

**Título:**

**Existe uma curva de Kuznets? Uma Análise para os Dados Nordestinos sob a Perspectiva do *Model Averaging***

**ÁREA 2 – ECONOMIA SOCIAL**

**Autores:**

**Fabício Linhares – Professor do CAEN/UFC. Email: [flinhares@caen.ufc.br](mailto:flinhares@caen.ufc.br)**

**Paulo Ernesto – Aluno do Curso de Economia da FEAAC/UFC e Bolsista de Iniciação Científica do CNPq. Email: [pauloernesto@caen.ufc.br](mailto:pauloernesto@caen.ufc.br)**

Endereço Postal: CAEN - Avenida da Universidade, 2700 • Benfica • CEP 60020-181 • Fortaleza / Ceará. Fone: (85)3366-7751.

### **RESUMO:**

A hipótese de Kuznets estabelece uma relação em forma de U invertido entre desigualdade de renda e crescimento econômico. Todavia, há uma considerável incerteza acerca de quais e quantas variáveis influenciam a distribuição de renda, o que pode influenciar significativamente na constatação dessa hipótese. O presente estudo tem como objetivo procurar evidências da hipótese de Kuznets diante dessa incerteza. Para tal, emprega-se uma metodologia econométrica que constrói uma média ponderada das estimativas dos coeficientes de interesse de vários modelos candidatos, em uma abordagem conhecida como *Jackknife Model Averaging*. Ao contrário da prática comum em grande parte da literatura, essa metodologia não especifica, através de um critério para seleção de variáveis, um modelo final onde a investigação é conduzida, mas baseia-se no conjunto de resultados de todos os modelos estimados. O estudo utiliza dados para os municípios nordestinos e considera seis diferentes cenários, todos com uma formulação distinta da equação de regressão. Verifica-se que somente em dois cenários os sinais dos coeficientes estavam de acordo com a referida curva. Nenhum modelo selecionado pelos critérios de informação de Akaike e Schwartz sustenta a hipótese de Kuznets. Percebe-se, portanto, que a verificação dessa hipótese depende da arbitrariedade da escolha de variáveis explicativas em modelos de regressão.

**Palavras-chave:** *Model Averaging*, Kuznets, modelos de regressão.

### **ABSTRACT:**

The Kuznets hypothesis provides an inverted U-shaped association between income inequality and economic growth. However, there is considerable uncertainty about what and how many variables influence the relationship between economic growth - the income distribution. This study aimed to find evidence to validate the Kuznets hypothesis for the municipalities in the Northeast Region of Brazil for the year 2000. Unlike much of the literature, this analysis seeks an alternative methodology for estimating the coefficients of the variables of interest. It was proposed to estimate all possible candidate models and compute a weighted average of the estimated coefficients of variables of interest, an approach known as Model Averaging. The choice of the weighting was based on the criterion of cross-validation, with the technique known as Jackknife Model Averaging. Given the uncertainty about the selection process of the model, it was stipulated six scenarios, each with a different basic econometric model. Only two scenarios coefficients were in agreement with that curve. No model suggested by the information criteria AIC and BIC validated the Kuznets hypothesis. The arbitrariness of the choice of explanatory variables eventually weakens the regression analysis.

**Keywords:** Model Averaging, Kuznets, regression models.

**JEL Classification Numbers:** R1; R11; R15

## 1. Introdução

Simon Kuznets, ao observar a fase de transição entre economias agrárias para industrializadas na Europa, afirmou que o processo de crescimento econômico é acompanhado por um aumento da desigualdade de renda. O aumento da concentração de renda teria explicação na transferência de pessoas e recursos de áreas rurais para áreas urbanas. Como a renda *per capita* média da população rural era menor que a da população urbana, e a desigualdade nas participações percentuais dentro da distribuição da população rural era menor que a da população urbana, a idéia do aumento da desigualdade advindo do processo migratório era bastante plausível. Todavia, a desigualdade crescia a taxas cada vez menores e, após atingir um determinado nível de desenvolvimento econômico, a concentração de renda passava a cair. Desta forma, relação entre desigualdade na distribuição de renda e crescimento econômico seria na forma de um “U invertido”.

A idéia geral por trás dessa hipótese tem como base a evolução de diferentes setores de produção na economia. No primeiro estágio, onde o setor dominante é pouco desenvolvido e proporciona uma baixa remuneração aos seus trabalhadores, a desigualdade também deve ser muito baixa, já que quase todo mundo adquire renda desse setor. À medida que o processo de crescimento se inicia, impulsionada por um setor mais dinâmico, a desigualdade de renda aumenta. Parte dos trabalhadores migra do setor tradicional (por exemplo, agrícola), onde a remuneração é menor, para o setor mais dinâmico (industrial), no qual tanto as expectativas quanto a diferenciação das remunerações são maiores. Nesse estágio inicial de crescimento, tanto o capital físico quanto o humano são escassos e desigualmente distribuídos, e os donos de tais capitais são capazes de obter altos retornos. Com o avanço do crescimento, à medida que os dois tipos de capital vão se acumulando e se tornando mais difundidos entre a população, a taxa de retorno do capital físico vai caindo, enquanto diferenças salariais entre trabalhadores qualificados e não qualificados vai diminuindo. A desigualdade de renda, conseqüentemente, tende a reduzir.

Passados cinquenta anos desse estudo, muitos economistas ainda buscam compreender os efeitos do crescimento econômico na desigualdade de renda. Na tentativa de aceitar ou refutar a hipótese de Kuznets, muitos estudos foram feitos para diversos países e regiões, inclusive para regiões do Brasil. No entanto, nem sempre o padrão da curva descrito anteriormente foi constatado, o que levou economistas a questionarem-se qual seria a melhor forma de análise do impacto a longo prazo do aumento do PIB na desigualdade. Uma possível causa da não evidência da curva, segundo a literatura, se apóia nos dados. Muitas vezes a análise é feita a partir de dados de vários países com diferentes níveis de desenvolvimento, ocasionando uma relação fraca entre o desenvolvimento e a concentração de renda. Caso a análise fosse feita para uma amostra mais homogênea, em que países e regiões possuíssem características específicas parecidas, a curva de Kuznets teria uma probabilidade maior de ser evidenciada. Este parece ser o caso dos municípios nordestinos, os quais não apresentam grandes divergências entre si, com exceção destes com as capitais estaduais.

A literatura acerca do tema é bastante extensa, com diferentes métodos de análise econométrica, diversas formas funcionais e diferentes estruturas de dados. Segundo e Tejada (2004), o fato de se obter resultados nem sempre favoráveis à existência de U invertido contribui para a especificação de diversas formas funcionais e do uso de diferentes métodos de estimação. Tendo em vista que os coeficientes das variáveis de interesse (renda *per capita* e renda *per capita* ao quadrado) variam de sobremaneira entre os diferentes modelos estimados, tais coeficientes são considerados sensíveis à especificação do modelo usado. Sendo assim, a inferência sobre os coeficientes é dita frágil.

Diante de uma grande quantidade de trabalhos empíricos que buscam comprovar a existência da curva de Kuznets, o pesquisador depara-se com vários modelos econométricos,

todos eles com suas formas funcionais, suas variáveis explicativas e seus diferentes métodos de estimação. Como, então, estimar o efeito do crescimento econômico na desigualdade de renda? Uma possível saída seria escolher um dos modelos apresentados ao pesquisador, e assim utilizá-lo para fazer inferência estatística. Quantificar-se-ia os coeficientes das variáveis de interesse e, após um teste de significância, poder-se-ia aceitar ou refutar a trajetória de U invertido.

Contudo, toda incerteza acerca do processo de seleção do modelo seria ignorada. Tal opção acaba por assumir implicitamente que: não só existe um modelo verdadeiro, mas também que esse modelo está entre os modelos apresentados ao pesquisador. Há, também, uma considerável incerteza acerca de quais variáveis influenciam a relação crescimento econômico – distribuição de renda. A seleção de variáveis explicativas em uma pesquisa empírica é uma questão importante, a qual dificulta o consenso sobre o modelo empírico a ser estimado. A arbitrariedade da escolha de variáveis de controle acaba fragilizando a análise da regressão.

O seguinte trabalho propõe uma saída alternativa. Estimam-se todos os possíveis modelos candidatos e conduz-se a análise a partir de uma média ponderada ótima de todas as estimativas dos coeficientes das variáveis de interesse. Esta abordagem é conhecida como *Model Averaging*. Não será preciso assumir que o verdadeiro modelo existe. Deste modo, está-se considerando não somente a incerteza associada à estimativa condicional do parâmetro, mas também a incerteza do parâmetro estimado dentre os diferentes modelos. O estudo atende a abordagem da *Frequentist Model Averaging* para estimar os pesos dos coeficientes de cada modelo candidato. Seguindo o trabalho de Hansen e Racine (2009), a escolha da ponderação será baseada no critério de *cross-validation*, com a técnica conhecida como *Jackknife Model Averaging* (JMA). Os autores demonstram que o estimador é assintoticamente ótimo no sentido de alcançar o menor erro quadrático de previsão possível. A opção pelo JMA se justifica pelo fato deste ser apropriado para mais modelos lineares gerais, assim, os erros aleatórios podem ter variância heteroscedástica e os modelos candidatos podem não ser “aninhados”.

Assim, pode-se chegar a conclusões mais confiáveis, ou pelo menos mais honestas, sobre a significância do efeito do crescimento econômico na desigualdade dos municípios do Nordeste.

O trabalho foi dividido da seguinte forma: a seção 2 trata de apresentar os mais relevantes estudos acerca da relação crescimento econômico – desigualdade de renda e da incerteza na estimação do modelo. A seção 3 descreve a metodologia utilizada para se estimar os coeficientes de interesse através da abordagem da *Frequentist Model Averaging*. Como não se sabe quais nem quantas variáveis explicativas devem compor o modelo, o trabalho expõe seis cenários, todos com diferentes modelos, explicados pormenorizadamente no capítulo. Uma análise descritiva das variáveis é feita na seção 4, bem como a apresentação destas e seus critérios de seleção. A seção 5 apresenta os resultados das estimações para todos os cenários. Por fim, a seção 6 descreve as conclusões que se podem tirar através dos resultados obtidos.

## **2. Referencial Teórico**

Existe uma vasta literatura, tanto nacional quanto internacional, que trata do referido tema e, usualmente, a hipótese de Kuznets é testada através de estimativas de Mínimos Quadrados Ordinários, obedecendo à modelagem econométrica clássica proposta por Ahluwalia (1976), a qual tem a renda per capita e o coeficiente de Gini como variáveis-chaves.

Em âmbito internacional, Ahluwalia (1976) utiliza dados em *cross-section* para uma amostra de 60 países (40 considerados subdesenvolvidos, 6 socialistas e 14 desenvolvidos). O resultado obtido foi a validação da hipótese de Kuznets.

Anand e Kanbur (1993) criticam a análise proposta por Ahluwalia, alegando que os dados devem ser formulados de maneira mais homogênea. Ao refazerem a análise, dando um melhor tratamento aos dados, os autores acabam por refutar a hipótese de Kuznets.

O estudo de Fields e Jackubson (1994) apresenta outra modelagem, com dados em *pooled cross-section* e em painel para o modelo quadrático do índice de Gini como função da renda *per capita*. Os resultados do *pooled cross-section* apóiam a hipótese do U invertido, já as estimações com efeitos fixos, não.

Contrastando com os trabalhos anteriormente citados, os quais utilizaram regressões condicionadas à média, Huang et al. (2007) faz uso da regressão quantílica, tanto paramétrica quanto semi-paramétrica, visando reexaminar a validade da hipótese de Kuznets em diferentes quantis da função desigualdade condicional. O que se constatou foi que, em países com “leve” desigualdade, a relação crescimento econômico – desigualdade se apresentou condizente com os estudos de Kuznets. No entanto, quando a mesma análise foi feita para países com desigualdade alta ou muito baixa, os resultados contradisseram Kuznets.

Já em âmbito nacional, Berni, Marquetti e Kloeckner (2002) utilizam dados do Theil-L e do Valor Adicionado Bruto *per capita* dos 333 municípios gaúchos relativos ao ano de 1990, com o objetivo de investigar a relação renda *per capita* setorial – desigualdade. Dada a importância da agricultura na economia estadual, faz-se necessário uma associação entre o indicador geral de desigualdade municipal e a renda setorial. A hipótese da curva de U invertido é investigada utilizando-se o método não paramétrico denominado regressão local, o qual permite que se visualize graficamente a relação entre a variável dependente e a independente. Outra virtude do método seria que nenhuma forma paramétrica é imposta aos dados, estes falam por si próprios. Os resultados comprovam a hipótese de Kuznets para o setor industrial e de serviços. O setor agropecuário, no entanto, não apresenta resultados estatisticamente significantes.

Bagolin, Gabe e Pontual (2003) afirmam que modelos econométricos com dados em *cross-section* carregam uma hipótese implícita não satisfatória: a inexistência de diferentes trajetórias de renda e desigualdade. Como características específicas da região podem ser determinantes na sua trajetória histórica, o modelo com dados em *cross-section* estariam obtendo resultados inconsistentes que não poderiam ser usados para a análise de longo prazo. Deste modo, utilizando dados em painel para municípios gaúchos, os autores validam a relação do U invertido entre renda *per capita* e índice de Theil, no período de 1970, 1980 e 1991.

Jacinto e Tejada (2004) utilizam dados *cross-section* e em painel para os municípios da região Nordeste do Brasil levando em conta os anos de 1970 e 1991, também encontrando evidências da referida curva. De acordo com os autores, o fato de se obter resultados nem sempre favoráveis à existência de U invertido contribui para a especificação de diversas formas funcionais e do uso de diferentes métodos de estimação.

Salvato et al. (2006) analisam a relação entre desigualdade e desenvolvimento econômico para os municípios mineiros em dados *cross-section* e painel, nos anos de 1991 e 2000. Os autores utilizam dois índices de desigualdade: Gini e Theil-L. Os resultados obtidos corroboram o “U” invertido para dados em *cross-section* no ano de 1991 e efeitos fixos, para ambos indicadores de desigualdade.

Gomes et al. (2010) propõem uma análise alternativa da curva de Kuznets. Em contraste com trabalhos anteriores, os quais utilizam diretamente a renda *per capita*, os autores optam por realizar uma decomposição vetorial da renda *per capita* dos municípios nordestinos, desagregando-a em vetores representativos da renda provinda do trabalho e da

renda provinda de transferências governamentais. Deste modo, espera-se uma análise mais profunda dos fatores que influenciam a relação crescimento – equidade. Seguindo o trabalho de Huang et al. (2007), a hipótese de Kuznets é validada para diferentes quantis da distribuição da desigualdade. Os resultados também sugerem que a redução da desigualdade será mais eficiente se o crescimento econômico ocorrer através da geração de emprego e renda do que se for ocasionado pelo aumento das transferências governamentais.

O foco da literatura acerca do tema tem sido buscar evidências para a validação ou não da hipótese de Kuznets. Entretanto, boa parte dessa literatura negligenciou a questão da fragilidade da análise de regressão dada a arbitrariedade da escolha de variáveis de controle. Deste modo, os autores buscaram uma metodologia alternativa para a estimação dos coeficientes das variáveis de interesse.

### 3. Metodologia

*Model Averaging* é vista como uma abordagem agnóstica no sentido em que pesquisadores, ao empregarem esta técnica, não necessitam opinar sobre o melhor e único modelo. O grande objetivo dessa abordagem seria apresentar resultados que são robustos para diferentes suposições.

Esta literatura é composta por duas vertentes distintas, a *Bayesian Model Averaging* (BMA) e a *Frequentist Model Averaging* (FMA). A BMA foi introduzida por Leamer (1978), mas só foi bastante utilizada nos anos 2000. Do mesmo modo, a FMA foi alvo de atenção somente na última década, com Hjort e Claeskens (2003) e Hansen (2007).

Os estimadores FMA dependem crucialmente dos pesos selecionados para a estimação. Hansen e Racine (2009) propõem um estimador conhecido por *Jackknife Model Averaging*, o qual seleciona os pesos através da minimização do critério de *cross-validation*. Os autores mostram que o estimador é assintoticamente ótimo, por obter o menor erro quadrado esperado possível.

A presente seção propõe-se a descrever a forma pela qual se estimou todos os possíveis modelos candidatos e computou-se uma média ponderada de todas as estimativas dos coeficientes das variáveis de interesse. Dada a incerteza das variáveis explicativas que compõem o modelo, serão aqui estipulados seis cenários. Cada cenário terá como base um modelo econométrico diferente a ser estimado empregando todas as variáveis explicativas.

Na subseção 3.1 tem-se uma explicação pormenorizada da abordagem de *Frequentist Model Averaging* e do estimador de FMA. A subseção 3.2 apresenta o estimador de *Jackknife Model Averaging*, utilizado na ponderação da média dos parâmetros estimados. E, por fim, a subseção 3.3 descreve cada cenário.

#### 3.1 Estimador de FMA

A técnica utilizada segue as idéias da abordagem conhecida como *Frequentist Model Averaging* (FMA). Para ilustrar a definição do estimador de FMA, considere o modelo linear geral comumente empregado no estudo da hipótese de Kuznets:

$$y = (\mathbf{1} \text{ } r_{pc} \text{ } r_{pc}^2)' \mathbf{c} + \mathbf{X} \boldsymbol{\beta} + \mathbf{U} \quad (1)$$

onde  $y$ ,  $r_{pc}$ ,  $r_{pc}^2$  e  $U$  são vetores  $N \times 1$  da variável que mensura a desigualdade de renda (índice de Gini), da variável renda *per capita*, da variável renda *per capita* ao quadrado e de choques aleatórios, respectivamente.  $X$  é uma matriz  $N \times q$  de possíveis regressores, variáveis tais como média de anos de estudo, disposição geográfica, acesso a crédito e receita tributária, as quais podem ou não serem incluídas no modelo. O conjunto  $\mathbf{c}$ ,  $\boldsymbol{\beta}_1$ ,  $\boldsymbol{\beta}_2$ ,  $\boldsymbol{\varphi}$  ( $q \times 1$ ) contém os parâmetros a serem estimados. Observe que, de acordo com a hipótese de Kuznets,  $\boldsymbol{\beta}_1$  deve ser positivo e  $\boldsymbol{\beta}_2$  negativo. Por fim,  $N$  é o número de observações da amostra.

Se todos os componentes de  $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_q)'$  podem ser zero, existe um total de  $2^q$  modelos candidatos a serem estimados. Defina  $\hat{\beta}_{kM}$  o estimador de  $\beta_k$  do modelo candidato  $M$ , sendo  $M \in \{M_1, M_2, \dots, M_{2^q}\}$  e  $k = 1, 2$ . O estimador de FMA para  $\beta_k$  é determinado a partir de uma média ponderada de todos esses estimadores,  $\{\hat{\beta}_{kM_1}, \hat{\beta}_{kM_2}, \dots, \hat{\beta}_{kM_{2^q}}\}$ . Assim, por exemplo,  $\hat{\beta}_{kFMA} = \hat{\beta}_{kM_1}$  se o primeiro modelo receber peso 1. Deste modo, pode-se escrever os estimadores de FMA da seguinte forma:

$$\hat{\beta}_{kFMA} = \sum_{j=1}^{2^q} \omega_{M_j} \hat{\beta}_{kM_j}, \text{ onde } \omega_{M_j} \in [0,1] \text{ e } \sum \omega_{M_j} = 1.$$

### 3.2 A escolha da ponderação baseada no critério de cross-validation

Segundo Hansen e Racine (2009), para modelos heteroscedásticos e não-aninhados, deve-se selecionar pesos dos estimadores de mínimos quadrados de *Model Averaging* através da abordagem de *Jackknife Model Averaging* (JMA). A explicação vem do fato de que o JMA é apropriado para mais modelos lineares gerais, assim, os erros aleatórios podem ter variância heteroscedástica e os modelos candidatos podem ser não aninhados. Através de simulações de Monte Carlo, Hansen e Racine demonstram ganhos potenciais de eficiência no uso do desse estimador, dado que o JMA alcança erros quadrados médios menores que outros métodos alternativos, como o *Mallows Model Averaging* (MMA). Na presença de erros homocedásticos, JMA e MMA são quase equivalentes, porém quando os erros são heterocedásticos, os erros de JMA são menores que os de MMA.

$$\hat{\beta}(Y) = \sum_{j=1}^J \lambda_j \hat{P}_j Y = \hat{P}(Y)$$

O estimador JMA é dado por:

onde  $\hat{P}_j = \hat{D}_j(P_j - I_N) + I_N$ ,  $P_j = X_j(X_j'X_j)^{-1}X_j'$  é a projeção da matriz sob o j-ésimo modelo candidato,  $\hat{D}_j$  é uma matriz diagonal  $N \times N$  com o i-ésimo elemento da diagonal sendo  $(1 - \kappa_{ii}^j)^{-1}$ ,  $\kappa_{ii}^j = X_{j,i}(X_j'X_j)^{-1}X_{j,i}'$ , e  $X_{j,i}$  é a i-ésima linha de  $X_j$ . O critério de deleted-1 cross-validation é definido como:

$$CV(Y) = (Y - \hat{\beta}(Y))'(Y - \hat{\beta}(Y))$$

Os pesos são dados por:

$$\hat{\lambda} = \underset{\lambda \in H_N}{\operatorname{argmin}} CV(\lambda)$$

### 3.4 Cenários

O presente estudo aborda a questão da incerteza acerca de quantas e quais variáveis devem compor um modelo visando estimar o efeito do crescimento econômico na concentração de renda. Com isso, organizou-se a análise em seis cenários. Dado que o estudo se concentra na análise dos coeficientes  $\beta_1$  e  $\beta_2$ , esses parâmetros e o intercepto são estimados em todos esses cenários, como na equação (1). Entretanto, cada um deles contém um modelo econométrico com sua estrutura base distinta.<sup>1</sup> A diferenciação ocorre no que tange ao número de regressores  $m$  e no fato de que alguns regressores, além da renda *per*

<sup>1</sup> A definição desses cenários foi limitada pela capacidade de processamento do computador.

*capita* e o quadrado da renda *per capita*, estão presentes em todas estimativas<sup>2</sup>. Esses regressores foram escolhidos pelo fato de serem encontrados na literatura com uma maior frequência do que o restante.

### Cenário I

Neste primeiro cenário tem-se o seguinte modelo:

$$y_t = c + \beta_1 r_{pc,t} + \beta_2 r_{pc,t}^2 + \varphi_1 Z_{1,t} + \varphi_2 Z_{2,t} + \varphi_3 Z_{3,t} + \varphi_4 Z_{4,t} + \varepsilon_t, \quad (2)$$

onde  $y$  é o índice de Gini,  $c$  é a constante e  $Z_1, Z_2, Z_3, Z_4$  são variáveis explicativas incluídas no modelo, das  $q$  levantadas no estudo pertencentes ao conjunto de regressores  $X$ . Portanto, o modelo contém duas variáveis explicativas fixas (renda *per capita* e renda *per capita* ao quadrado) mais a constante e quatro regressores variáveis, regressores estes que podem ser escolhidos do total de  $q$ .

### Cenário II

O modelo do segundo cenário se assemelha bastante com o primeiro, porém diferencia-se no número de variáveis explicativas. Nesse caso o número de variáveis incluídas no modelo é 15. Tem-se, então:

$$y_t = c + \beta_1 r_{pc,t} + \beta_2 r_{pc,t}^2 + \varphi_1 Z_{1,t} + \varphi_2 Z_{2,t} + \dots + \varphi_{13} Z_{13,t} + v_t, \quad (3)$$

### Cenário III

O modelo do terceiro cenário apresenta-se da seguinte maneira:

$$y_t = c + \beta_1 r_{pc,t} + \beta_2 r_{pc,t}^2 + \varphi_1 edu_t + \varphi_2 dcorr_t + \varphi_3 ocr_t + \varphi_4 Z_{4,t} + \varphi_5 Z_{5,t} + \dots + \varphi_8 Z_{8,t} + w_t \quad (4)$$

Enquanto que nos cenários anteriores somente os regressores de interesse renda *per capita* e renda *per capita* ao quadrado e o intercepto eram fixos neste, entretanto, além das variáveis de interesse, as variáveis média de anos de estudo de pessoas com 25 anos ou mais ( $edu$ ), razão entre despesa corrente e PIB ( $dcorr$ ) e operações de crédito ( $ocr$ ) também são fixadas. Assim, tem-se  $Z_1 = edu$ ,  $Z_2 = dcorr$  e  $Z_3 = ocr$ . Já no que tange a regressores variáveis, o modelo apresenta 5 regressores, escolhido dentre os  $q-3$  restantes.

A escolha das variáveis fixadas seguiram os trabalhos de Ahluwalia (1976), Huang et al. (2007) e Beck et al. (2007) os quais utilizaram, além da renda *per capita* e renda *per capita* ao quadrado, *proxys* para o capital humano, tamanho do setor público e desenvolvimento financeiro, respectivamente.

### Cenário IV

O modelo do cenário se assemelha com o do cenário anterior, diferenciando-se no número de variáveis explicativas não-fixadas. Estas, para esse cenário, são de número 11. Tem-se:

$$y_t = c + \beta_1 r_{pc,t} + \beta_2 r_{pc,t}^2 + \varphi_1 edu_t + \varphi_2 dcorr_t + \varphi_3 ocr_t + \varphi_4 Z_{4,t} + \varphi_5 Z_{5,t} + \dots + \varphi_{11} Z_{11,t} + \pi_t \quad (5)$$

### Cenário V

<sup>2</sup> Em todos os casos, o número de variáveis incluídas no modelo é menor ou igual ao número total de variáveis levantadas no estudo,  $m \leq q$ .

No quinto cenário tem-se um modelo com 15 variáveis não-fixadas, semelhante com o modelo do segundo. No entanto, ao invés de ter-se somente três termos fixos, o presente modelo também apresenta as variáveis média de anos de estudo de pessoas com 25 anos ou mais, despesa corrente pelo PIB e operações de crédito fixadas. Portanto, tem-se:

$$y_i = \alpha + \beta_1 rpc_i + \beta_2 rpc_i^2 + \varphi_1 edu_i + \varphi_2 dcrr_i + \varphi_3 ocn_i + \varphi_4 Z_{1i} + \varphi_5 Z_{2i} + \dots + \varphi_{15} Z_{15i} + \zeta_i \quad (6)$$

### Cenário VI

O último cenário trata de todos os modelos aninhados que podem ser formados a partir dos  $q$  regressores em  $X$ , levando em consideração sua ordem de relevância na explicação da desigualdade. Nesse caso, a matriz  $X$  é organizada de tal forma que o regressor  $X_1$  é mais relevante do que o  $X_2$ , o  $X_2$  é mais relevante do que o  $X_3$ , e assim sucessivamente. Todos os modelos estimados seguem o seguinte procedimento. O primeiro subconjunto desses modelos é formado por equações do tipo  $y_i = \alpha + \beta_1 rpc_i + \beta_2 rpc_i^2 + \varphi_1 Z_i + w_{1i}$ , onde  $Z_i \in \{X_{1i}, \dots, X_{qi}\}$ , o segundo por  $y_i = \alpha + \beta_1 rpc_i + \beta_2 rpc_i^2 + \varphi_1 X_{1i} + \varphi_2 Z_i + w_i$ , onde  $Z_i \in \{X_{2i}, \dots, X_{qi}\}$ , o terceiro por  $y_i = \alpha + \beta_1 rpc_i + \beta_2 rpc_i^2 + \varphi_1 X_{1i} + \varphi_2 X_{2i} + \varphi_3 Z_i + w_i$ , onde  $Z_i \in \{X_{3i}, \dots, X_{qi}\}$ , e assim por diante até o modelo que inclua todos os  $q$  regressores.

Conforme mencionado anteriormente, o critério de relevância adotado foi baseado na frequência com que tais regressores são utilizados na literatura. Desta forma, *proxys* para o capital humano, tamanho do setor público, desenvolvimento financeiro, demografia e disposição geográfica são ditas as mais “relevantes”.

## 4. Base de Dados e Análise Descritiva das Variáveis

A base de dados empregada na pesquisa é obtida junto ao Ipeadata, vinculado ao Instituto de Pesquisa Econômica Aplicada (IPEA). Outros dados, por sua vez, são coletados do Atlas do Desenvolvimento Humano do Brasil, o qual é provido pelo Programa das Nações Unidas para o Desenvolvimento (PNUD). São coletadas informações referentes ao ano 2000, para os 1786 municípios nordestinos.

Dada a incerteza acerca de quantas e quais variáveis deveriam compor o modelo, os autores optaram por selecionar um grande número de variáveis as quais ele intuitivamente acreditava ter qualquer relação com a concentração de renda. Foi feito também um levantamento dos regressores utilizados nos principais estudos acerca do tema. Ao todo, obteve-se 52 variáveis, todas referentes ao ano 2000, que poderiam ou não compor os modelos.

De posse das 52 variáveis, optou-se por agrupá-las em indicadores de mesmo enfoque ou tema. Os indicadores eram de educação, finanças públicas, desenvolvimento financeiro, saúde, infraestrutura domiciliar, demografia e disposição geográfica. Como variáveis de um mesmo tema geralmente são bastante correlacionadas, poderia-se estar incorrendo no problema da multicolinearidade. Sendo assim, fez-se uso de matrizes de correlação para as variáveis de mesmo tema. Aquelas que apresentavam um alto grau de correlação com a maioria das outras variáveis de seu grupo foram escolhidas enquanto as outras foram descartadas. Deste modo, das 52 variáveis, restaram 22.

A seguir tem-se um quadro com todas as variáveis levantadas neste estudo.

VARIÁVEL	DESCRIÇÃO
gini	Mede o grau de desigualdade existente na distribuição de indivíduos segundo a renda domiciliar <i>per capita</i> .
rpc	Renda domiciliar <i>per capita</i> .
rpc <sup>2</sup>	Quadrado da renda domiciliar <i>per capita</i> .
edu	Razão entre o somatório do número de anos de estudo completados pelas pessoas que tem 25 ou mais anos de idade e o número de pessoas nessa faixa etária.
dcorr	Despesas correntes municipais referentes à manutenção e funcionamento de serviços públicos, pagamento de pessoal, aquisição de material de consumo, operação de escolas e de centros de saúde, dentre outras, em proporção do PIB.
rcorrPIB	Receitas correntes municipais tributárias, de contribuição patrimonial, industrial, agropecuária, de serviços, de transferências correntes, dentre outras, em proporção do PIB.
tribestPIB	Transferências correntes de tributos estaduais para os municípios em proporção do PIB.
ocr	Operações de crédito. Dados de captação e aplicação de recursos de agências bancárias - por município.
fnerpc	Valor <i>per capita</i> dos empréstimos do Fundo Constitucional de Financiamento do Nordeste (FNE).
credrural	Estoque de crédito rural. Valor referente à agregação do saldo contábil das operações de crédito realizadas pelas agências das instituições financeiras do município.
espvida	Expectativa de anos de vida de uma pessoa nascida no ano de referência supondo que as taxas de mortalidade por idade estimadas para anos anteriores se mantivessem constantes nos anos posteriores.
mortinf	Número de pessoas de cada mil nascidas vivas que não deverão completar um ano de vida
fecundidade	Número médio de filhos que uma mulher teria ao terminar o período reprodutivo.
energia	Percentual de pessoas que vivem em domicílios com iluminação elétrica, proveniente ou não de uma rede geral, com ou sem medidor.
dom_agua_enc_ban	Percentual de pessoas que vivem em domicílios com água encanada em pelo menos um de seus cômodos e com banheiro, definido como cômodo que dispõe de chuveiro ou banheira e aparelho sanitário.

VARIÁVEL	DESCRIÇÃO
dom_dens2	Percentual de pessoas que vivem em domicílios com densidade superior a 2. A densidade do domicílio é dada pela razão entre o total de moradores do domicílio e o número total de cômodos do mesmo, excluídos o(s) banheiro(s) e mais um cômodo, destinado à cozinha.
popurbana	População residente – urbana.
poprural	População residente – rural.
distcap	Distância à capital.
area	Área geográfica em km <sup>2</sup> .
densd	Medida de densidade populacional.
semi_arido	Variável binária que assume valor 1 se o município se localiza na região semi-árida.

#### Quadro 1 – Apresentação das variáveis

Fonte: Ipeadata (2011), Atlas do Desenvolvimento Humano (2000). Elaboração Própria.

No intuito de se desenvolver algum conhecimento de como se dá a evolução da concentração de renda dos municípios nordestinos, na medida em que se obtém valores cada vez maiores para cada variável, construiu-se uma tabela a qual contém a média do coeficiente de Gini dos municípios nordestinos referentes a diferentes quantis da distribuição de todas as variáveis. A exceção foi a variável semi\_arido, por se tratar de uma variável binária.

Tabela 1 - Média do índice de Gini dos municípios nordestinos correspondentes a diferentes quantis de cada variável – 2000

Média do índice de Gini						
VARIÁVEL	1º Decil	1º Quartil	2º Quartil	3º Quartil	4º Quartil	10º Decil
rendapc	0,59	0,58	0,58	0,58	0,59	0,59
estudo25	0,58	0,58	0,58	0,58	0,58	0,59
dcorrPIB	0,58	0,58	0,58	0,58	0,57	0,56
rcorrPIB	0,58	0,58	0,58	0,58	0,57	0,56
tribestPIB	0,58	0,58	0,58	0,58	0,58	0,60
opcr00	0,58	0,58	0,58	0,58	0,58	0,58
FNE00rpc	0,58	0,57	0,58	0,58	0,58	0,59
credrural	0,57	0,57	0,58	0,58	0,60	0,60
fecundidade	0,57	0,57	0,58	0,58	0,59	0,60
espvida00	0,58	0,57	0,58	0,58	0,58	0,57
mortinf00	0,57	0,58	0,58	0,58	0,58	0,59
energia	0,58	0,58	0,58	0,58	0,57	0,58
dom_agua_enc_ban	0,57	0,58	0,58	0,58	0,58	0,58
dom_dens2	0,57	0,57	0,58	0,58	0,58	0,58
Popurbana	0,56	0,56	0,58	0,58	0,59	0,60
Poprural	0,54	0,56	0,58	0,58	0,60	0,60
distcap	0,56	0,57	0,58	0,58	0,59	0,61
area00	0,56	0,56	0,58	0,58	0,60	0,61
densd00	0,59	0,58	0,58	0,58	0,57	0,57

Fonte: Ipeadata (2011). Elaboração Própria.

Percebe-se que a média do índice de Gini do primeiro decil é a mesma para o décimo decil, e que essa média cai para os quartis. Tal fato sugere que os municípios mais pobres possuem, em média, uma concentração de renda do mesmo nível dos municípios de renda per capita maiores, e que se tem uma curva mais parecida com um “U” propriamente dito do que com um “U invertido”. A Tabela 1 também sugere que quanto maior for a área do município, ou quanto mais longe for o município de sua capital estadual, mais concentrada tende ser a sua renda. Para níveis cada vez mais altos de fecundidade e mortalidade infantil, mais alta tende a ser a média do índice de Gini.

Por outro lado, a desigualdade tende a cair quanto maiores forem as receitas e despesas municipais correntes em proporção do PIB e a densidade populacional. A evolução da média do Gini diferencia-se nas variáveis relativas ao acesso ao crédito. A média permanece constante para os diferentes quantis de operações de crédito, enquanto que a renda tende a se concentrar mais para municípios que usufruem de maiores volumes de crédito rural.

Na sessão seguinte analisa-se o resultado da validação ou não da hipótese de Kuznets para cada tipo de modelo de todos os cenários estipulados.

## 5. Resultados

Nesta seção, as estimativas referentes aos coeficientes das variáveis de interesse são apresentadas. Espera-se que no caso da hipótese do “U invertido” ser validada,  $\hat{\beta}_1$  seja positivo e  $\hat{\beta}_2$  seja negativo. Qualquer resultado diferente questionaria a validação da hipótese.

O exame dos sinais dos coeficientes  $\beta_1$  e  $\beta_2$  considera, além dos casos descritos nos cenários I a VI, os modelos que apresentaram as maiores estatísticas da razão t e os menores valores para os critérios de Akaike e Schwartz, já que essas referências são muitas vezes utilizadas na procura de um modelo econométrico ideal em pesquisas similares. Para efeito de comparação, também é reportado os pesos que esses modelos recebem no cálculo do estimador de JMA.

Tendo em vista o grande número de modelos estimados em alguns cenários, e dada por finda a estimação da média ponderada dos coeficientes de interesse, são selecionados os modelos candidatos que receberam uma ponderação maior ou igual a 0,01 em cada cenário. Para esses modelos é analisada a frequência em que os sinais dos coeficientes  $\hat{\beta}_1$  e  $\hat{\beta}_2$  estão de acordo com a hipótese de Kuznets. Por fim, investigam-se o tamanho das razões t das estimativas pelo JMA e das estimativas nesses modelos, apenas a título de informação. Vale ressaltar que testes de hipóteses com base no estimador de JMA não são possíveis porque sua distribuição ainda é desconhecida.

### 5.1. Estimativas de *Jackknife Model Averaging*

A seguir são apresentadas as estimativas dos coeficientes das variáveis de interesse de todos os cenários, bem como suas respectivas razões t ponderadas.

Tabela 2 – Resultados obtidos pelos estimadores de FMA

Cenários	$\hat{\beta}_{1FMA}$	$\hat{\beta}_{2FMA}$
Cenário I	0,542814 6,60328	-0,14557 -0,51018
Cenário II	0,181704 2,36524	0,139616 0,52360

Cenário III	0,501135 6,22562	-0,03681 -0,13175
Cenário IV	0,258316 3,34674	0,106517 0,39760
Cenário V	-0,00245 -0,03202	0,289048 1,08840
Cenário VI	0,070567 0,91904	0,242413 0,90959

Elaboração Própria

Considerando apenas os sinais dos coeficientes, os estimadores de *JMA* sugerem que a validação da hipótese do “U invertido” ocorre somente no primeiro e terceiro cenário. No entanto, pode-se observar que as razões *t* calculadas com a ponderação obtida nesse método são muito baixas para o  $\beta_2$ . Apesar delas não poderem ser usadas para o teste de significância individual dos coeficientes, elas servem ao menos como indicativo de que esses coeficientes talvez sejam estatisticamente zeros. Se esse for o caso, a hipótese de Kuznets não seria sustentada em nenhum cenário.

Para o segundo, quarto e sexto cenário, os sinais de  $\beta_1$  e  $\beta_2$  estimados sugerem que o crescimento econômico proporciona um aumento da concentração de renda dos municípios nordestinos, e esse aumento dá-se a taxas crescentes. No quinto cenário, os sinais desses coeficientes estão invertidos em relação à proposta de Kuznets,  $\hat{\beta}_{1,FMA} < 0$  e  $\hat{\beta}_{2,FMA} > 0$ . Ou seja, o aumento da renda *per capita* estaria proporcionando uma queda na desigualdade, e essa queda estaria acontecendo a taxas cada vez maiores.

De forma geral, no que tange os valores das razões *t*, talvez somente o coeficiente  $\hat{\beta}_{1,FMA}$  poderia ser estatisticamente significativo em quase todos os cenários (considerando o nível de 5% de significância do teste *t* convencional). Nessa perspectiva, os resultados em conjunto indicariam apenas uma relação positiva entre os níveis de desigualdade da renda e renda *per capita*.

Tal resultado contrasta com o trabalho de Jacinto e Tejada (2004), o qual validou a curva de Kuznets para os municípios nordestinos para o ano de 1991.

## 5.2 Estimativas Seleccionadas pelos Critérios de Informação

A Tabela 3 apresenta as estimativas de  $\beta_1$  e  $\beta_2$  dos modelos seleccionados em cada cenário pelos critérios de informação de Akaike e Schwarz, bem como suas respectivas ponderações calculadas pelo critério de *cross-validation*. Em cada cenário o modelo sugerido pelo critério de Akaike coincidiu com o de Schwarz. A exceção foi o sexto cenário, o qual apresentou modelos diferentes.

Tabela 3 – Modelos seleccionados segundo os critérios de Akaike e Schwarz para cada cenário

Cenários	Peso	$\hat{\beta}_1$	$\hat{\beta}_2$	$t-\hat{\beta}_1$	$t-\hat{\beta}_2$	AIC	BIC
Cenário I	0	0,242676	1,029	1,45875	1,6178	-6,0675	-6,03653
Cenário II	0	0,216873	1,53605	0,82293	1,40411	-6,10042	-6,02052
Cenário III	0	0,216873	1,53605	0,82293	1,40411	-6,10042	-6,02052
Cenário IV	0,0005	0,233259	1,39306	0,89248	1,29194	-6,10062	-6,02516
Cenário V	0	0,214894	1,52301	0,81207	1,38415	-6,09527	-6,00206
Cenário VI AIC	0	0,209655	1,55595	0,79032	1,41571	-6,09522	-6,00201
Cenário VI BIC	0	0,241214	1,38373	0,90857	1,25981	-6,0897	-6,00537

Elaboração Própria

Em todos os seis cenários estipulados, os modelos sugeridos pelos critérios de informação possuíam ambos os coeficientes  $\hat{\beta}_1$  e  $\hat{\beta}_2$  positivos. No entanto, esses coeficientes não são estatisticamente significantes ao nível de significância de 5%. Deste modo, não houve modelo que validasse a Hipótese de Kuznets.

Observe que pelo método do *cross-validation*, esses modelos receberam ponderações nulas. A exceção foi o quarto cenário cujo modelo recebeu um peso não nulo, mas também muito pequeno. Isso implica que, quando consideramos um processo de seleção de modelos que vise o menor erro de previsão possível, os critérios de Akaike e Schwartz poderiam não ser apropriados na definição de modelos candidatos para averiguação da curva de Kuznets (empregando os dados do Nordeste).

### 5.3 Estimativas com Base nas Maiores Razões t

A seguir serão apresentados os modelos com as maiores estatísticas t para os dois coeficientes de interesse. Vale mencionar que, enquanto os testes de hipóteses com base nos estimadores de FMA  $\hat{\beta}_{1,FMA}$  e  $\hat{\beta}_{2,FMA}$  não são possíveis pois sua distribuição ainda é desconhecida, os mesmos não ocorrem com os coeficientes  $\hat{\beta}_1$  e  $\hat{\beta}_2$  dos modelos individualmente selecionados em cada cenário pelos diferentes critérios anteriormente citados.

Tabela 4 – Modelos com as maiores estatísticas  $t-\hat{\beta}_1$  para cada cenário

Cenários	Peso	$\hat{\beta}_1$	$\hat{\beta}_2$	$t-\hat{\beta}_1$	$t-\hat{\beta}_2$	AIC	SIC
Cenário I	0,00019	0,915225	-0,43482	4,563832	-0,61931	-5,98905	-5,95798
Cenário II	0	0,702977	-0,34561	3,498519	-0,49633	-6,06393	-5,98404
Cenário III	0,00112	0,873311	-0,57224	4,258649	-0,80798	-5,98741	-5,93858
Cenário IV	0	0,838255	-0,55788	4,162914	-0,79386	-6,04284	-5,96738
Cenário V	0	0,372539	1,082406	1,404297	0,977308	-6,07758	-5,98437
Cenário VI	0	0,881053	-0,68933	4,35181	-0,9753	-6,03009	-5,95908

Elaboração Própria

A Tabela 4 reúne os modelos que apresentaram as maiores estatísticas t para o primeiro coeficiente. Percebe-se que os modelos receberam valores muito pequenos de ponderação. A maior ponderação foi dada ao modelo do terceiro cenário, 0,00112. Com exceção do modelo do quinto cenário, todos os outros validaram a hipótese do “U” invertido. No entanto, todos os modelos apresentam o segundo coeficiente estatisticamente insignificante.

Já a Tabela 5 reúne os modelos que apresentaram as maiores estatísticas para o segundo coeficiente. Ao contrário dos modelos de maiores estatísticas  $t-\hat{\beta}_1$ , nenhum modelo com as maiores estatísticas  $t-\hat{\beta}_2$  validou a Hipótese de Kuznets. A ponderação dada aos modelos foi muito pequena, tais como no primeiro caso. Conclui-se que mesmo quando o pesquisador procurasse um modelo com base na estatística t dos coeficientes de interesse, ainda não seria possível obter Kuznets.

Tabela 5 – Modelos com as maiores estatísticas  $t-\hat{\beta}_2$  para cada cenário

Cenários	Peso	$\hat{\beta}_1$	$\hat{\beta}_2$	$t-\hat{\beta}_1$	$t-\hat{\beta}_2$	AIC	SIC
Cenário I	0	-0,90466	-0,90406	4,376522	6,126264	-5,89939	-5,86832

Cenário II	0	-0,61764	4,5603	-2,42541	4,158363	-5,99828	-5,91839
Cenário III	0,00002	-0,69884	5,154983	-3,08073	5,846893	-5,90064	-5,85182
Cenário IV	0	-0,71254	4,813998	-2,79271	4,369123	-5,98629	-5,91083
Cenário V	0	-0,18518	2,725332	-0,68829	2,419437	-6,03511	-5,9419
Cenário VI	0	-0,80929	5,020379	-4,04575	5,782084	-5,86937	-5,85162

Elaboração Própria

#### 5.4 Análise de modelos candidatos com pesos acima ou igual a 0,1%

Nas seções anteriores observou-se os coeficientes das variáveis de interesse de todos os modelos candidatos para cada cenário. De acordo com diferentes critérios, pôde-se verificar quais modelos validavam a hipótese de Kuznets e a ponderação recebida de acordo com o critério de *cross-validation*. No entanto, os autores também tiveram a intenção de verificar o sinal dos coeficientes beta dos modelos com pesos igual ou acima de 0,1%, bem como as respectivas estatísticas t para testes de significância, ao nível de significância de 5%.

Para o primeiro cenário, apenas cerca 35% desses modelos candidatos apresentaram o sinal do  $\hat{\beta}_1$  condizente com a hipótese de Kuznets, e somente 37% dos coeficientes dos modelos com ponderação mínima apresentaram-se estatisticamente significantes ao nível de significância de 5%. Já no que tange ao coeficiente  $\hat{\beta}_2$ , 99% destes apresentaram-se positivos, e 30% dos coeficientes são ditos estatisticamente significantes.

Para o segundo cenário, tem-se que cerca de 75% destes coeficientes são positivos. No entanto, apenas 3% do total são considerados estatisticamente significantes. As estimativas do segundo coeficiente revelam que todos os coeficientes apresentaram-se positivos e que 28% são considerados estatisticamente significantes.

No terceiro cenário, 45% dos modelos possuem o sinal requerido para o coeficiente estimado  $\hat{\beta}_1$ . Contudo, somente 25% dos coeficientes estimados podem ser considerados estatisticamente significantes. Para o segundo coeficiente, apenas 1% apresentou o sinal requerido para a validação da hipótese de Kuznets. 68% dos coeficientes são considerados estatisticamente significantes.

No quarto cenário percebe-se que cerca de 65% dos modelos possuem o sinal requerido do primeiro coeficiente para a validação da hipótese. Apenas 11% dos estimadores  $\hat{\beta}_1$  dos modelos de ponderação mínima para esse cenário são considerados estatisticamente significantes. Como ocorreu em cenários anteriores, somente 1% dos estimadores  $\hat{\beta}_2$  apresentaram-se negativos. E apenas 37% podem ser ditos estatisticamente significantes.

No quinto cenário, os dois estimadores  $\hat{\beta}_1$  e  $\hat{\beta}_2$  apresentaram-se positivos e estatisticamente insignificantes ao nível de 5% de significância para os modelos com ponderação de 0,1% ou mais.

Por fim, no sexto cenário, para os modelos selecionados, 83% dos coeficientes da renda *per capita* apresentaram-se positivos, enquanto que 33% dos coeficientes da renda *per capita* ao quadrado apresentaram-se negativos. No que tange as estatísticas t, 50% dos coeficientes  $\hat{\beta}_1$  podem ser considerados estatisticamente significantes, contra apenas 17% do segundo coeficiente.

## 6. Considerações finais

O presente estudo teve como objetivo buscar evidências para a validação, ou não, da hipótese de Kuznets para os municípios nordestinos. No entanto, ao contrário de grande parte

da literatura acerca do tema, buscou-se uma metodologia alternativa para a estimação dos coeficientes das variáveis de interesse.

Considerando a incerteza acerca do processo de seleção do modelo, estipulou-se seis cenários. Estes continham tipos de modelos diferentes não só no número de variáveis, mas também por conter alguns regressores fixos em todos os modelos, e outros que poderiam ou não estar presente. Propôs-se então estimar todos os possíveis modelos candidatos e computar uma média ponderada de todas as estimativas dos coeficientes das variáveis de interesse, abordagem conhecida como *Model Averaging*. A escolha da ponderação foi baseada no critério de *cross-validation*, com a técnica conhecida como *Jackknife Model Averaging*.

Para a seleção das variáveis, os autores tiveram a preocupação de selecionar um grande número de variáveis as quais ele intuitivamente acreditava ter qualquer relação com a concentração de renda. Outras foram selecionadas por estarem presentes em muitos trabalhos anteriormente citados. Dado que muitas vezes variáveis de um mesmo tema se encontram correlacionados, fez-se uso de matrizes de correlação. Aquelas que apresentavam um alto grau de correlação com a maioria das outras variáveis de seu grupo foram escolhidas, enquanto as outras foram descartadas.

De posse dos cenários e das variáveis incertas, chegou-se aos resultados. Dos seis cenários estipulados, somente em dois a curva de Kuznets foi validada. O cenário I foi um dos que apresentaram os coeficientes de FMA condizente com uma curva côncava. No entanto, somente em 35% dos modelos de peso igual ou superior a 0,1% o coeficiente estimado  $\hat{\beta}_1$  apresentou-se positivo. Já para o segundo coeficiente estimado, quase nenhum destes apresentou o  $\hat{\beta}_2 < 0$ . O cenário III também validou hipótese de Kuznets. Semelhante ao Cenário I, 45% dos modelos de peso igual ou superior a 0,1% apresentaram o coeficiente estimado  $\hat{\beta}_1$  positivo. E para o segundo coeficiente, quase nenhum destes apresentou-se negativo também.

Tal fato sugere que os modelos cujos coeficientes estavam de acordo com a referida curva devem ter recebido uma ponderação maior. Ou então que a maioria dos modelos que receberam ponderação quase nula apresentou coeficientes condizentes com a curva, a ponto de compensar a sua pequena ponderação no cômputo da média.

No que tange aos cenários que não apresentaram resultados condizentes com a curva de Kuznets, a justificativa deu-se apenas pelo fato de ter-se  $\hat{\beta}_{2,FMA} > 0$ . A exceção foi o quinto cenário, no qual  $\hat{\beta}_{2,FMA}$  era negativo. Contudo, em todos os cenários os quais apontaram a não-validação da curva o coeficiente estimado  $\hat{\beta}_{2,FMA}$  era maior que zero. Verificou-se que em 3 cenários dos 6 estipulados há evidências que o crescimento econômico proporciona um aumento da concentração de renda, e que tal aumento cresce a taxas positivas. Até mesmo em cenários que apresentaram os coeficientes de FMA condizente com uma curva côncava, a maior parte dos modelos apresentou  $\hat{\beta}_2 > 0$ .

Dependendo do critério de seleção do modelo, a hipótese pode ser rejeitada ou não. Nenhum modelo sugerido pelos critérios de informação de Akaike e Schwarz validou a curva de Kuznets. De todos os modelos sugeridos por esses critérios, nenhum destes recebeu uma ponderação maior que 0,1%. Na verdade, quase todos receberam ponderação nula.

Este estudo contrasta com os trabalhos anteriores de Jacinto e Tejada (2004), Salvato et al. (2006) e Bangolin, Gabe e Ribeiro (2003). Ao contrário desses estudos, não se tem evidências da curva de Kuznets. Os resultados sugerem que, para o caso nordestino, o crescimento econômico proporciona um aumento da concentração de renda dos municípios nordestinos, e esse aumento dá-se a taxas crescentes. Provavelmente a economia nordestina ainda não atingiu certo grau de dinamismo, o qual já atravessou o período inicial do

crescimento econômico acompanhado por aumento da desigualdade. A alta concentração do capital físico e humano ainda é uma realidade nordestina.

## 7 Referências

AHLUWALIA, M. S. Income distribution and development: some stylized facts *American Economic Review*, v. 66, p. 128-135, 1976.

ARNAND, S.; KAMBUR, S. M. R. Inequality and development: a critique. *Journal of Development Economics*, v. 41, p. 19-43, 1993.

**Atlas do Desenvolvimento Humano no Brasil**. 2003. Fundação João Pinheiro (FJP-MG).

BANGOLIN, I.; GABE, J.; RIBEIRO, E. Crescimento e desigualdade no Rio Grande do Sul: uma revisão da Curva de Kuznets para os municípios gaúchos (1970-1991). **Anais do 2º Encontro de Economia Gaúcha**. Porto Alegre, FEE, v. 1, 2004.

BECK, T.; DEMIRGÜÇ-KUNT, A.; LEVINE, R. Finance, Inequality and Poor. *Journal of Economic Growth*, v.12, p. 27-49, 2007.

BÊRNI, D. A.; MARQUETTI, A.; KLOECKNER, R. A Desigualdade Econômica do Rio Grande do Sul: Primeiras Investigações sobre a Curva de Kuznets. **Anais do 1º Encontro de Economia Gaúcha**. Porto Alegre, FEE, 2002.

CLAESKENS, G.; HJORT, N. Model Selection and Model Averaging, **Cambridge University Press**, 2008.

FIELDS, G. S.; JACKUBSON, G. H. New evidence on the Kuznets curve. (Mimeo). **Cornell University**, 1994.

GOMES, P. E. M.; PENNA, C. M.; TROMPIERI NETO, N.; LINHARES, F. Trabalho, Transferências do Governo, Desigualdade: A Curva de Kuznets para o Nordeste. **Anais do VIII Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos**, Juiz de Fora, 2010.

GREENE, W. H. **Econometric Analysis**. 5 ed. New Jersey: Prentice Hall, 2003.

GUJARATI, D. **Econometria Básica**. 4 ed. Rio de Janeiro: Campus, 2006.

HANSEN, B.; RACINE, J. **Jackknife Model Averaging**. Manuscrito não publicado, 2009.

(RIVER) HUANG, H. C.; LIN, S. C.; SUEN, Y. B.; YEH, C. C. A Quantile Inference of the Kuznets Hypothesis, **Economic Modelling**, Elsevier, v. 24, n. 4, p. 559-570, jul, 2007.

IPEADATA. **Regional**. Disponível em: < <http://www.ipeadata.gov.br/Default.aspx> > Acesso em: 24 fev. 2011.

JACINTO, P. A.; TEJADA, C. A. O; FIGUEIREDO, E. Desigualdade de renda e crescimento econômico para o Nordeste do Brasil: o que os dados têm a dizer? 2004, Mimeo.

KMENTA, J. **Elementos de Econometria**. 1 ed. São Paulo: Atlas, 1990.

KUZNETS, S. Economics Growth and Income Inequality. **American Economic Review**. V. 45, p. 1-28, 1955.

LEAMER, E.; LEONARD, H. Reporting the Fragility of Regression Estimates. **Review of Economics and Statistics**, v.65, p. 306-317, 1983.

MORAL-BENITO, E. Model Averaging in Economics. **Working Papers**. wp2010\_1008, CEMFI, 2010

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.

SALVATO, M. A.; ALVARENGA, P. S.; FRANÇA, C. S.; ARAÚJO JÚNIOR, A. F. Crescimento e Desigualdade: Evidências da Curva de Kuznets para os Municípios de Minas Gerais – 1991/2000. **Revista Economia & Gestão da PUC-Minas**, v.6, n.13, 2006.

TANG, K. K.; LIM, A. S. K.; Education Inequality, Human Capital Inequality and the Kuznets Curve," **MRG Discussion Paper Series 0506**, School of Economics, University of Queensland, Australia.