

O capital humano nos municípios paranaenses: uma análise com regressões quantílicas

Kassya Christina Keppe*

Luciano Nakabashi**

RESUMO – O presente trabalho trata-se do capital humano como um dos fatores determinantes para explicar a diferença no nível e crescimento do PIB por trabalhador nos municípios paranaenses. Muitos estudos mostraram que realmente existe uma relação positiva entre o capital humano e o crescimento da renda. Porém, o presente estudo mostra o efeito do capital humano nos diferentes quantis de renda, no período de 1980 a 2000. A análise empírica será realizada através de uma comparação entre os métodos de Regressões Quantílicas (LAD) e Mínimos Quadrados Ordinários (MQO).

Palavras-chave: Capital humano. Municípios paranaenses. Regressões quantílicas.

1 INTRODUÇÃO

O estado do Paraná apresenta grande disparidade na renda dos seus diferentes municípios. Muitos fatores podem explicar o crescimento da renda e, portanto, essa disparidade de renda entre os municípios, sendo que o capital humano é um dos possíveis candidatos. Isso ocorre porque a informação, o conhecimento e o grau de habilidade dos trabalhadores são fatores muito importantes no processo de produção, inovação e difusão de tecnologia.

Visando mensurar o impacto do capital humano no crescimento da renda, Mankiw, Romer e Weil¹ (1992), incluíram o capital humano no modelo de crescimento endógeno de Solow (1956). Desde então, muitos estudos empíricos comprovaram esta relação.

O objetivo do presente estudo é apresentar os elementos que podem explicar a grande diferença de renda existente entre os municípios paranaenses para os anos de 1980, 1991 e 2000, utilizando o Método de Regressões Quantílicas, pois este permite analisar o impacto do capital humano em cada quantil da renda.

* Economista pela Universidade Federal do Paraná (UFPR). Endereço eletrônico: kkeppe@hotmail.com.

** Doutor em economia pelo Centro de Desenvolvimento e Planejamento Regional da Universidade Federal de Minas Gerais (CEDEPLAR/UFMG), professor do Departamento de Economia da Universidade Federal do Paraná (UFPR) e coordenador do Boletim Economia & Tecnologia. Endereço eletrônico: Luciano.nakabashi@ufpr.br.

¹ MRW doravante.

Além desta introdução, o estudo apresenta, na segunda seção a metodologia e a base de dados utilizados na análise empírica. Na terceira seção, fazemos uma análise dos resultados das regressões estimadas, sendo esta seguida por uma breve conclusão.

2 METODOLOGIA E DADOS

2.1. O MÉTODO DE REGRESSÕES QUANTÍLICAS

O método de Regressões Quantílicas fornece uma estrutura linear conveniente para examinar como os quantis de uma variável dependente muda em resposta a um conjunto de variáveis independentes. Dito de outro modo, a regressão quantílica permite a estimativa linear de funções quantílicas condicional.

Diferentemente do método de Mínimos Quadrados Ordinários², que minimiza a soma dos quadrados dos resíduos, o método de Regressões Quantílicas, desenvolvido por Koenker e Basset (1978), minimiza a soma dos desvios absolutos, sendo, por esse motivo, também conhecido com estimador *Least Absolute Deviations*³.

Supondo a seguinte equação:

$$y_i = x_i' \beta + u_{qi} \quad (1)$$

onde:

y_i é o vetor da variável dependente;

x_i' é o vetor de variáveis explanatórias;

β é o vetor de coeficientes;

μ_{qi} é o resíduo para o q-ésimo quantil.

O valor mínimo dos desvios absolutos $(y_i - x_i' \beta)$ é dado pela solução do problema:

$$\min_{\beta} \sum_{t=1}^n |y_t - x_t' \beta| \quad (2)$$

Esta fórmula pode ser generalizada para que o coeficiente β se ajuste a qualquer

² MQO daqui em diante.

³ LAD doravante.

outro quantil da distribuição da variável dependente, tem-se:

$$\begin{aligned} \min_{\beta} \frac{1}{n} \left\{ \sum_{i \in \{i::y_i \geq x_i' \beta\}} q |y_i - x_i' \beta| + \sum_{i \in \{i::y_i < x_i' \beta\}} (1-q) |y_i - x_i' \beta| \right\} = \\ \min_{\beta} \frac{1}{n} \sum \rho_q(y_i - x_i' \beta) = \min_{\beta} \frac{1}{n} \sum \rho_q(u_{qi}) \end{aligned} \quad (3)$$

onde a função *check*, $\rho_q(\mu_{qi})$, é definida como:

$$\begin{cases} \rho_q(\mu_{qi}) = q\mu_{qi}; \text{ se } \mu_{qi} \geq 0 \\ \rho_q(\mu_{qi}) = (q-1)\mu_{qi}; \text{ se } \mu_{qi} < 0 \end{cases} \quad q \in [0,1] \quad (4)$$

A variável dependente y é ordenada de forma crescente, ou seja, quanto mais elevado for o quantil – quanto maior for q – mais alto é o valor da variável y dos indivíduos presentes nesse quantil.

O modelo LAD generalizado tem por objetivo minimizar, dada a variável dependente y e suas variáveis explanatórias x , a soma dos desvios absolutos para cada quantil da variável y . Desse modo, tem-se que a interpretação dos coeficientes é análoga a do método de MQO.

Enquanto este último nos fornece, como resultado, coeficientes que medem a influência das variáveis explanatórias sobre a média condicional da variável dependente, no método de regressões quantílicas, os coeficientes representam a influência das variáveis explanatórias sobre a variável dependente condicional ao q -ésimo quantil. Em outras palavras, a análise através de regressões quantílicas permite analisar toda a distribuição enquanto o método de MQO só permite analisar a média da relação entre as variáveis.

A função *check*, nada mais é do que a solução para um problema que minimiza a soma ponderada dos desvios absolutos. Desse modo, para quantis acima da mediana, digamos $\theta = 0,75$, um maior peso é colocado sobre resíduos acima da mediana que em resíduos abaixo da mediana. Estes ajustes empurram a minimização acima da mediana, e impedem os desvios acima da mediana a se minimizarem mais que os desvios abaixo dela, tornando a estimação mais aderente ao quantil 0,75.

O estimador LAD generalizado se mostra mais adequado nos caso em que o pesquisador está interessado em analisar partes específicas da distribuição condicional da variável dependente.

No presente estudo, a estimação do modelo foi realizada utilizando o *software* estatístico *Stata* 10.0 para os quantis 0,25; 0,50; 0,75 e 0,90. Desse modo, têm-se resultados para quatro diferentes níveis de renda dos municípios. A regressão do quantil 0,25 nos fornece resultados aplicados aos municípios com os 25% menores rendimentos da amostra; já a regressão do quantil 0,90 nos fornece resultados para os municípios com os 90% maiores rendimentos da amostra. Análoga é a interpretação dos resultados para as regressões dos quantis 0,50 e 0,75.

2.2 A BASE DE DADOS

Neste modelo serão utilizadas *proxies* de estoques, consumo de energia não residencial por trabalhador para a variável de capital físico, e anos de escola para a variável de capital humano. Todas as variáveis foram divididas pelo número de trabalhadores, que é a população com mais de 15 anos de idade.

O PIB dos municípios foi obtido junto à base de dados do IPARDES e IBGE, e se encontra em preços constantes de 2000. Os dados de consumo de energia não residencial, em MWh, também foram obtidos junto ao IPARDES, sendo o consumo de energia não residencial a diferença entre o consumo total de energia e consumo residencial em cada município. Esta é a variável utilizada como *proxy* para o estoque de capital físico, por faltarem dados sobre investimentos nos municípios estudados.

Os dados pra anos de estudo foram obtidos junto ao IPEA. Esta variável é utilizada como *proxy* para capital humano. Os dados para a população com mais de 15 anos são provenientes do IBGE e IPEA, e todas as demais variáveis que estão em nível por trabalhador foram obtidas com base nesta variável.

Quando trabalhamos com dados no estado estacionário, a base de dados é composta por três períodos: 1980, 1991 e 2000. No caso do estado de convergência, utilizamos os períodos de 1991 e 2000, e a diferença do PIB por trabalhador é composta por $\ln(1991) - \ln(1980)$ e $\ln(2000) - \ln(1980)$. Adicionalmente, o PIB por trabalhador inicial é para o ano de 1980.

A base de dados é composta por 290 municípios em 3 períodos de análise. Como o número de municípios em 1980 era menor do que nos outros anos (1991 e 2000), os municípios que foram desmembrados a partir de 1980, foram somados aos municípios de origem. As variáveis que já estavam em média, como a variável anos de estudo, foram

encontradas a partir de uma média ponderada do município de origem com os municípios que foram desmembrados.

3 RESULTADOS

Na Tabela 1, podemos ver os resultados para o ano de 1980. Em toda a análise para este período é feita a suposição de que os municípios se encontram no estado estacionário⁴. Nas quatro primeiras colunas, os resultados foram obtidos através do método LAD para os quantis 0,25; 0,50; 0,75 e 0,99. Na coluna seguinte foi utilizado o método de MQO.

A variável de capital humano apresenta o seu maior valor no quantil 0,25, onde 1% de aumento nesta variável reflete um incremento de 0,463% na renda do município. Como a média de anos de estudo neste período é de 2,38 anos, um ano a mais de estudo leva a um aumento de 42,07% na quantidade de capital humano. Considerando que a elasticidade da renda é de 0,463 no quantil 0,25, isso implica que para os municípios que estão entre os 25% com menor renda, um ano a mais de estudo resulta em um aumento na renda de 19,47%. Somente no quantil 0,99, o coeficiente desta variável apresenta um sinal negativo.

TABELA 1 – RESULTADOS DAS REGRESSÕES PARA 1980

Quantil	0,25	0,50	0,75	0,99	MQO
H	0,463 (0,083)**	0,308 (0,078)**	0,222 (0,089)**	-0,559 (0,053)**	0,287 (0,084)**
K	0,201 (0,026)**	0,290 (0,025)**	0,304 (0,028)**	0,771 (0,119)**	0,281 (0,027)**
C	1,497 (0,093)**	1,901 (0,088)**	2,142 (0,099)**	4,404 (0,129)**	1,910 (0,094)**
N	290	290	290	290	290
R²	0,242	0,300	0,342	0,3370	0,475

NOTA: erro padrão está entre parêntese. * significativo ao nível de 5%. ** significativo ao nível de 1%. A variável dependente é renda por trabalhador. H é anos de escola, K é o consumo de energia não residencial por trabalhador, C é a constante, N é o tamanho da amostra.

Na Tabela 2, podemos ver o resultado da regressão para o ano de 1991. Nas quatro primeiras colunas é feita a suposição de que o município está no estado estacionário. Desse modo, testamos a adequação ao método LAD com os quantis 0,25; 0,50; 0,75 e 0,99. Na quinta coluna testamos a adequação ao método de MQO, com os municípios no estado estacionário.

⁴ Para um estudo da forma funcional utilizada no presente estudo ver Mankiw, Romer e Weil (1992).

Na sétima, oitava, nona e décima coluna, assumimos que os municípios estão no estado de convergência. É utilizado o método LAD com os quantis 0,25; 0,50; 0,75 e 0,99. Na décima primeira coluna, apresentamos os resultados da estimação por MQO, com os municípios no estado de convergência.

Considerando que os municípios estão no estado estacionário, podemos ver que o capital físico é mais importante para explicar os níveis de renda por trabalhador dos diferentes municípios. Para todos os quantis de renda, os coeficientes são positivos e significativos a 1%. O capital físico tem um coeficiente maior, para o quantil 0,75, onde um aumento de 1% no consumo de energia elétrica não residencial leva a um incremento de 0,532% no PIB por trabalhador do município. Os coeficientes da variável de capital humano, não são significativos a 5% para nenhum dos quantis.

TABELA 2 - RESULTADOS DAS REGRESSÕES PARA 1991

Quantil	0,25	0,50	0,75	0,99	MQO(1)	Convergência				MQO (2)
						0,25	0,50	0,75	0,99	
H	0,198 (0,149)	0,169 (0,163)	-0,116 (0,129)	0,015 (0,126)	0,115 (0,113)	-0,122 (0,140)	-0,068 (0,157)	-0,200 (0,161)	-0,010 (0,068)	-0,113 (0,110)
K	0,461 (0,050)**	0,487 (0,048)**	0,532 (0,030)**	0,293 (0,011)**	0,457 (0,034)**	0,412 (0,053) **	0,369 (0,049) **	0,419 (0,037) **	0,204 (0,011) **	0,371 (0,034) **
C	1,422 (0,199)**	1,707 (0,214)**	2,120 (0,166)**	2,615 (0,145)**	1,743 (0,148)**	1,202 (0,200) **	1,250 (0,211) **	1,636 (0,215) **	2,245 (0,029) **	-0,659 (0,049) **
Y80						-0,655 (0,077) **	-0,620 (0,070) **	-0,630 (0,067) **	-0,818 (0,029) **	1,384 (0,147) **
N	290	290	290	290	290	290	290	290	290	290
R ²	0,224	0,296	0,322	0,403	0,377	0,231	0,229	0,220	0,274	0,442

NOTA: erro padrão está entre parêntese. * significativo ao nível de 5%. ** significativo ao nível de 1%. A variável dependente é renda por trabalhador nas quatro primeiras regressões, nas regressões seguintes a variável dependente é a variação da renda por trabalhador. H é anos de escola, K é o consumo de energia não residencial por trabalhador. C é a constante, Y80 é o nível de renda por trabalhador em 1980 e N é o tamanho da amostra.

Na Tabela 3, podemos ver os resultados das regressões para o ano 2000. Nas quatro primeiras colunas é feita a suposição de que os municípios estão no estado estacionário. As quatro primeiras colunas foram testadas através do Método LAD, com os quantis 0,25; 0,50; 0,75 e 0,99. Na quinta coluna testamos a adequação ao método de MQO, com os municípios no estado estacionário. Da sétima coluna em diante, assumimos que os municípios se encontram em estado de convergência.

Para o ano de 2000, os resultados se mostraram muito mais expressivos, Com os municípios no estado estacionário, a variável de capital humano tem um efeito maior sobre a renda. O maior coeficiente desta variável se dá no quantil 0,50 (mediana), onde o aumento de 1% na variável, reflete em um aumento de 0,536% na renda do município. Como a média de anos de escola neste período é de 4,67, um ano a mais de escola representa um aumento de 21,41% na quantidade de capital humano. Como a elasticidade da renda em relação ao capital humano é de 0,536%, em média, um ano a mais de estudo para os municípios que estão se quantil reflete em um aumento na renda de 11,47%.

TABELA 3 - RESULTADOS DAS REGRESSÕES PARA 2000

Quantil	0,25	0,50	0,75	0,99	MQO(1)	Convergência				MQO(2)
						0,25	0,50	0,75	0,99	
H	0,242 (0,105)*	0,536 (0,124)**	0,506 (0,148)**	0,411 (0,308)	0,453 (0,123)**	0,358 (0,164) *	0,365 (0,125) **	0,224 (0,175)	-1,700 (0,197) **	0,281 (0,128*)
K	0,300 (0,026)**	0,295 (0,029)**	0,364 (0,031)**	0,252 (0,030)**	0,285 (0,029)**	0,197 (0,038) **	0,252 (0,027) **	0,336 (0,038) **	0,527 (0,024) **	0,252 (0,030)**
C	1,376 (0,162)**	1,065 (0,190)**	1,298 (0,227)**	2,230 (0,534)**	1,232 (0,188)**	0,893 (0,233) **	0,969 (0,180) **	1,363 (0,263) **	4,031 (0,274) **	1,192 (0,184)**
Y80						-0,830 (0,067) **	-0,804 (0,043) **	-0,800 (0,058) **	-0,286 (0,046) **	-0,832 (0,044)**
N	290	290	290	290	290	290	290	290	290	290
R ²	0,230	0,246	0,237	0,137	0,377	0,396	0,336	0,259	0,218	0,570

NOTA: erro padrão está entre parênteses. * significativo ao nível de 5%. ** significativo ao nível de 1%. A variável dependente é a renda por trabalhador nas quatro primeiras regressões, nas regressões seguintes a variável dependente é a variação da renda por trabalhador. H é anos de escola, K é o consumo de energia não residencial por trabalhador. C é a constante, Y80 é o nível de renda por trabalhador em 1980 e N é o tamanho da amostra.

4 CONCLUSÃO

A principal conclusão do presente estudo é que ocorreu uma grande variação na importância do capital humano em cada quantil de 1980 até 2000. Enquanto o capital humano era mais importante para os níveis de renda mais baixos em 1980, este passou a ser mais relevante para os níveis mais elevados de renda em 2000, principalmente para o quantil 0,50.

Desse modo, os resultados apresentados indicam que esse fator tem se tornado mais relevante para as classes de renda mais elevadas. Ou seja, a educação formal tem apresentado maior interação com cargos de renda mais elevada que, em média, são mais complexos.

REFERÊNCIAS

KOENKER, R. HALLOCK, K. F. Quantile Regression. **Journal of Economic Perspectives**, v. 15, n. 4, p. 143–156, 2001.

KOENKER, R.; BASSET, G. Regression Quantiles. **Econométrica**, n. 46, 1978.

MANKIW, N.G. ROMER, D. WEIL, D.N. (1992), A Contribution to the empirics of Economic Growth, **The Quarterly Journal of Economics**, vol. 107, no. 2, pp. 407-37, 1992.

SOLOW, R. M. A contribution to the theory of economic growth. **The Quarterly Journal of Economics**, v.70, n.1, p.65-94, 1956.