	Frágil	-5,814	2,271	-2,560 (72%)	Robusta
		Inferior:	-6,140	1,692	-3,630	0,859					
X13	Proxy2 para capital humano	Superior:	37,39	56,20	0,665	0,846	Frágil	-100,987	60,506	-1,669 (35%)	Não Robusta
		Inferior:	-173,13	47,34	-3,657	0,869					
X14	Proxy3 para capital humano	Superior:	0,632	0,274	2,309	0,858	Frágil	0,08	0,45	0,178 (21%)	Não Robusta
		Inferior:	-1,433	0,382	-3,751	0,865					
X15	Proxy1 para benefício de transporte	Superior:	0,003	0,001	2,232	0,879	Frágil	-0,0003	0,001	-0,197 (0,5%)	Não Robusta
		Inferior:	-0,002	0,001	-1,820	0,848					
X16	Proxy2 para benefício de transporte	Superior:	232083,1	81563,2	2,845	0,846	Frágil	54240,5	70271,9	0,772 (14%)	Não Robusta
		Inferior:	-165956,6	60385,7	-2,748	0,862					
X17	Taxa de mortalidade infantil	Superior:	-0,006	0,005	-1,291	0,858	Frágil	-0,018	0,005	-3,473 (97%)	Robusta
		Inferior:	-0,027	0,007	-3,649	0,857					

Fonte: Elaboração Própria dos Autores.

Observação: \*Porcentagem das regressões nas quais o coeficiente apresenta uma estatística *t-student* maior que |2|.

**Tabela 1 – Resultados dos Testes de Robustez - (Variável Dependente = Taxa de Crescimento do PIB Per capita dos Estados)**

Código	Variável	Teste 1					Frágil/ Robusta	Teste 2			Conclusão
			$\beta$	Desvio-padrão	Teste t	R <sup>2</sup>		$\beta$ (médio)	Desvio-padrão (médio)	Teste t (% $t >  2 $ )*	
X18	Taxa de fecundidade	Superior:	0,072	0,109	0,654	0,854	Frágil	-0,313	0,129	-2,425 (70%)	Robusta
		Inferior:	-0,548	0,136	-4,017	0,861					
X19	Expectativa de vida	Superior:	0,263	0,062	4,213	0,880	Frágil	0,154	0,083	1,851 (53%)	Não Robusta
		Inferior:	0,070	0,118	-2,248	0,842					
X20	Temperatura	Superior:	0,266	0,096	2,753	0,865	Frágil	0,033	0,098	0,334 (9%)	Não Robusta
		Inferior:	-0,265	0,118	-2,248	0,842					
X21	Latitude	Superior:	0,113	0,049	2,298	0,856	Frágil	-0,063	0,04	-1,551 (39%)	Não Robusta
		Inferior:	-0,271	0,047	-5,754	0,884					
X22	Índice pluviométrico	Superior:	-0,0004	0,0005	-0,886	0,895	Frágil	-0,0014	0,0005	-2,656 (88%)	Robusta
		Inferior:	-0,0019	0,0005	-3,517	0,854					
X23	Consumo de energia elétrica	Superior:	2,72E-06	2,1E-05	0,135	0,845	Frágil	-0,00002	0,00002	-1,062 (4%)	Não Robusta
		Inferior:	-5,15E-05	2,59E-05	-1,991	0,830					
X24	Instalações sanitárias	Superior:	0,027	0,016	1,658	0,862	Frágil	-0,005	0,018	-0,280 (4%)	Não Robusta
		Inferior:	-0,044	0,014	-3,108	0,871					
X25	Água canalizada	Superior:	0,026	0,016	1,651	0,852	Frágil	0,001	0,026	0,051 (14%)	Não Robusta
		Inferior:	-0,059	0,022	-2,712	0,851					
X26	Iluminação elétrica	Superior:	0,021	0,015	1,426	0,850	Frágil	-0,008	0,024	-0,328 (18%)	Não Robusta
		Inferior:	-0,051	0,016	-3,166	0,854					
X27	Índice L de Theil	Superior:	4,348	2,640	1,647	0,909	Frágil	1,242	2,688	0,462 (0%)	Não Robusta
		Inferior:	-1,510	3,150	-0,479	0,894					
X28	Carga tributária global	Superior:	0,006	0,067	0,089	0,911	Frágil	-0,078	0,06	-1342 (13%)	Não Robusta
		Inferior:	-0,157	0,052	-3,033	0,904					
X29	(Carga tributária global) <sup>2</sup>	Superior:	-0,0018	0,0020	-0,894	0,912	Frágil	-0,0038	0,0018	-2,080 (57%)	Robusta
		Inferior:	-0,0061	0,0019	-3,241	0,911					
X30	Taxa líquida de migração	Superior:	0,128	0,031	4,186	0,881	Frágil	0,068	0,031	2,156 (58%)	Robusta
		Inferior:	0,021	0,033	0,637	0,854					

Fonte: Elaboração Própria dos Autores.

Observação: \*Porcentagem das regressões nas quais o coeficiente apresenta uma estatística *t-student* maior que  $|2|$ .

total de 21 variáveis, foram estimados  $M=1.140$  [ $M=20!/(3!17!)$ ] modelos. Para outras quatro variáveis (X27, X28, X29, X30), utilizou-se um *pool* de 24 variáveis e foram estimados  $M=2.024$  [ $M=24!/(3!21!)$ ] modelos.<sup>27</sup> Portanto, no total foram feitas 32.096 regressões. Já para o teste 1, ou teste EBA (LEVINE; RENELT, 1992), eliminaram-se das combinações anteriores aquelas variáveis com uma correlação superior a 0,70. Os resultados dos testes encontram-se na Tabela 2 e serão discutidos a seguir.

Na Tabela 1, para o teste 1 (EBA) foram indicados os “dois limites extremos” (como definido na seção anterior). Este estudo limitou-se a informar o menor e o maior coeficiente, não fazendo a soma ou subtração dos dois desvios-padrão, tendo em vista que os coeficientes trocaram de sinal antes mesmo da operação algébrica, exceto para as variáveis X11, X17, X19, X22, X29 e X30. Entretanto, como o limite extremo superior de todas as variáveis é positivo e o limite extremo inferior é negativo, concluiu-se que todas são frágeis. Tais resultados mostram de maneira clara como a estimação de uma única equação pode levar a uma inferência espúria.

O teste 1 foi questionado por Sala-i-Martin, visto que o autor critica a severidade do teste<sup>28</sup> e argumenta que em vez de se analisar os extremos das estimativas de uma variável específica, é necessário fazer a análise de toda a distribuição da variável explicativa, calculando-se, assim, a média dos coeficientes e dos desvios-padrão. Na coluna do teste 2, tem-se a porcentagem das 1.140 (ou 2.024) regressões nas quais o coeficiente da variável testada foi estatisticamente diferente de zero (definido como uma estatística *t-student* maior que  $|2|$ ). Pode-se observar que algumas variáveis foram estatisticamente significantes em mais de 60% ou 70% das vezes, enquanto outras foram significantes em menos de 10% das vezes. Entretanto, o teste dos

27. Essa diferença no número de modelos estimados deve-se a diferenças no período utilizado para o teste das variáveis. Por exemplo, quando testadas as primeiras 21 variáveis, não foram incluídas X27, X28, X29 e X30 porque estas contavam com apenas três décadas (54 observações), enquanto aquelas abrangiam quatro décadas (72 observações).

28. Segundo Brock; Durlauf e West (2003) a maior crítica ao método “*extreme bound analysis*” é sua insensibilidade relativa ao bom ajuste (*goodness of fit*) de diferentes especificações de modelos. Dessa forma, modelos mal especificados podem estar prejudicando o teste proposto.

limites extremos dá a mesma classificação para todas as variáveis: frágil. Portanto, a seguir serão analisados os resultados baseados nos resultados do segundo teste.

#### 4.1 – Variáveis que Têm uma Correlação Robusta com o Crescimento

As colunas referentes ao teste 2 nos trazem as estimativas médias de  $\beta z$  e do desvio-padrão de cada variável  $z$  testada. Com essas médias, efetuou-se o teste *t-student* ao nível de significância de 5%. De acordo com o resultado do teste, classificou-se cada variável como robusta ou não-robusta. Em resumo, sete variáveis de um pool de 25 foram consideradas robustas e são analisadas a seguir.

(1) Nível do PIB *per capita* no início do período, X4. Verificou-se que o coeficiente do nível do PIB *per capita* é negativo e significativo, confirmando estudos anteriores da vasta literatura sobre o assunto. O valor médio do coeficiente é -1,476 (e o desvio-padrão é 0,736). Assim, este resultado revela que os estados com níveis iniciais de renda mais baixos tendem a crescer mais rapidamente que os estados com maiores níveis de renda. Entretanto, como se está controlando as regressões com uma série de outras variáveis, pode-se afirmar a ocorrência de convergência beta condicional, ou seja, os estados não estão convergindo para um mesmo nível de PIB *per capita*, e sim para níveis próprios de estado estacionário.

(2) Taxa de urbanização, X11. O coeficiente apresentou uma relação significativa e negativa (-5,814) com as taxas de crescimento econômico dos estados. Esse resultado sugere que a concentração de pessoas em áreas urbanas afetou negativamente o crescimento econômico dos estados, visto que a grande concentração de pessoas em áreas urbanas pode estar causando prejuízos ao desenvolvimento dessas regiões. É importante salientar que o sinal esperado para esse coeficiente era positivo. Tendo em vista que a taxa de urbanização é uma *proxy* para economias de aglomeração, a Nova Geografia Econômica prediz que regiões aglomeradas tendem a apresentar maior taxa de crescimento econômico. Portanto, para o caso dos estados brasileiros essa previsão da teoria parece não ter uma validade empírica.

(3) Taxa de mortalidade infantil, X17. O coeficiente

da taxa de mortalidade infantil é negativo (-0,018), mostrando que um pior estado de saúde prejudica o crescimento econômico dos estados. Esse pior estado de saúde pode ser traduzido em maior mortalidade infantil; logo, estados que apresentaram menores níveis de mortalidade tiveram maiores taxas de crescimento do PIB *per capita*. Como visto, nesse caso a teoria não foi negada.

(4) Taxa de fecundidade, X18. Taxas de fecundidade mais baixas contribuíram para o maior crescimento do PIB *per capita* dos estados, ou seja, o coeficiente desta variável é negativo (-0,313). Entretanto, aqui deve-se ter cuidado para tratar essa questão. Como salientado por Temple (1999), as decisões sobre fecundidade podem estar fortemente correlacionadas com outros aspectos do desenvolvimento social. Assim, a variável fecundidade pode estar representando uma *proxy* para os mais diversos arranjos sociais, que passam por questões como religião, educação e tradição. Outra questão que se deve salientar é o fato de se dispor apenas de dados para os anos de 1991 e 2000. Assim, os dados para os anos de 1960, 1970 e 1980 foram calculados por extrapolação. Mesmo assim, considerou-se importante incluir essa variável, pois de maneira geral a taxa de fecundidade apresentou uma tendência de queda no Brasil no período estudado.

(5) Índice pluviométrico, X22. Entre as variáveis geográficas (X20, X21, X22), a única que apresentou um coeficiente estatisticamente significativo foi o índice pluviométrico (-0,0014). Pode-se inferir que regiões em que o índice de chuva é maior têm um menor crescimento econômico. Porém, é importante ressaltar que a medida dessa variável tem problemas, visto que foi usado o volume total de precipitação na capital como uma *proxy* representativa de precipitação em todo o estado. Apesar disso, uma interpretação para o resultado encontrado pode vir de Hall e Jones (1996). Hall e Jones (1996) fazem um estudo empírico sobre os determinantes do nível da renda *per capita* de um total de 133 países. Hall e Jones (1996) testam a importância do clima na determinação do nível da renda *per capita* dos países, e também de outras variáveis explicativas (idioma, abertura ao comércio exterior e tipos de organização econômica). Apesar de não existir um modelo teórico que justifique a inclusão da variável

clima no modelo econométrico, os autores encontraram uma relação negativa e estatisticamente significativa entre o nível da renda *per capita* e clima. Esses resultados para o clima dão suporte à hipótese de que clima temperado favorece a produtividade. Azzoni et al. (2000) argumentam que a variável índice pluviométrico está relacionada à natureza e, provavelmente, tem seus efeitos na produtividade da agricultura.

(6) Carga tributária global ao quadrado, X29. Com base na relação não-linear entre crescimento e taxação proposta por Barro (1990) estabeleceu-se uma especificação polinomial de segundo grau para a variável carga tributária global. A estimativa para o coeficiente foi significativa e negativa (-0,0038). O resultado do teste de robustez não nega a validade de uma especificação não-linear, em forma de “U-invertido”, envolvendo crescimento e taxação. Logo, verifica-se a existência de um ponto ótimo para a carga tributária, capaz de maximizar as taxas crescimento econômico estaduais.

(7) Taxa líquida de migração, X30. O processo imigratório contribuiu para o aumento nas taxas de crescimento econômico estaduais. Ao apresentar um coeficiente com um sinal positivo (+0,068), a variável migração valida a predição feita pela Nova Geografia Econômica (NGE). Como salientado anteriormente, a imigração afeta positivamente o crescimento econômico, tendo em vista que uma maior quantidade de trabalhadores implica maior mercado consumidor, atraindo mais firmas e reduzindo seus custos.

## 4.2 – Variáveis que não Têm Correlação Robusta com o Crescimento

As variáveis participação do setor industrial (X6), participação do setor de comércio (X8) e participação do setor de serviços (X9) no PIB fizeram parte deste estudo, no intuito de verificar se a presença de economias de escala e/ou de escopo, oriundas dos encadeamentos para frente e para trás nestes setores teve impacto positivo sobre o crescimento econômico dos estados, como previsto pela NGE. Além disso, testou-se a relação entre participação do setor agropecuário (X7) no PIB e crescimento econômico, buscando verificar se existe uma força centrífuga que bloqueia a concentração espacial e inibe o crescimento regional. O resultado dos testes

mostrou que nenhuma dessas variáveis tem uma correlação robusta com o crescimento. Entretanto, deve-se salientar, a variável participação do setor de serviços (X9) no PIB foi estatisticamente significativa ao nível de 6%, aparecendo em 47% das regressões com uma estatística *t-student* maior que |2|. Portanto, é importante destacar que essa variável pode ser considerada marginalmente relacionada com as taxas de crescimento econômico dos estados brasileiros, visto que sua significância estatística fica fora, mas bem próxima do limite estabelecido para uma variável ser considerada robusta.

O coeficiente da variável de densidade demográfica (X10) não foi significativo. Portanto, não se captaram efeitos de congestão, ou seja, áreas densamente habitadas podem ter custos elevados, causando deseconomias externas e, assim, menor crescimento econômico. Outras duas variáveis referentes ao capital humano (X13, X14) foram não-significativas. Vale ressaltar que a taxa de matrícula no ensino secundário (X13) foi estatisticamente significativa a 10%, apresentando um coeficiente significativo e negativo em 35% das estimações, ou seja, uma estatística *t-student* maior do que |2|. Assim, tal variável é considerada marginalmente relacionada com as taxas de crescimento econômico dos estados brasileiros. Entretanto, uma *proxy* de capital humano, taxa de matrícula no ensino primário (X12), já estava presente em todas as regressões, tendo em vista que era uma variável fixa. Com relação às variáveis benefícios de transporte (X15 e X16), estas não foram significativas. Já a segunda *proxy* para o estado de saúde (X19), isto é, expectativa de vida, apresentou-se significativa no nível de 8%, sendo que em 53% das regressões o coeficiente teve uma estatística *t-student* maior que |2|. Essa variável, também, pode ser considerada marginalmente relacionada com as taxas de crescimento econômico dos estados brasileiros. Assim, confirma-se a importância que o estado de saúde da população tem sobre o crescimento econômico, visto que a primeira variável do estado de saúde (X17) tinha mostrado uma correlação robusta com o crescimento econômico dos estados.

O efeito sobre o crescimento é nulo para as variáveis geográficas, temperatura (X20) e latitude (X21). Vale ressaltar que embora essas duas *proxies*

não sejam significativas, uma terceira (X22) teve seu coeficiente significativo. Isso mostra a importância de, ao se fazerem os testes, usar-se mais de uma *proxy* para cada variável. Muitas vezes, ao se escolher apenas uma *proxy* para determinada variável, pode-se verificar sua insignificância estatística, mas esta se deve ao uso de uma má *proxy*.

Nenhuma das três *proxies* para infraestrutura (X23, X24 e X25) apresentou sinais significativos. Assim, parece não existir, em nível agregado, uma relação direta entre infraestrutura e crescimento estadual. Logo, são importantes estudos microrregionais para entender essas relações. Também, verificou-se que não existe relação entre desigualdade de renda (X27) e crescimento. Por último, a especificação linear e negativa entre taxação e crescimento econômico, proposta pelo modelo de Rebelo (1991), não foi aceita, tendo o coeficiente da carga tributária global (X28) apresentado um sinal estatisticamente não significativo.

## 5 – CONCLUSÕES

---

A teoria econômica sugere uma variedade de determinantes para o crescimento econômico. A literatura empírica mostra um número ainda maior de determinantes para as variações nas taxas de crescimento da renda *per capita* dos estados brasileiros. No intuito de clarificar tal assunto, este trabalho procurou responder à seguinte pergunta: quais variáveis possuem uma correlação robusta com as taxas de crescimento do PIB *per capita* dos estados brasileiros? Para essa resposta foram utilizados dois testes de robustez: *Extreme Bounds Analysis* (EBA), proposto por Levine e Renelt (1992) e a análise de toda a distribuição da variável explicativa, proposta por Sala-i-Martin (1997).

Os resultados mostraram que, de acordo com o teste EBA, nenhuma variável é robusta. Entretanto, ao se usar o teste em que toda a distribuição dos coeficientes é analisada, chega-se à conclusão de que migração contribui positivamente para o crescimento do PIB *per capita* dos estados; enquanto urbanização, taxa de mortalidade infantil, fecundidade, pluviometria e carga tributária (elevada ao quadrado) estão correlacionadas negativamente com as taxas de crescimento do PIB *per capita* dos estados brasileiros. Além disso, confirma-se a ocorrência de convergência condicional dos PIBs *per capita* dos estados.

Em relação à migração, é importante salientar o resultado encontrado. Ao contrário do que a teoria tradicional (o modelo de Solow com migração apresenta o trabalho saindo da região pobre para a região rica, diminuindo a taxa de crescimento dessa última) prevê, a relação encontrada entre crescimento e migração foi positiva. Logo, ao se analisar as quatro décadas, pode-se concluir que os estados que tiveram uma entrada líquida de migrantes cresceram mais. Nesse sentido, pode-se argumentar que se as migrações foram das regiões pobres para as mais desenvolvidas, as migrações contribuíram para um crescimento das regiões mais ricas, corroborando previsões da Nova Geografia Econômica (NGE). Por outro lado, a concentração de pessoas em áreas urbanas afetou negativamente o crescimento econômico dos estados. É o que infere-se a partir do coeficiente da variável taxa de urbanização. A grande concentração de pessoas em áreas urbanas pode estar causando prejuízos ao desenvolvimento dessas regiões. Outra variável que merece destaque é a carga tributária global, que apresentou uma relação de longo prazo não-linear com o crescimento econômico dos estados. Foi encontrada uma relação em forma de “U-invertido”, mostrando que o crescimento dos estados brasileiros pode ser maximizado a partir de uma carga tributária global ótima.

Um outro grupo de variáveis que foram classificadas como não-robustas não pode ser esquecido: participação do setor de serviços no PIB, taxa de matrícula no segundo grau e expectativa de vida. Essas três variáveis são significativas nos níveis de 6%, 10% e 8%, respectivamente. Pode-se dizer que, essas variáveis são marginalmente relacionadas com as taxas de crescimento econômico dos estados brasileiros, visto que suas significâncias estatísticas ficam fora, mas bem próximas do limite estabelecido para uma variável ser considerada robusta. Vale destacar que uma variável de capital humano, mais especificamente a taxa de matrícula no primeiro grau foi mantida fixa em todas as regressões, o que pode ter afetado o sinal e a significância da variável testada, a taxa de matrícula no segundo grau. Por fim, salienta-se que, apesar da variável expectativa de vida não ser considerada robusta, outra *proxy* do estado de saúde da população, mortalidade infantil, já havia sido

considerada robustamente correlacionada com o crescimento econômico dos estados brasileiros.

Este trabalho buscou salientar questões importantes quanto ao estudo dos determinantes do crescimento econômico dos estados brasileiros. Aparentemente, tais questões podem ser divididas em duas áreas: uma metodológica e outra relacionada ao crescimento dos estados, discutida anteriormente. A questão metodológica é mais um alerta que se quis fazer aos estudos econométricos baseados em regressões únicas. O que se quer ressaltar é que quando se tem um *pool* de variáveis sugerido tanto pela literatura teórica quanto empírica, pode-se chegar a qualquer resultado. Isso quer dizer que, dependendo da combinação de variáveis que se escolhe, qualquer variável pode ser estatisticamente significativa. Esse fato pode ser constatado neste estudo, visto que todas as variáveis em pelo menos um dos modelos estimados tiveram coeficientes significativos. Logo, uma análise menos sistemática poderia ter levado a erros de interpretação dos resultados. Portanto, quando a literatura teórica e/ou empírica sugere variáveis explicativas para a determinação de algum fenômeno econômico (neste caso, a variação da PIB *per capita* estadual), é importante a execução de testes de robustez nas variáveis estudadas.

Outra questão metodológica a salientar é a importância de se trabalhar com dados em painel quando se estuda crescimento econômico, visto que análises em *cross-section*, ao omitir pelo menos uma variável importante, qual seja, o nível da tecnologia, apresentam resultados viesados. Por sua vez, regressões utilizando dados em painel, permitem fazer um controle para a omissão de variáveis que estão presentes ao longo do tempo, não viesando os coeficientes.

Em resumo, este estudo logrou avançar na discussão sobre os determinantes do crescimento econômico dos estados brasileiros em dois aspectos. Em primeiro lugar, construiu-se uma base de dados para os estados brasileiros compreendendo 27 variáveis para os anos 1960-2000. Com tal base de dados será possível, aprofundar a pesquisa a respeito da robustez de variáveis para o caso brasileiro, por exemplo, aplicando o teste bayesiano sugerido por Sala-i-Martin; Doppelhofer e Miller (2004). Segundo, o presente estudo foi o primeiro

a discutir a robustez dos determinantes do crescimento econômico dos estados brasileiros utilizando métodos (LEVINE; RENELT, 1992; SALA-I-MARTIN, 1997) empregados na discussão do crescimento econômico entre países. Com tal abordagem foi possível lidar de forma pragmática com o problema de incerteza de modelos no intuito de avaliar os determinantes do crescimento econômico dos estados brasileiros.

Por fim, deve-se salientar que estudos como este, ao lado de estudos de casos, estudos em nível microrregional ou municipal são importantes para que se compreenda a dinâmica do desenvolvimento regional brasileiro, suas possibilidades e seus desafios. Pode-se mencionar, assim, Temple (1999, p. 152): *“At least some knowledge of the average pattern is the beginning of wisdom, and although we have not learnt as much as might be hoped, it is always worth remembering how little we knew when we started”*<sup>29</sup>.

## AGRADECIMENTO

Os autores agradecem os comentários dos dois pareceristas anônimos. Os erros remanescentes são dos autores.

## ABSTRACT

The main goal of this paper is to determine which variables have a robust correlation with the growth rate of per capita Gross Domestic Product (GDP) of Brazilian states between 1960 and 2000. We have run two robustness tests suggested by the literature. The first approach is proposed by Levine e Renelt (1992), that used the Extreme Bounds Analysis (EBA) test to identify robust variables related to the economic growth. An alternative approach to the previous one was considered by Sala-i-Martin (1997). In sum, based on those tests, we can affirm that urbanization, mortality rates, fertility rates, rain (climate), tax burden and migration have a robust correlation with the growth rates of per capita GDP of the Brazilian states. Moreover, according to the tests, it was confirmed the occurrence of conditional convergence for the Brazilian states GDPs.

<sup>29</sup>Tradução feita pelos autores: “Ao menos algum entendimento do padrão médio é o começo da sabedoria, e embora nós não tenhamos apreendido o esperado, é sempre válido lembrar o quão pouco sabíamos quando começamos”.

## KEY WORDS:

Robustness Test. Economic Growth. Convergence.

## REFERÊNCIAS

- ALESINA, A.; RODRICK, D. Distributive politics and economic growth. **Quarterly Journal of Economics**, v. 109, n. 2, p. 465-490, May 1994.
- AZZONI, C. R. Crescimento econômico e convergência das rendas regionais: o caso brasileiro à luz da nova teoria do crescimento. In: ENCONTRO NACIONAL DE ECONOMIA, 22., 1994, Florianópolis. **Anais...** 1994. Florianópolis: Anpec, 1994. V. 1.
- AZZONI, C. et al. Geografia e convergência da renda entre os estados brasileiros. In: HENRIQUES, R. (Org.). **Desigualdade e pobreza no Brasil**. Rio de Janeiro: Ipea, 2000. p. 299-343.
- BARRO, R. J. **Determinants of economic growth: a cross-country empirical study**. Cambridge: NBER, 1996. 145 p. (Working Papers, n. 5698).
- BARRO, R. J. Economic growth in a cross section of countries. **Quarterly Journal of Economics**, v. 106, n. 2, p. 407-43, May 1991.
- BARRO, R. J. Government spending in a simple model of endogenous growth. **Journal of Political Economy**, v. 98, n. 8, p. 103-125, Oct. 1990.
- BARRO, R.; SALA-I-MARTIN, X. **Economic growth**. 2. ed. Cambridge: MIT, 2003. 654 p.
- BLANCO, F. A.; REIS, E. J. **Capacidade tributária dos estados brasileiros, 1970/90**. Brasília, DF: Ipea, 1996. 26 p. (Texto para Discussão, n. 404).
- BLEANEY, M. F.; FIGUEIRÊDO, L. Can the economic geography explain spatial inequality in Brazil?. In: CONFERENCE ON SPATIAL INEQUALITY IN LATIN AMERICA, 2002, Cholula, México. **Proceedings...** Cholula, México: WIDER, 2002.
- BLOOM, D. E.; CANNING, D.; SEVILHA, J. **The effect of health on economic growth: theory and evidence**. Cambridge: NBER, 2001. 26 p. (Working Papers, n. 8587).
- BOSWORTH, B; COLLINS, S. The empirics of growth:

- an update. **Brookings Papers on Economic Activity**, v. 2003, n. 2, p. 113-179, 2003.
- BROCK, W.; DURLAUF, S. Growth empirics and reality. **World Bank Economic Review**, v. 15, p. 229-272, 2001.
- BROCK, W.; DURLAUF, S.; WEST, K. Policy evaluation in uncertain economic environments. **Brookings Papers of Economic Activity**, p. 235-322, 2003.
- CHRISTALLER, W. **Central Places in Southern Germany**. [S.l.]: Prentice-Hall, 1966. Primeira publicação em Alemão em 1933.
- CRESPO CUARESMA, J.; DOPPELHOFER, G. Nonlinearities in cross-country growth regressions: A Bayesian Averaging of Thresholds (BAT) approach. **Journal of Macroeconomics**, v. 29, p. 541-554, 2007.
- DURLAUF, S.; FRANKEL, J. The empirics of growth: an update: comments and discussion. **Brookings Papers on Economic Activity**, v. 2003, n. 2, p. 180-206, 2003.
- DURLAUF, S.; JOHNSON, P.; TEMPLE, J. **Growth econometrics**. [S.l.: s.n.], 2004. 185 p. (Working Paper). Mimeografado.
- EICHER, T. S.; PAPAGEORGIOU, C.; ROEHN, O. Unraveling the fortunes of the fortunate: an Iterative Bayesian Model Averaging (IBMA) approach. **Journal of Macroeconomics**, v. 29, p. 494-514, 2007.
- FERNÁNDEZ, C.; LEY, E.; STEEL, M. F. Model uncertainty in cross-country growth regressions. **Journal of Applied Econometrics**, v. 16, p. 563-576, 2001.
- FERREIRA, A. H. B. **A distribuição interestadual e inter-regional da renda no Brasil: tendências recentes**. 1995. 101f. Tese (Doutorado em Economia) - Departamento de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo Horizonte, 1995.
- FERREIRA, A. H. B.; DINIZ, C. C. **Convergência entre as rendas per capita estaduais no Brasil**. Belo Horizonte: UFMG, 1994. 27 p. (Texto para Discussão, n. 79).
- FERREIRA, P. C. G.; ELLERY JR., R. G. Convergência entre a renda per capita dos estados brasileiros. **Revista de Econometria**, v. 16, n. 1, p. 83-103, abr. 1996.
- FIGUEIRÊDO, L.; NORONHA, K. V. K.; ANDRADE, M. V. **Os impactos da saúde sobre o crescimento econômico na década de 90: uma análise para os estados brasileiros**. Belo Horizonte: UFMG, 2003. 26 p. (Texto para Discussão, n. 219). Disponível em: <<http://www.cedeplar.ufmg.br/pesquisas/td/TD%20219.pdf>>. Acesso em: 2010.
- FUJITA, M.; KRUGMAN, P.; VENABLES, A. **The spatial economy**. Cambridge: MIT, 1999. 367 p.
- FUJITA, M.; THISSE, J. F. **Economics of agglomeration: cities, industrial location and regional growth**. Cambridge: Cambridge University, 2002. 466 p.
- GREENE, W. **Econometric analysis**. 5. ed. [S.l.]: Upper Saddle River, 2003. 1026 p.
- HALL, R. E.; JONES, C. I. **The productivity of nations**. Cambridge: NBER, 1996. 47 p. (Working Papers, n. 5812).
- IBGE. **Anuário estatístico do Brasil**. [S.l.], 1960.
- IPEADATA. **Dados macroeconômicos e regionais**. Disponível em: <<http://www.ipeadata.gov.br>>. Acesso em: 10 abr. 2004.
- ISLAM, N. What have we learnt from the convergence debate?. **Journal of Economic Surveys**, v. 17, p. 309-362, 2003.
- JONES, C. I. **Introdução à teoria do crescimento econômico**. Rio de Janeiro: Campus, 2000. 178 p.
- JUDGE, G. G. et al. **Introduction to the theory and practice of econometrics**. 2. ed. New York: John Wiley & Sons, 1988.
- LEAMER, E. E. **Specification searches: ad hoc inference with nonexperimental data**. New York: Wiley & Sons, 1978.
- LEAMER, E. E. Let's take the con out of econometrics. **American Economic Review**, v. 73, p. 31-43, 1983.
- LEAMER, E. E. Sensitive analyses would help. **American Economic Review**, v. 75, p. 308-313, 1985.

- LEVINE, R.; RENELT, D. A Sensitivity analysis of cross-country growth regressions. **American Economic Review**, v. 82, n. 4, p. 942-963, Sept. 1992.
- LEY, E.; STEEL, M. Jointness in variable selection with applications to growth regression. **Journal of Macroeconomics**, v. 29, p. 476-493, 2007.
- LLÉDO, V. D.; FERREIRA, P. C. Crescimento endógeno, distribuição de renda e política fiscal: uma análise cross-section para os estados brasileiros. **Pesquisa e Planejamento Econômico (PPE)**, v. 27, n. 1, p. 41-70, abr. 1997.
- LÖSCH, A. **The economics of location**. New Haven: Yale University Press, 1954.
- LUCAS, R. On the mechanics of economic development. **Journal of Monetary Economics**, v. 22, n. 1, p. 3-42, Jul. 1988.
- MANKIW, N. G.; ROMER, D.; WEIL, D. N. A contribution to the empirics of economic growth. **Quarterly Journal of Economics**, v. 107, n. 2, p. 407-37, May 1992.
- MARSHALL, A. **Principles of economics**. Londres: Macmillan, 1920.
- REBELO, S. Long run policy analysis and long run growth. **Journal of Political Economy**, v. 99, n. 3, p. 500-521, Jun. 1991.
- ROMER, P. M. Endogenous technological change. **Journal of Political Economy**, v. 98, n. 5 (part 2), p. 71-102, Oct. 1990.
- ROMER, P. M. Increasing returns and long-run growth. **Journal of Political Economy**, v. 94, n. 5, p. 1002-1037, Oct. 1986.
- SALA-I-MARTIN, X. The classical approach to convergence analysis. **The Economic Journal**, v. 106, n. 437, p. 1019-1036, 1996.
- SALA-I-MARTIN, X. I just ran two million regressions. **American Economic Review**, v. 87, p. 178-183, 1997.
- SALA-I-MARTIN, X.; DOPPELHOFER, G.; MILLER, R. Determinants of long-term growth: a Bayesian Averaging of Classical Estimates (BACE) approach. **American Economic Review**, v. 94, p. 813-835, 2004.
- SOLOW, R. M. A contribution to the theory of economic growth. **Quarterly Journal of Economics**, v. 70, n. 1, p. 65-94, Feb. 1956.
- TEMPLE, J. Growth regressions and what the textbooks do not tell you. **Bulletin of Economic Research**, v. 52, p. 181-205, 2000.
- TEMPLE, J. The new growth evidence. **Journal of Economic Literature**, v. 37, n. 1, p. 112-156, Mar. 1999.
- THOMAS, R. L. **Modern Econometrics: an introduction**. 2. ed. Harlow: Addison-Wesley, 1997.

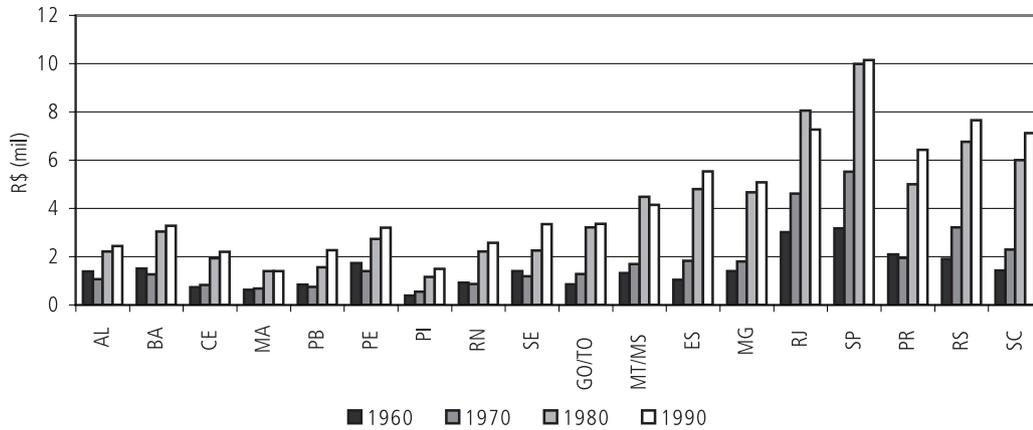
---

Recebido para publicação em: 15.10.2008



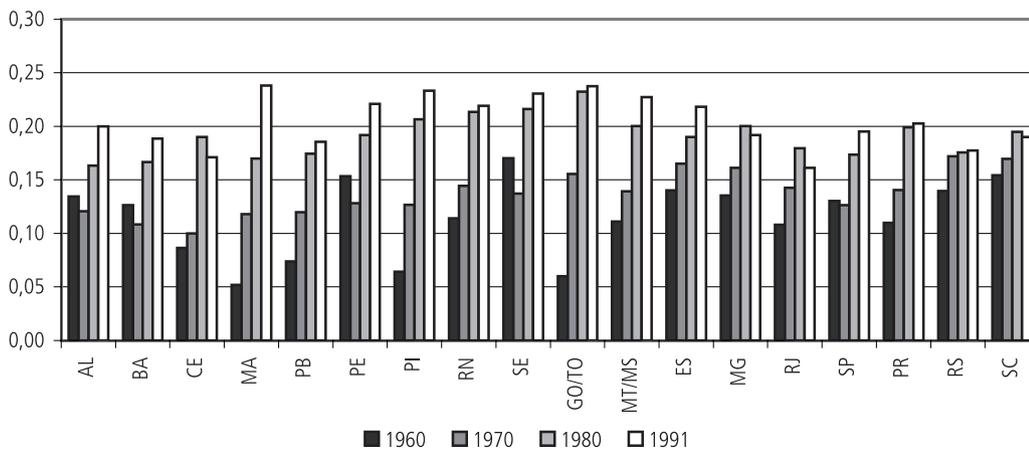


## ANEXO B



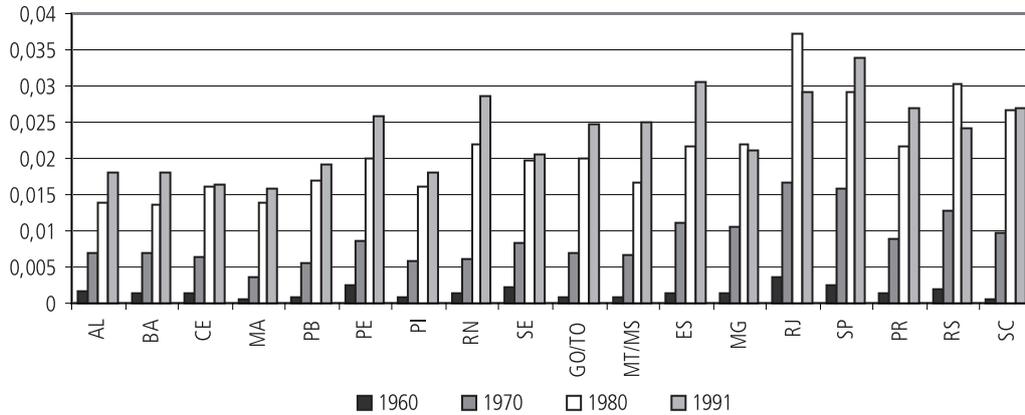
**Gráfico 1B – PIB Per capita Estadual no Início das Décadas (X4)**

Fonte: Elaboração Própria dos Autores.



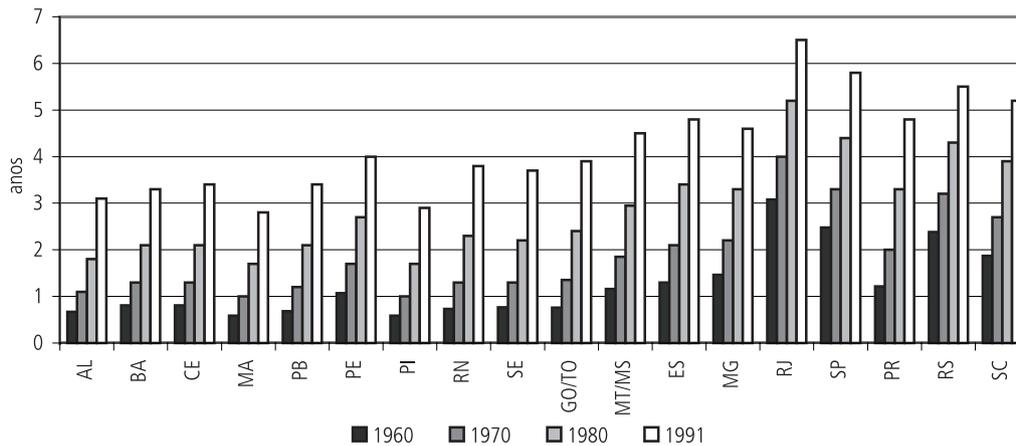
**Gráfico 2B – Taxa de Matrícula no Ensino Primário (X12)**

Fonte: Elaboração Própria dos Autores.



**Gráfico 3B – Taxa de Matrícula no Ensino Secundário (X13)**

Fonte: Elaboração Própria dos Autores.



**Gráfico 4B – Número Médio de Anos de Estudo das Pessoas de 25 ou + Anos de Idade (X14)**

Fonte: Elaboração Própria dos Autores.

