

# Análise do dividendo demográfico na matrícula escolar no Brasil numa abordagem hierárquica e hierárquica-espacial

Juliana de Lucena Ruas Riani\*  
Eduardo Luiz Gonçalves Rios-Neto\*\*

*Esse artigo procura investigar os determinantes da matrícula escolar no Brasil, nos ensinos fundamental e médio, combinando duas tradições dos estudos educacionais. A primeira refere-se ao dividendo demográfico e a segunda aos estudos dos determinantes educacionais que utilizam o arcabouço teórico da Função de Produção Educacional. Somado a isso, busca-se incluir a dinâmica espacial na análise educacional, através da construção de uma metodologia alternativa, que combina os modelos hierárquico e espacial. Os principais resultados indicaram que uma menor pressão demográfica constitui um dos fatores contextuais mais importantes para a matrícula no fundamental. Já para o ensino médio, a disponibilidade de oferta educacional, representada pela razão professor/população em idade escolar, é o fator municipal mais relevante. Quando se considera o espaço na análise através do modelo hierárquico-espacial, observa-se que a razão de dependência e as demais variáveis contextuais das cidades vizinhas afetam a matrícula de determinado município.*

**Palavras-chave:** Educação. Função de Produção Educacional. Dividendo demográfico.

## Introdução

O objetivo desse artigo é investigar os determinantes da matrícula escolar no Brasil nos níveis de ensino fundamental e médio, combinando duas tradições dos estudos educacionais.

A primeira vem da literatura econômica demográfica associada com o dividendo demográfico. A principal contribuição dessa área é o clássico debate entre Coale e Hoover (1958) e Schultz (1987) sobre os efeitos de um aumento ou queda da fecundidade na educação. De acordo com os primeiros, uma menor fecundidade leva à diminuição da taxa de dependência de jovens, reduzindo o número de pessoas

matriculadas na escola. Isto provoca um crescimento da poupança do governo devido à redução nos gastos educacionais. Porém, se nem toda criança em idade escolar estiver matriculada, o dividendo demográfico pode acarretar aumento na cobertura escolar, via taxa de matrícula. Schultz, por sua vez, questiona essa relação com base na idéia de que o orçamento governamental não é elástico no curto prazo, ou seja, os gastos do governo podem não diminuir na mesma proporção da redução da coorte em idade escolar, podendo incorrer numa melhora na qualidade do ensino, via aumento dos salários dos professores ou diminuição do tamanho das turmas. Ressalta-se que a maioria dos

\*Pesquisadora da Escola de Governo da Fundação João Pinheiro.

\*\* Professor titular do curso de Demografia do Centro de Desenvolvimento e Planejamento Regional – Cedeplar, da Universidade Federal de Minas Gerais – UFMG.

estudos que testam a hipótese do dividendo demográfico utiliza dados agregados, sem nenhum controle sobre o ambiente familiar do aluno.

A segunda tradição está relacionada aos estudos dos determinantes educacionais que utilizam o arcabouço teórico da Função de Produção Educacional (FPE). A FPE analisa como os diversos insumos do processo educacional afetam os indicadores de educação, podendo ser especificada da seguinte forma:<sup>1</sup>  $A = f(F, S, O, e)$ , onde o resultado educacional (A) é função de: um vetor das características e *background* da família do indivíduo (F); um vetor de insumos escolares e dos professores (S); um vetor de outros insumos relevantes, tal como fatores comunitários (O); e o termo de erro aleatório que reflete os fatores não mensurados que contribuem para o resultado educacional (e).

Assim, esse artigo busca avaliar até que ponto as profundas mudanças demográficas ocorridas no Brasil, com a significativa queda da taxa de fecundidade total e, por conseguinte, a diminuição da razão de dependência dos jovens podem trazer algum bônus demográfico para a educação brasileira, no sentido de abrir uma “janela de oportunidades”. Porém, essa análise é feita considerando também outros fatores, que não só os demográficos, tais como o *background* familiar, a restrição de oferta e outros relacionados ao perfil escolar dos municípios.

Nesse estudo, foram utilizadas duas metodologias diferentes: a hierárquica e a hierárquica-espacial. A primeira é importante por incorporar a estrutura hierárquica dos dados, proporcionando melhor obtenção das estimativas dos parâmetros. Já a segunda inclui os efeitos espaciais na análise, advindos do fato de trabalhar com dados agregados por município, e constitui uma importante contribuição metodológica para os estudos educacionais.

## **Análise hierárquica dos determinantes da matrícula**

### *Construindo o modelo*

Para analisar o dividendo demográfico e outros fatores familiares e contextuais na matrícula escolar, utilizou-se o arcabouço teórico da FPE como ponto de partida. Porém, foram feitas algumas modificações para ajustá-lo à estrutura e disponibilidade dos dados.

O primeiro ajuste foi a utilização de modelos hierárquicos de dois níveis na estimação da FPE, em que o primeiro nível considera o indivíduo e o segundo, o município. A vantagem de estimar a FPE utilizando os modelos hierárquicos, em vez dos modelos clássicos, é que, nesses últimos, os resultados podem ser enviesados, devido à quebra de seus pressupostos, que são erros independentes e identicamente distribuídos. Somado a isso, esses modelos resolvem o problema da unidade de análise, quando se têm variáveis independentes medidas em níveis distintos de agregação.

Numa análise agregada do Brasil é bem provável que estes pressupostos não se apliquem, uma vez que a relação entre as características dos indivíduos e sua frequência escolar pode variar segundo os municípios ou quando se considera que há uma dependência entre a frequência escolar dos indivíduos que são atendidos em um mesmo município. Isso pode acontecer devido às diferentes disponibilidades de oferta educacional.

A modelagem hierárquica consiste em considerar que o intercepto e/ou a inclinação não são os mesmos para todas as unidades de nível 2 e que a variação pode se dar pelo efeito de alguma variável explicativa de nível 2 e/ou por componente aleatório. Na formalização dos modelos hierárquicos, cada nível é representado por meio de um conjunto de equações. Dessa forma, a equação no primeiro nível é descrita da seguinte maneira:

<sup>1</sup> Para maiores detalhes da Função de Produção Educacional, ver: Hanushek (2002), Hanushek, Gomes-Neto e Harbison (1996), Lee e Barro (1992), Lazear (1999), Todd e Wolpin (2003) e Krueger (1999).

$$\ln \left( \frac{\pi_{ij}}{1 - \pi_{ij}} \right) = \beta_{0j} + \beta_{1j} X_{1ij} + e_{ij} \quad (1)$$

Assumindo-se que a variabilidade do nível 2 se dá apenas no intercepto, as equações do segundo nível ficam descritas assim:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} \quad (3)$$

Em que:

$\pi_{ij}$  = a probabilidade de sucesso para o indivíduo  $i$ ;

$i = 1, 2, \dots, n_j$  unidades de nível 1, que nesse caso são os indivíduos;

$j = 1, 2, \dots, J$  unidades de nível 2, que são os municípios;

$\beta_{0j}$  é o intercepto assumido aleatório;

$\beta_{1j}$  é o coeficiente da variável independente do nível 1;

$\gamma_{00}, \gamma_{01}$  e  $\gamma_{10}$  são os parâmetros da parte fixa do modelo, comum a todos os indivíduos;

$W_j$  é o vetor de variáveis independentes medidas no nível 2;

$u_{0j}$  e  $e_{ij}$  são os efeitos aleatórios associados ao nível 2 e nível 1, respectivamente. Assume-se que  $u_{0j}$  e  $e_{ij}$  são independentes e seguem distribuição normal com média zero e variâncias constantes  $\sigma_{u_0}^2$  e  $\sigma_e^2$ .

Utilizou-se o modelo logístico hierárquico devido ao fato de as variáveis dependentes serem medidas ao nível do aluno. Dessa forma, a frequência à escola foi dimensionada através de variáveis categóricas que tentaram modelar as probabilidades de frequentar a escola tanto no ensino fundamental quanto no médio. Para tanto, foram utilizadas as seguintes categorias para a variável resposta: "1", se o indivíduo estiver frequentando a escola; e "0", em caso contrário, além de duas amostras para cortes etários – filhos de 7 a 14 e de 15

a 17 anos –, que são as faixas de idade consideradas adequadas de cursar o ensino fundamental e o médio, respectivamente.

Quanto às variáveis independentes, no primeiro nível do modelo logístico hierárquico, encontram-se aquelas relacionadas às características dos indivíduos e ao seu ambiente familiar. As que representam as características individuais, tais como sexo, cor e localização do indivíduo, têm a principal função de controle na regressão, enquanto aquelas referentes ao ambiente familiar são extremamente importantes para a probabilidade de frequência escolar, sendo selecionadas as que normalmente são utilizadas nos trabalhos sobre determinantes educacionais (SILVA; HASENBALG, 2001; RIANI, 2005; LAM; BARROS, 1995; RIOS-NETO; CESAR; RIANI, 2001; BARROS *et alii*, 2001) e que comprovadamente causam impactos no desempenho educacional do indivíduo: escolaridade da mãe; categoria de ocupação do chefe;<sup>2</sup> chefia feminina; família convivente.

Já as variáveis de nível 2 refletem os aspectos demográficos e a disponibilidade de oferta escolar do município. Para os primeiros, foi considerado o tamanho relativo da coorte em idade escolar adequada para cursar determinado nível de ensino, como forma de captar possíveis ganhos demográficos causados pela transição demográfica que vem ocorrendo no Brasil nas últimas décadas. O tamanho da população do município também foi incluído como controle para as regressões.

A disponibilidade ou restrição de oferta escolar está relacionada ao perfil escolar do município. Com esse indicador é possível averiguar até que ponto a disponibilidade de serviços educacionais em cada município interfere no acesso à escola. Para medi-la, foi utilizada a razão entre o número de professores em cada nível de ensino e a população em idade adequada de cursar determinado nível de ensino.

<sup>2</sup> Foram construídas categorias socioocupacionais para o chefe da família, considerando-se o nível educacional exigido para o exercício de cada ocupação, o tipo de especialização de suas funções e o nível de rendimento, estratificando-os em três níveis ocupacionais, quais sejam: nível superior e gerencial, nível médio e nível manual. Essa agregação foi proposta por SILVA (1973 e 1985), que busca uma homogeneização das categorias ocupacionais. Para não perder informação das famílias cujos chefes não estão ocupados, criou-se uma quarta categoria para os inativos, ou seja, os desempregados ou aposentados. Nas regressões, a categoria de referência foi a de nível manual.

Outro fator importante do perfil da rede escolar nos municípios é a concentração dos alunos em estabelecimentos de grande porte. Uma discussão recente da política educacional brasileira, por parte do Ministério da Educação, é a nucleação das escolas, procurando aglomerar aquelas pequenas em um único espaço físico, de forma a tornar mais efetiva a atenção do governo sobre estas, facilitando assim o investimento em infra-estrutura. Dessa forma, foram criadas três variáveis, segundo o tamanho dos estabelecimentos: aqueles com 50 a 200 matrículas; com 201 a 1.000 matrículas; e com mais de 1.000 matrículas.

Embora a FPE considere outros insumos educacionais, como a qualidade da infra-estrutura das escolas e dos professores, para o acesso das crianças e jovens ao sistema de ensino é provável que apenas alterações no nível da demanda e oferta escolar sejam os fatores que mais o influenciam. No nível micro, as variáveis de *background* familiar exercem influência na demanda, como, por exemplo, o maior nível educacional da mãe, que aumenta a percepção da importância da escola para seu filho.

No nível macro, o tamanho relativo da coorte em idade escolar e a restrição de oferta são os fatores mais importantes que causam alterações na demanda e oferta escolar, já que podem dificultar a obtenção de vagas em escolas, principalmente públicas. Dessa forma, optou-se por não incluir os demais fatores escolares. De fato, trabalhos anteriores comprovam que a qualidade da infra-estrutura escolar e dos professores não possui impactos na frequência escolar.<sup>3</sup>

Os dados foram retirados do Censo Demográfico de 2000, do IBGE, com exceção do número de professores em cada nível de ensino, que foi extraído do Censo Escolar de 2000, obtido junto ao Instituto Nacional de Estudos e Pesquisas Educacionais – Inep.

## Resultados

Para o acesso do indivíduo no sistema de ensino foram estimados cinco modelos, conforme pode ser visto nas Tabelas 1 e 2, respectivamente, para os ensinos fundamental e médio. No modelo 1, foram incluídas somente covariáveis do primeiro nível, possibilitando mensurar a variabilidade não-condicional do nível 2. Nos modelos 2, 3 e 4 consideraram-se, respectivamente, as seguintes variáveis: tamanho relativo da coorte em idade escolar; razão professor/população em idade escolar; e a proporção de estabelecimentos com determinado número de matrículas (variáveis de nucleação).

A razão para incluir separadamente as variáveis de nível 2 mais relevantes é a possibilidade de se verificar o quanto cada uma contribui para a redução da variabilidade não-condicional do intercepto estimada no modelo 1, que pode ser obtida através do cálculo da proporção da variância explicada, dada pela seguinte fórmula:

$$\% \text{ da variância explicada} = \frac{\hat{\sigma}_{00}(\text{n\~{a}o-condicional}) - \hat{\sigma}_{00}(\text{condicional})}{\hat{\sigma}_{00}(\text{n\~{a}o-condicional})} (4)$$

Por fim, o modelo 5 – modelo final – incorpora todas as covariáveis incluídas nos modelos anteriores, além das variáveis categóricas para o tamanho de população, como controle.

Primeiramente será comentado o resultado para o ensino fundamental (Tabela 1). Através da análise do efeito aleatório do modelo 1 (parte inferior da tabela), percebe-se que é aceitável a hipótese de o intercepto ter efeito aleatório, já que este foi significativo. Em outras palavras, é aceitável que os municípios apresentem valores distintos para a probabilidade de frequentar a escola para as crianças de 7 a 14 anos de idade. Os modelos 2 a 4 mostram que o tamanho relativo da coorte é a variável que mais explica a variabilidade do intercepto

<sup>3</sup> Ver Riani (2005).

**TABELA 1**  
**Resultado da regressão para a probabilidade de frequentar a escola, para indivíduos de 7 a 14 anos**  
**Brasil – 2000**

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4		Modelo 5	
	Coefficiente	Sig.	Coefficiente	Sig.	Coefficiente	Sig.	Coefficiente	Sig.	Coefficiente	Sig.
<b>Efeito fixo</b>										
Intercepto	3,202	0,000	6,349	0,000	2,336	0,000	2,562	0,000	4,120	0,000
Tamanho relativo da coorte de 7 a 14 anos			-18,005	0,000					-11,765	0,000
Razão professor/pop. 7 a 14 anos					13,844	0,000			10,845	0,000
Proporção de estabelecimentos com 50 a 200 matrículas							0,543	0,000	0,607	0,000
Proporção de estabelecimentos com 201 a 1.000 matrículas							1,636	0,000	1,111	0,000
Proporção de estabelecimentos com mais de 1.000 matrículas							0,723	0,000	0,924	0,000
População até 4.999 pessoas									-0,010	0,868
População entre 5.000 e 9.999 pessoas									-0,028	0,648
População entre 10.000 e 19.999 pessoas									-0,051	0,399
População entre 20.000 e 49.999 pessoas									-0,103	0,087
População entre 50.000 e 99.999 pessoas									-0,003	0,959
População entre 100.000 e 199.999 pessoas									0,019	0,793
Idade	-0,024	0,000	-0,024	0,000	-0,024	0,000	-0,024	0,000	-0,024	0,000
Situação de domicílio: 1 =urbano e 0=rural	0,613	0,000	0,612	0,000	0,614	0,000	0,611	0,000	0,611	0,000
Cor: 1 =branco e amarelo e 0=preto e pardo	0,138	0,000	0,139	0,000	0,138	0,000	0,139	0,000	0,139	0,000
Sexo: 1 =homem e 0 =mulher	-0,199	0,000	-0,199	0,000	-0,199	0,000	-0,199	0,000	-0,199	0,000
Educação da mãe	0,177	0,000	0,178	0,000	0,177	0,000	0,178	0,000	0,178	0,000
Domicílio chefiado por mulher	-0,228	0,000	-0,228	0,000	-0,228	0,000	-0,228	0,000	-0,229	0,000
Família convivente	-0,216	0,000	-0,217	0,000	-0,216	0,000	-0,216	0,000	-0,217	0,000
Categoria de ocupação do chefe com nível superior de ensino	0,297	0,000	0,308	0,000	0,297	0,000	0,302	0,000	0,310	0,000
Categoria de ocupação do chefe com nível médio de ensino	0,013	0,136	0,014	0,091	0,012	0,141	0,014	0,102	0,015	0,081
Chefe não ocupado	-0,190	0,000	-0,188	0,000	-0,190	0,000	-0,189	0,000	-0,187	0,000
<b>Efeito aleatório</b>										
Coefficiente	0,581	0,000	0,350	0,000	0,537	0,000	0,427	0,000	0,287	0,000
% da variância explicada			39,70		7,54		26,43		50,65	

Fonte: IBGE. Censo Demográfico 2000; Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

(39,70%), enquanto a razão professor/população em idade escolar é a que menos explica (7,54%).

Partindo para a análise dos efeitos fixos, as variáveis de nível 1 mostraram-se significativas para praticamente todos os casos e com o sinal esperado. Os resultados para os atributos individuais (idade, sexo, raça e situação de residência) mostraram que a idade tem relação negativa com a frequência escolar, sendo que mulheres, brancos, amarelos e moradores da área urbana possuem maior probabilidade de freqüentar a escola.

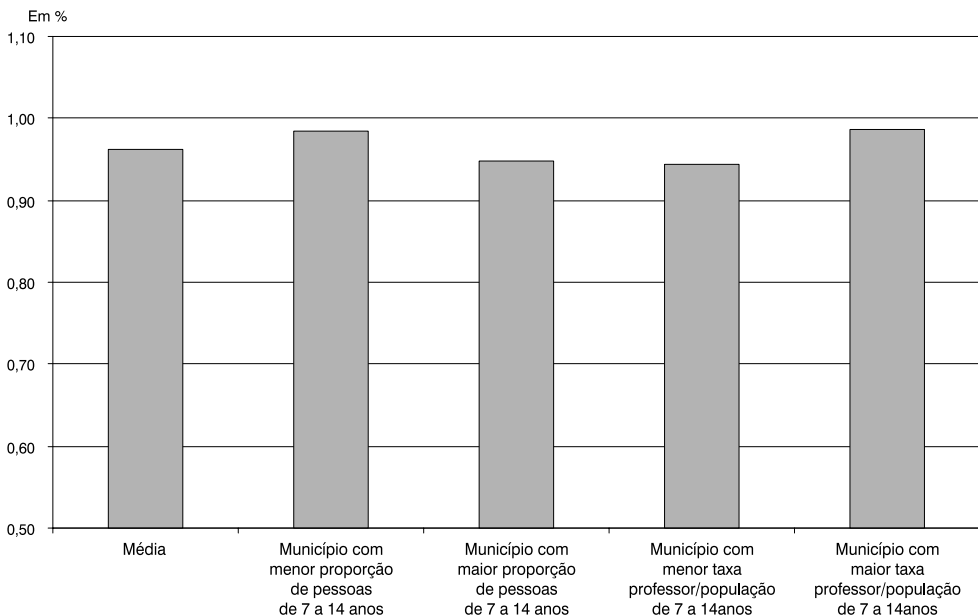
Com relação aos fatores de *background* familiar, observa-se que o nível educacional da mãe é mais importante do que as demais variáveis. Para melhor ilustrar o impacto desses fatores, a Tabela 3 apresenta as variações no valor predito segundo mudanças nos fatores de *background* familiar, utilizando os coeficientes do modelo 5. O aumento de um desvio padrão em relação

à média nos anos de estudo da mãe provoca a elevação de 1,97% na probabilidade de as crianças de 7 a 14 anos freqüentarem a escola, enquanto o efeito das demais variáveis é menor. Cabe ressaltar que o pequeno aumento dessa probabilidade deve-se ao fato de esse nível de ensino estar praticamente universalizado, ou seja, próximo de chegar aos 100%.

Analisando as variáveis de nível 2, observa-se que, para o acesso das crianças de 7 a 14 anos à escola, uma menor pressão demográfica, derivada de uma menor razão de dependência de jovens, é um dos fatores contextuais mais importantes. Menores coortes nessa faixa etária provocam aumento na probabilidade de freqüentar a escola numa magnitude próxima à verificada na simulação da variação na educação materna. Entretanto, a restrição de oferta, embora apresente coeficiente significativo, tem impacto menor. O Gráfico 1 ilustra o impacto de variações na proba-

**GRÁFICO 1**

**Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola, para pessoas de 7 a 14 anos, considerando os municípios com menor e maior valores  
Brasil – 2000**



Fonte: IBGE. Censo Demográfico 2000; Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

bilidade de freqüentar a escola para as pessoas de 7 a 14 anos, decorrente de variações no tamanho relativo da coorte e restrição de oferta. Verifica-se o maior impacto da razão de dependência de jovens, evidenciando que o acesso das crianças de 7 a 14 anos à escola é beneficiado por um maior dividendo demográfico, proveniente da redução da fecundidade.

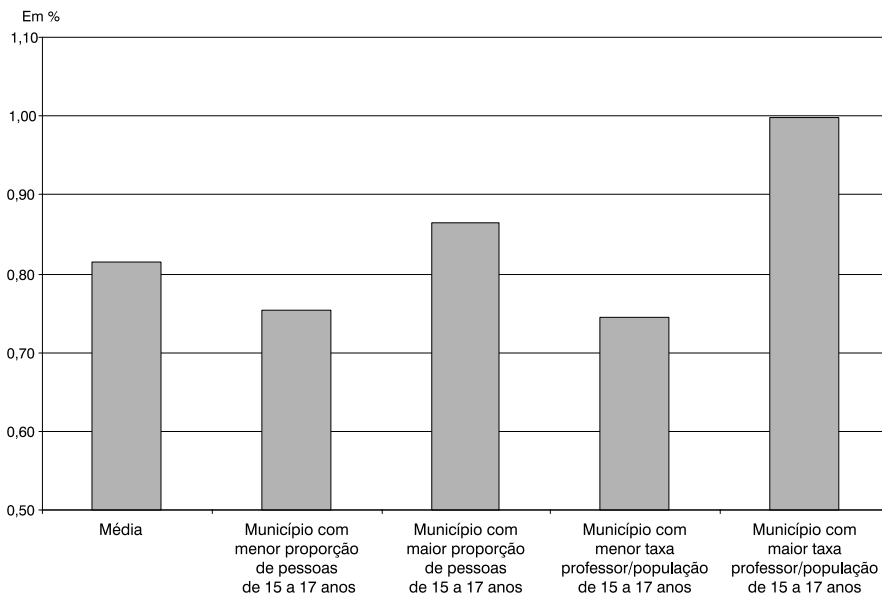
Os resultados para os indivíduos de 15 a 17 anos, referentes ao ensino médio, encontram-se na Tabela 2. Observa-se que, para esse nível de ensino, também é aceitável a hipótese de o intercepto ser considerado com efeito aleatório. Quando se acrescentam variáveis de nível 2, ao contrário do caso anterior, a restrição de oferta é o fator que mais explica a variabilidade do intercepto. Como as variáveis de nucleação não são significativas para

essa probabilidade, foi estimado mais um modelo de regressão que não as considera (modelo 6).

Com relação às variáveis de nível 1, os resultados são muito semelhantes aos encontrados para as pessoas de 7 a 14 anos, com as seguintes principais diferenças: impacto negativo mais forte da idade, sexo e famílias chefiadas por mulher; maior impacto das variáveis de posição da ocupação do chefe da família, sendo o efeito positivo superior para aquelas com maior *status* socioeconômico e negativo para chefe fora do mercado de trabalho; e impacto positivo para família convivente, inverso do ocorrido no caso anterior.

Para o ensino médio, a variação decorrente das simulações<sup>4</sup> de mudanças nas variáveis de nível 1 é bem maior do que para o fundamental, principalmente a

**GRÁFICO 2**  
Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola, para as pessoas de 15 a 17 anos, considerando os municípios com menor e maior valores  
Brasil – 2000



Fonte: IBGE. Censo Demográfico 2000; Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

<sup>4</sup> Para o ensino médio as simulações foram feitas considerando os coeficientes do modelo 6.

**TABELA 2**  
**Resultado da regressão para a probabilidade de frequentar a escola, para pessoas de 15 a 17 anos**  
**Brasil – 2000**

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3	
	Coefficiente	Sig.	Coefficiente	Sig.	Coefficiente	Sig.
<b>Efeito fixo</b>						
Intercepto	1,440	0,000	1,912	0,000	1,141	0,000
Tamanho relativo da coorte de 15 a 17 anos			-7,081	0,000		
Razão professor/pop. 15 a 17 anos					8,115	0,000
Proporção de estabelecimentos com 50 a 200 matrículas						
Proporção de estabelecimentos com 201 a 1.000 matrículas						
Proporção de estabelecimentos com mais de 1.000 matrículas						
População até 4.999 pessoas						
População entre 5.000 e 9.999 pessoas						
População entre 10.000 e 19.999 pessoas						
População entre 20.000 e 49.999 pessoas						
População entre 50.000 e 99.999 pessoas						
População entre 100.000 e 199.999 pessoas						
Idade	-0,441	0,000	-0,441	0,000	-0,441	0,000
Situação de domicílio: 1 = urbano e 0 = rural	0,470	0,000	0,470	0,000	0,469	0,000
Cor: 1 = branco e amarelo e 0 = preto e pardo	0,162	0,000	0,162	0,000	0,162	0,000
Sexo: 1 = homem e 0 = mulher	-0,445	0,000	-0,445	0,000	-0,445	0,000
Educação da mãe	0,158	0,000	0,158	0,000	0,158	0,000
Domicílio chefiado por mulher	-0,362	0,000	-0,362	0,000	-0,362	0,000
Família convivente	0,069	0,003	0,069	0,003	0,069	0,003
Categoria de ocupação do chefe com nível superior de ensino	0,479	0,000	0,481	0,000	0,483	0,000
Categoria de ocupação do chefe com nível médio de ensino	0,141	0,000	0,141	0,000	0,141	0,000
Chefe não ocupado	-0,048	0,000	-0,048	0,000	-0,047	0,000
<b>Efeito aleatório</b>						
Coefficiente	0,294	0,000	0,290	0,000	0,262	0,000
% da variância explicada			1,43		10,88	

(continua)



(continuação)

Variáveis Independentes	Modelo 4		Modelo 5		Modelo 6	
	Coefficiente	Sig.	Coefficiente	Sig.	Coefficiente	Sig.
<b>Efeito fixo</b>						
Intercepto	1,518	0,000	1,000	0,000	1,113	0,000
Tamanho relativo da coorte de 15 a 17 anos			9,421	0,000	9,241	0,000
Razão professor/pop. 15 a 17 anos			10,310	0,000	10,285	0,000
Proporção de estabelecimentos com 50 a 200 matrículas	-0,136	0,016	-0,015	0,783		
Proporção de estabelecimentos com 201 a 1.000 matrículas	-0,098	0,081	0,416	0,119		
Proporção de estabelecimentos com mais de 1.000 matrículas	0,637	0,000				
População até 4.999 pessoas			-0,693	0,000	-0,788	0,000
População entre 5.000 e 9.999 pessoas			-0,725	0,000	-0,781	0,000
População entre 10.000 e 19.999 pessoas			-0,665	0,000	-0,711	0,000
População entre 20.000 e 49.999 pessoas			-0,609	0,000	-0,635	0,000
População entre 50.000 e 99.999 pessoas			-0,406	0,000	-0,413	0,000
População entre 100.000 e 199.999 pessoas			-0,191	0,003	-0,197	0,003
Idade	-0,441	0,000	-0,441	0,000	-0,441	0,000
Situação de domicílio: 1=urbano e 0=rural	0,470	0,000	0,468	0,000	0,468	0,000
Cor: 1=branco e amarelo e 0=preto e pardo	0,162	0,000	0,162	0,000	0,162	0,000
Sexo: 1=homem e 0=mulher	-0,445	0,000	-0,444	0,000	-0,444	0,000
Educação da mãe	0,158	0,000	0,158	0,000	0,158	0,000
Domicílio chefiado por mulher	-0,362	0,000	-0,362	0,000	-0,362	0,000
Família convivente	0,069	0,003	0,069	0,003	0,069	0,003
Categoria de ocupação do chefe com nível superior de ensino	0,479	0,000	0,485	0,000	0,485	0,000
Categoria de ocupação do chefe com nível médio de ensino	0,141	0,000	0,141	0,000	0,141	0,000
Chefe não ocupado	-0,048	0,000	-0,047	0,000	-0,047	0,000
<b>Efeito aleatório</b>						
Coefficiente	0,282	0,000	0,220	0,000	0,223	0,000
% da variância explicada	3,93		25,03		24,08	

Fonte: IBGE. Censo Demográfico 2000, Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

educação materna (Tabela 3). Tal fato é atribuído à menor proporção de pessoas da amostra de 15 a 17 anos que freqüentam a escola (81%, contra 95% da amostra de filhos de 7 a 14 anos).

Partindo para a análise das variáveis municipais, percebe-se que, para o ensino médio, a restrição de oferta tem impacto positivo importante. Quando se considera o município com maior número de professores deste ensino por população de 15 a 17 anos, a probabilidade de freqüentar a escola chega a quase 100%, conforme pode ser visto no Gráfico 3. Destaca-se também o sinal positivo encontrado para o tamanho relativo da coorte em idade escolar, resultado não esperado. Porém, quando se considera apenas essa variável de nível 2 na análise (modelo 2), seu sinal torna-se negativo. Através dos exercícios não mostrados nesse artigo, verifica-se que o sinal dessa variável torna-se positivo quando a restrição de oferta é incluída na regressão. Uma possível explicação é o baixo nível de atendimento dessa faixa etária, bem distante do de saturação, o que faz com que um aumento da demanda, quando controlado pela restrição de oferta, tenha efeito positivo.

Em síntese, os resultados mostraram que uma menor pressão demográfica, derivada de uma menor razão de dependência de jovens, é um dos fatores contextuais mais importantes para a freqüência escolar das crianças de 7 a 14 anos. Já para a fre-

qüência escolar dos jovens de 15 a 17 anos, a disponibilidade de oferta educacional, representada pela razão professor/população em idade escolar, é o fator municipal mais relevante, sinalizando a necessidade de políticas públicas que aumentem a oferta desse nível de ensino nos municípios brasileiros.

Além disso, os resultados comprovam a grande importância da educação materna no atendimento para as duas faixas etárias, evidenciando a desigualdade de oportunidades educacionais no Brasil.

### Análise hierárquica-espacial dos determinantes da matrícula

#### Construindo o modelo

Os modelos hierárquicos levam em conta a heterogeneidade espacial, considerando a existência de um efeito aleatório ligado a cada unidade de nível 2, nesse caso os municípios. Porém, eles não englobam a autocorrelação espacial, já que um dos seus pressupostos é a independência entre os efeitos das áreas do nível 2, ou seja, os erros desse nível são independentes e com variância constante. Isto pode não ser realista quando se considera que o Brasil possui uma forte heterogeneidade social e econômica, que se traduz numa grande heterogeneidade espacial, podendo-se esperar que as áreas próximas sejam mais similares do que as

**TABELA 3**  
Variações na probabilidade de freqüentar a escola, por faixa etária, segundo variações nas variáveis de *background* familiar para diferentes amostras (1)  
Brasil – 2000

Variáveis Independentes	7 a 14 anos		15 a 17 anos		15 a 17 anos com ensino fundamental concluído	
	Média mais desvio padrão	Média menos desvio padrão	Média mais desvio padrão	Média menos desvio padrão	Média mais desvio padrão	Média menos desvio padrão
Educação da mãe	1,97	-3,83	9,75	-14,59	3,70	-5,65
Domicílio chefiado por mulher	-0,34	0,31	-2,88	2,62	-1,35	1,20
Família convivente	-0,15	0,14	0,15	-0,15	0,08	-0,09
Categoria de ocupação do chefe nível superior	0,28	-0,30	2,24	-2,43	1,29	-1,46
Categoria de ocupação do chefe nível médio	0,02	-0,02	0,95	-0,98	0,59	-0,63
Chefe não ocupado	-0,33	0,30	-0,40	0,40	-0,04	0,04

Fonte: IBGE. Censo Demográfico 2000. Elaboração própria.

(1) Utilizaram-se os coeficientes do modelo 5 para a amostra de 7 a 14 anos e o modelo 6 para a amostra de 15 a 17 anos.

**TABELA 4**  
**Estatística I de Moran, segundo variáveis**  
**Brasil – 2000**

Variáveis	I de Moran	Prob.
Taxa de atendimento de 7 a 14 anos	0,579	0,000
Taxa de atendimento de 15 a 17 anos	0,392	0,000
<b>Restrição de oferta</b>		
Razão professor do ensino fundamental por população em idade escolar	0,428	0,000
Razão professor do ensino médio por população em idade escolar	0,439	0,000
<b>Variáveis de nucleação – ensino fundamental</b>		
Proporção de estabelecimentos com 50 a 200 matrículas	0,206	0,000
Proporção de estabelecimentos com 201 a 1.000 matrículas	0,530	0,000
Proporção de estabelecimentos com mais de 1.000 matrículas	0,215	0,000
<b>Variáveis de nucleação – ensino médio</b>		
Proporção de estabelecimentos com 50 a 200 matrículas	0,077	0,000
Proporção de estabelecimentos com 201 a 1.000 matrículas	0,085	0,000
Proporção de estabelecimentos com mais de 1.000 matrículas	0,100	0,000
<b>Dividendo demográfico</b>		
Tamanho relativo da coorte de 7 a 14 anos	0,811	0,000
Tamanho relativo da coorte de 15 a 17 anos	0,652	0,000
<b>Tamanho da população</b>		
De 0 a 4.999 habitantes	0,161	0,000
De 5.000 a 9.999 habitantes	0,030	0,000
De 10.000 a 19.999 habitantes	0,037	0,000
De 20.000 a 49.999 habitantes	0,053	0,000
De 50.000 a 99.999 habitantes	0,014	0,000
De 100.000 a 199.999 habitantes	0,052	0,000
Acima de 200.000 habitantes	0,193	0,000

Fonte: IBGE. Censo Demográfico 2000; Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

distantes, ocorrendo, portanto, correlação espacial no nível macro.

Os efeitos espaciais advindos do fato de se trabalhar com dados agregados são a dependência e a heterogeneidade espacial. A dependência espacial, chamada de autocorrelação espacial na sua forma empírica, é a correlação do valor de uma variável na localização  $i$  com o seu valor no local  $j$ . Já a heterogeneidade espacial, num modelo econométrico, corresponde à instabilidade estrutural nas médias e variância das variáveis no espaço.

Em outras palavras, as variáveis agregadas de um município podem estar sendo influenciadas por aquelas dos municípios vizinhos, levando a uma autocorrelação espacial positiva. Por outro lado, podem existir unidades espaciais com comportamento

diferenciado de seus vizinhos, sugerindo autocorrelação espacial negativa. Quando não se corrigem esses efeitos espaciais, os coeficientes do modelo hierárquico podem ser inconsistentes e ineficientes.

Para detectar a autocorrelação espacial, foi feito um diagnóstico das variáveis independentes do nível 2 dos modelos hierárquicos vistos na seção anterior, bem como da taxa de atendimento, que representa a variável dependente do modelo hierárquico. Porém, como a análise espacial requer que as medidas sejam agregadas em unidades geográficas, estas são medidas em termos de taxas municipais, em vez de dicotômicas, como no modelo anterior.

O diagnóstico foi feito através da estatística I de Moran,<sup>5</sup> que é uma ferramenta de análise exploratória espacial dos dados.

<sup>5</sup> A estatística I de Moran é dada pela fórmula: 
$$I = \frac{N}{S_0} \frac{\sum_i \sum_j w_{ij} (x_i - \mu)(x_j - \mu)}{\sum (x_i - \mu)^2}$$
 onde: N é o número de observações;  $w_{ij}$  é o elemento da Matriz W correspondente ao par  $i$  e  $j$ ;  $x_i$  e  $x_j$  são os valores das variáveis nos locais  $i$  e  $j$ ;  $S_0$  é a soma de todos os pesos ( $S_0 = \sum_i \sum_j w_{ij}$ ) e para matrizes padronizadas  $S_0 = N$ ;  $\mu$  é a média da variável  $x$ .

Os resultados apresentados na Tabela 4 mostram que todas as variáveis possuem autocorrelação espacial positiva, pois o I de Moran foi positivo e significativo a 1% para todos os casos. Dessa forma, foi feita uma outra especificação da FPE, procurando incluir os efeitos espaciais na análise.

A econometria espacial incorpora os dois efeitos espaciais descritos anteriormente, com modelos de regressões. Porém, esses modelos trabalham apenas com variáveis agregadas, não possibilitando incluir na análise variáveis individuais e relacionadas ao *background* familiar.

Dessa forma, buscou-se desenvolver uma metodologia que combine as metodologias hierárquica e espacial. Tal modelo, denominado hierárquico-espacial, é uma contribuição metodológica importante desse trabalho. Assim, as características do modelo espacial – autocorrelação e heterogeneidade espacial – são incorporadas no modelo hierárquico, que possui uma natureza não espacial, mas que permite considerar diferentes níveis de agregação em uma mesma análise.

A vantagem de combinar as duas metodologias é que a modelagem hierárquica possibilita analisar conjuntamente as variáveis relacionadas aos indivíduos e as contextuais (municipais) em níveis diferentes de agregação, diminuindo o problema relacionado à falácia ecológica. Por sua vez, o modelo espacial permite trabalhar com a possível correlação espacial das variáveis dependentes e independentes e/ou dos erros no segundo nível do modelo hierárquico, o que possibilitará obter melhores estimativas dos parâmetros das regressões.

Este campo de estudo está começando a ser desenvolvido por meio de métodos *baeysianos*, que utilizam o programa computacional WinBUGS juntamente com MLwiN. Porém, a proposta aqui é desen-

volver uma técnica usando os programas HLM e SpaceStat. Tentativa semelhante foi desenvolvida no trabalho de Morenoff (2003) para o estudo do peso da criança ao nascer. Assim, esse trabalho foi o ponto de partida para a metodologia da presente análise.

Morenoff (2003) parte do pressuposto de que o processo espacial ocorre através do processo auto-regressivo de defasagem espacial,<sup>6</sup> cuja fórmula é dada por:

$$y = \rho W y + X \beta + \varepsilon \quad (5)$$

onde:  $\rho$  é o parâmetro auto-regressivo espacial;  $W$  é a matriz de pesos espaciais;  $X$  é a matriz de variáveis independentes;  $\beta$  é a matriz de coeficientes da regressão; e  $\varepsilon$  é o vetor do termo de erro aleatório.

Esta equação é a forma estrutural do modelo de defasagem espacial. Com a forma reduzida, é mais fácil perceber como age o processo espacial, também conhecido, neste caso, como multiplicador espacial, qual seja:

$$y = (I - \rho W)^{-1} X \beta + (I - \rho W)^{-1} \varepsilon \quad (6)$$

O termo  $(I - \rho W)^{-1}$  é a inversa de Leontief, que liga a variável  $y_i$  para todos  $x_i$  do sistema através do multiplicador espacial, e  $y_i$  ao e de todas as localizações no sistema, não apenas ao erro em  $i$  (ANSELIN, 2003).

Desta forma, a equação (6) é importante por mostrar que o efeito espacial opera através das covariáveis  $X$  e do termo de erro não observado. Em outras palavras, existe uma dependência espacial nas variáveis observadas e não-observadas. A alternativa encontrada por Morenoff, para introduzir o efeito espacial na análise hierárquica, foi a inclusão de defasagens espaciais das variáveis independentes<sup>7</sup> no modelo hierárquico, ou seja, a inclusão de  $WX$ . Esse procedimento permite corrigir a dependência espacial apenas das variáveis observadas.

<sup>6</sup> Existe outro modelo espacial – o de erro espacial – que pressupõe a existência de correlação espacial em variáveis que não foram incluídas no modelo. Sua fórmula é dada por:

$$y = X \beta + \lambda W \varepsilon + u$$

onde:  $\lambda$  é o coeficiente auto-regressivo;  $\varepsilon$  é o termo de erro que possui dependência espacial e  $u$  é o termo de erro aleatório.

<sup>7</sup> A defasagem espacial de determinada variável na localidade  $j$  é a sua média nos seus vizinhos, que são determinados pela matriz de peso espacial.

A estratégia adotada nesse artigo consiste em três passos. Em primeiro lugar são obtidos os resíduos do segundo nível do modelo hierárquico não-condicional, ou seja, quando não se incluem variáveis de nível 2 na análise, mas apenas as de nível 1.

Em seguida, são feitos testes econométricos espaciais nesses resíduos,<sup>8</sup> o que possibilita perceber se o efeito espacial opera na forma de defasagem ou de erro espacial. Os testes mostraram que o processo espacial ocorre mais intensamente na forma de defasagem espacial – resultado não mostrado nesse artigo.

Na terceira etapa, é estimado o modelo hierárquico-espacial, com duas especificações diferentes. A primeira consiste em acrescentar a defasagem dos resíduos descritos anteriormente no segundo nível do modelo hierárquico, ou seja, na equação (2), para corrigir a dependência espacial tanto das variáveis observadas quanto das não-observadas. A segunda compreende a inclusão, no nível 2 da regressão hierárquica, da defasagem espacial das variáveis independentes contextuais. Essa é a mesma estratégia adotada por Morenoff. Cabe ressaltar que, com essa estratégia, não se está controlando os efeitos espaciais das variáveis não-observadas, mas apenas das observadas. Porém, ela é útil por permitir verificar qual covariável é mais espacialmente correlacionada com as variáveis independentes.

### Resultados

Os resultados do modelo hierárquico-espacial para a frequência escolar dos indivíduos de 7 a 14 e de 15 a 17 anos encontram-se nas Tabelas 5 e 6. Para facilitar as comparações, os resultados do modelo hierárquico da seção anterior também são apresentados nas tabelas. Visto que nessa parte as variáveis de interesse são as contextuais, os resultados para as variáveis de nível 1 não são mostrados; entretanto não diferem dos anteriores.

Analisando primeiramente os indivíduos de 7 a 14 anos (Tabela 5), percebe-se

que a inclusão da defasagem espacial dos resíduos diminui bastante o impacto da pressão demográfica, nos dois tipos de resíduo – *bayesiano* e MQO. Para exemplificar, no modelo hierárquico-espacial que inclui a defasagem espacial do resíduo *bayesiano*, a diminuição de um desvio padrão em relação à média, do tamanho relativo da coorte, provoca variação de 0,31% na probabilidade de as crianças de 7 a 14 anos freqüentarem a escola, passando de 0,961 para 0,964. Quando não se incluiu o termo espacial na análise, a variação foi de 1,13%. Com relação à variável de restrição de oferta, a redução do seu efeito é bem menor, de forma que ela passa a ser mais importante do que o tamanho relativo da coorte.

Quando se acrescentam as defasagens espaciais das variáveis independentes, observa-se que o tamanho relativo da coorte em idade escolar dos municípios vizinhos é o fator defasado que mais exerce influência no acesso das crianças à escola de determinado município. Uma variação negativa de um desvio padrão em relação à média no tamanho relativo da coorte dos municípios vizinhos provoca aumento de 0,46% na probabilidade de freqüentar a escola, alcançando 0,966. Vale lembrar que essa pequena variação na probabilidade deve-se ao fato de a freqüência das crianças de 7 a 14 anos já estar praticamente no seu nível máximo, ou seja, 100%.

No que diz respeito aos jovens de 15 a 17 anos (Tabela 6), observa-se que a inclusão dos efeitos espaciais também diminuiu o impacto das covariáveis. Um fato importante é que a inclusão da defasagem do resíduo muda o sinal da variável de pressão demográfica, tornando-o negativo, o que é esperado. Já no modelo que inclui apenas as defasagens das variáveis independentes, o sinal continua positivo, apesar de o coeficiente ser menor do que no modelo não-espacial (apenas hierárquico). Tal resultado dá indícios de que o tamanho relativo da coorte está captando outros efeitos contextuais, que são autocorrelacionados espacialmente, já que a defasagem espacial do resíduo capta os

<sup>8</sup> O HLM gera dois tipos de resíduo: *bayesiano* e MQO. A análise foi feita para ambos.

efeitos espaciais das variáveis observadas e não-observadas.

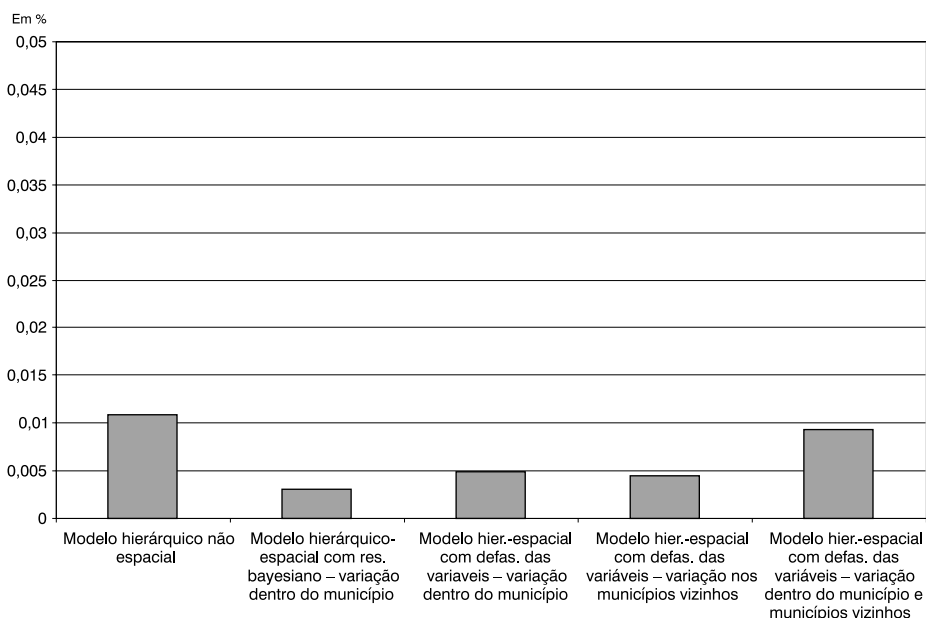
Entre as variáveis independentes defasadas, verifica-se que as de pressão demográfica e restrição de oferta são as mais importantes, sendo o efeito da primeira bastante significativo. Um aumento do tamanho relativo da coorte de 15 a 17 anos nos municípios vizinhos, de um desvio padrão em relação à média, aumenta em 2,32% a probabilidade de frequência escolar dessa faixa etária no município local.

Assim, ao se tentar controlar a autocorrelação espacial nos modelo hierárquicos, diminuindo os problemas de viés e eficiência das estimativas, observam-se, por um lado, redução do efeito das variáveis de nível 2 e, por outro, um significativo efeito das defasagens espaciais das variáveis contextuais (modelo hierárquico-espacial com a defasagem das variáveis independentes).

Esses resultados sugerem que as variáveis de nível 2 possuem um efeito direto e

indireto sobre as probabilidades de frequência escolar. O efeito direto seria o impacto na probabilidade de cursar a escola decorrente de mudanças nas variáveis contextuais no próprio município. Já o indireto seria o impacto decorrente de mudanças das variáveis contextuais nos municípios vizinhos. O primeiro efeito é captado pelos coeficientes das covariáveis do nível 2 nas duas especificações do modelo hierárquico-espacial (com defasagem dos resíduos e com defasagem das covariáveis) e, o segundo, pelos coeficientes das defasagens espaciais das variáveis do modelo hierárquico-espacial que as inclui. Pode-se dizer que o resultado final de uma variação do tamanho relativo da coorte e da razão professor/população em idade escolar – principais fatores contextuais para esse caso – é a soma desses dois efeitos, já que essas variáveis são altamente correlacionadas espacialmente, de tal forma que podem variar conjuntamente.

**GRÁFICO 3**  
**Variação na probabilidade de cursar a escola decorrente da diminuição de um desvio-padrão em relação à média do tamanho relativo da coorte em idade escolar – simulações segundo os diferentes modelos para as pessoas de 7 a 14 anos**  
**Brasil – 2000**



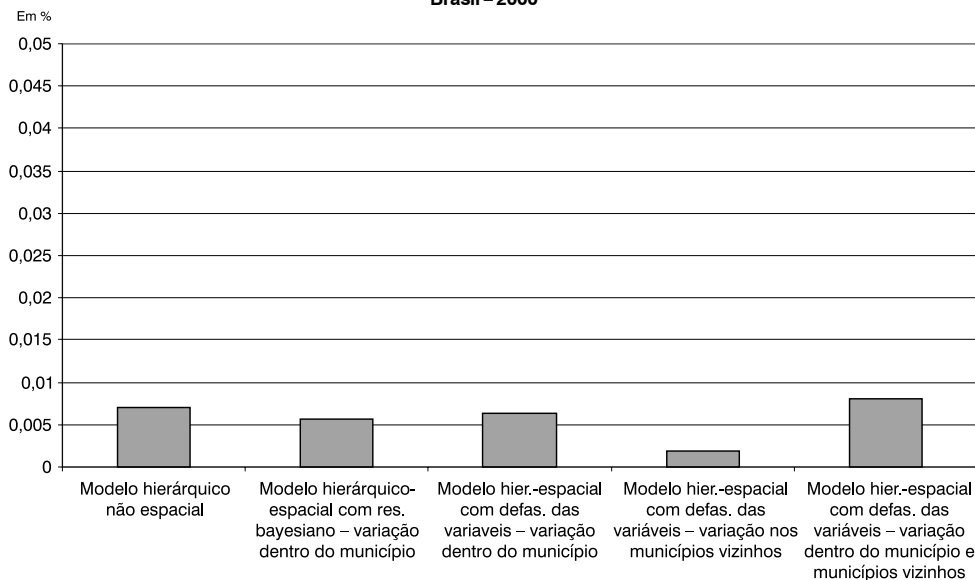
Fonte: IBGE. Censo Demográfico 2000. Elaboração própria.

Os Gráficos 3 a 6 ajudam a compreender melhor os efeitos diretos e indiretos. O Gráfico 3 mostra a mudança, em pontos percentuais, na probabilidade de cursar a escola para as pessoas de 7 a 14 anos decorrente de diminuição de um desvio padrão em relação à média do tamanho relativo da coorte em idade escolar no próprio município e em seus vizinhos, para os diferentes modelos. A três primeiras colunas medem o aumento percentual quando ocorre redução na pressão demográfica no próprio município, sendo que na primeira as simulações foram feitas com base nos coeficientes do modelo hierárquico da seção anterior,<sup>9</sup> que não separa os efeitos diretos e os indiretos. Na segunda e terceira colunas, as simulações foram calculadas por meio dos modelos hierárquico-espaciais com a defasagem do resíduo *bayesiano* e com as defasagens das covariáveis de nível 2, respectivamente. Nesse caso, pode-se dizer que esse seria o efeito

direto do aumento provocado pela diminuição do tamanho relativo da coorte. No entanto, como o primeiro modelo está captando a autocorrelação espacial das variáveis observadas e não-observadas e o segundo modelo apenas das observadas, os resultados são diferentes. Quanto maior a correlação das variáveis incluídas no modelo com as omitidas e quanto maior a autocorrelação espacial dessas últimas, maior será a diferença dos efeitos diretos encontrados nos dois modelos hierárquico-espaciais. A quarta coluna mede o aumento percentual quando ocorre diminuição no valor dessa variável nos municípios vizinhos, que foi considerado o efeito indireto do dividendo demográfico. Por fim, a última coluna é a soma da terceira e quarta colunas, ou seja, é a soma dos efeitos direto e indireto do tamanho relativo da coorte em idade escolar.

No Gráfico 4 esse mesmo exercício foi realizado para a razão professor/popula-

**GRÁFICO 4**  
**Variação na probabilidade de cursar a escola decorrente do aumento de um desvio padrão em relação à média da razão professor/população em idade escolar – simulações segundo os diferentes modelos para as pessoas de 7 a 14 anos**  
**Brasil – 2000**



Fonte: Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

<sup>9</sup> Para a frequência escolar dos indivíduos de 7 a 14 anos, utilizou-se o modelo 5 e, para os jovens de 15 a 17 anos, o modelo 6.

**TABELA 5**  
**Resultado dos diferentes modelos para a probabilidade de frequentar a escola, para as pessoas de 7 a 14 anos**  
 Brasil – 2000

Variáveis independentes	Modelo hierárquico (modelo 5)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
	Coefficiente	Sig.	Coefficiente	Sig.	Coefficiente	Sig.	Coefficiente	Sig.
<b>Efeito fixo</b>								
Intercepto	4,120	0,000	3,134	0,000	3,047	0,000	3,613	0,000
Tamanho relativo da coorte de 7 a 14 anos	-11,765	0,000	-2,802	0,000	-2,438	0,000	-4,744	0,000
Razão professor/pop. 7 a 14 anos	10,845	0,000	8,144	0,000	8,289	0,000	9,295	0,000
Proporção de estabelecimentos com 50 a 200 matrículas	0,607	0,000	0,350	0,000	0,352	0,000	0,460	0,000
Proporção de estabelecimentos com 201 a 1.000 matrículas	1,111	0,000	0,629	0,000	0,619	0,000	0,846	0,000
Proporção de estabelecimentos com mais de 1.000 matrículas	0,924	0,000	0,210	0,175	0,317	0,040	0,718	0,000
População até 4.999	-0,010	0,868	-0,286	0,000	-0,329	0,000	-0,283	0,000
População entre 5.000 e 9.999 pessoas	-0,028	0,648	-0,236	0,000	-0,268	0,000	-0,258	0,000
População entre 10.000 e 19.999 pessoas	-0,051	0,399	-0,217	0,000	-0,249	0,000	-0,236	0,000
População entre 20.000 e 49.999 pessoas	-0,103	0,087	-0,218	0,000	-0,250	0,000	-0,252	0,000
População entre 50.000 e 99.999 pessoas	-0,003	0,959	-0,055	0,302	-0,083	0,116	-0,093	0,148
População entre 100.000 e 199.999 pessoas	0,019	0,793	-0,058	0,345	-0,063	0,301	-0,038	0,601
Defasagem espacial do resíduo			0,781	0,000	0,690	0,000		
Defasagem espacial tamanho relativo da coorte de 7 a 14 anos							-6,418	0,000
Defasagem espacial razão professor/pop. 7 a 14 anos							5,040	0,000
Defasagem espacial prop. de estabel. com 50 a 200 matrículas							0,696	0,000
Defasagem espacial prop. de estabel. com 201 a 1.000 matrículas							0,477	0,000
Defasagem espacial prop. de estabel. com mais de 1.000 matrículas							-0,129	0,733
Defasagem espacial população até 4.999 pessoas							0,519	0,000
Defasagem espacial população entre 5.000 e 9.999 pessoas							0,124	0,269
Defasagem espacial população entre 10.000 e 19.999 pessoas							0,106	0,337
Defasagem espacial população entre 20.000 e 49.999 pessoas							0,068	0,537
Defasagem espacial população entre 50.000 e 99.999 pessoas							0,031	0,796
Defasagem espacial população entre 100.000 e 199.999 pessoas							0,241	0,089
<b>Efeito aleatório</b>								
Coefficiente	0,287	0,000	0,192	0,000	0,190	0,000	0,268	0,000

Fonte: IBGE. Censo Demográfico 2000; Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

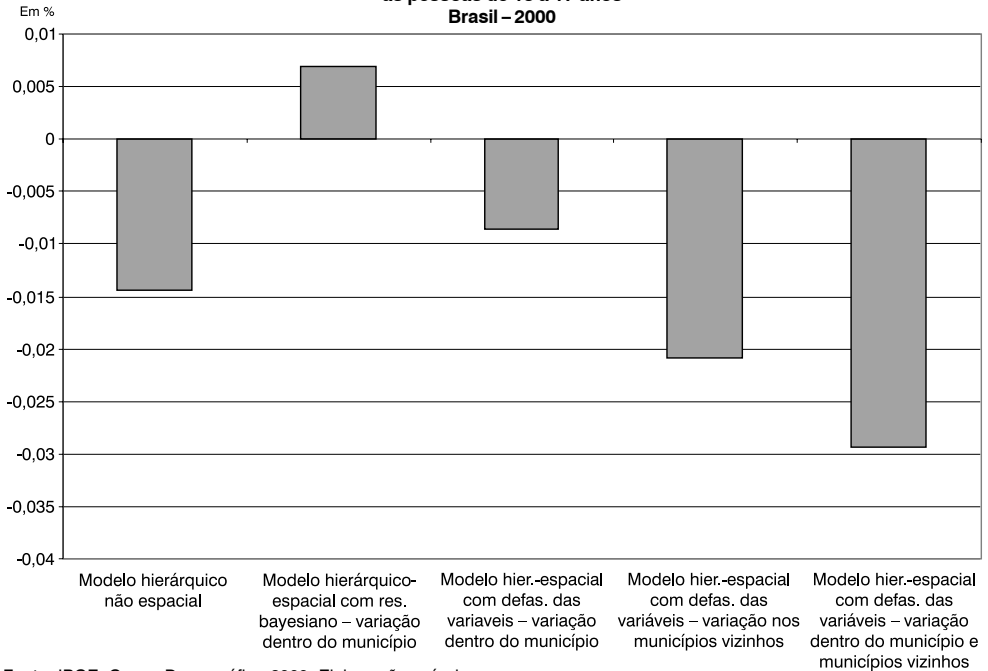


**TABELA 6**  
**Resultado dos diferentes modelos para a probabilidade de frequentar a escola, para as pessoas de 15 a 17 anos**  
**Brasil – 2000**

Variáveis Independentes	Modelo Hierárquico (Modelo 6)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
	Coefficiente	Sig	Coefficiente	Sig	Coefficiente	Sig	Coefficiente	Sig
<b>Efeito fixo</b>								
Intercepto	1,113	0,000	1,909	0,000	1,910	0,000	0,637	0,000
Tamanho relativo da coorte de 15 a 17 anos	9,241	0,000	-4,501	0,000	-4,157	0,000	5,762	0,000
Razão professor/pop. 15 a 17 anos	10,285	0,000	9,602	0,000	9,277	0,000	9,284	0,000
População até 4.999	-0,788	0,000	-0,624	0,000	-0,658	0,000	-0,714	0,000
População entre 5.000 e 9.999 pessoas	-0,781	0,000	-0,589	0,000	-0,622	0,000	-0,696	0,000
População entre 10.000 e 19.999 pessoas	-0,711	0,000	-0,557	0,000	-0,584	0,000	-0,625	0,000
População entre 20.000 e 49.999 pessoas	-0,635	0,000	-0,503	0,000	-0,525	0,000	-0,548	0,000
População entre 50.000 e 99.999 pessoas	-0,413	0,000	-0,284	0,000	-0,301	0,000	-0,328	0,000
População entre 100.000 e 199.999 pessoas	-0,197	0,003	-0,129	0,009	-0,138	0,006	-0,163	0,010
Defasagem espacial do resíduo			1,004	0,000	0,789	0,000		
Defasagem esp. tamanho relativo da coorte de 15 a 17 anos							13,695	0,000
Defasagem espacial razão professor/pop. 15 a 17							4,308	0,000
Defasagem espacial população até 4.999 pessoas							-0,166	0,062
Defasagem espacial população entre 5.000 e 9.999 pessoas							-0,578	0,000
Defasagem esp. população entre 10.000 e 19.999 pessoas							-0,513	0,000
Defasagem esp. população entre 20.000 e 49.999 pessoas							-0,404	0,000
Defasagem esp. população entre 50.000 e 99.999 pessoas							-0,370	0,001
Defasagem esp. população entre 100.000 e 199.999 pessoas							-0,049	0,695
<b>Efeito aleatório</b>								
Coefficiente	0,223	0,000	0,125	0,000	0,126	0,000	0,213	0,000

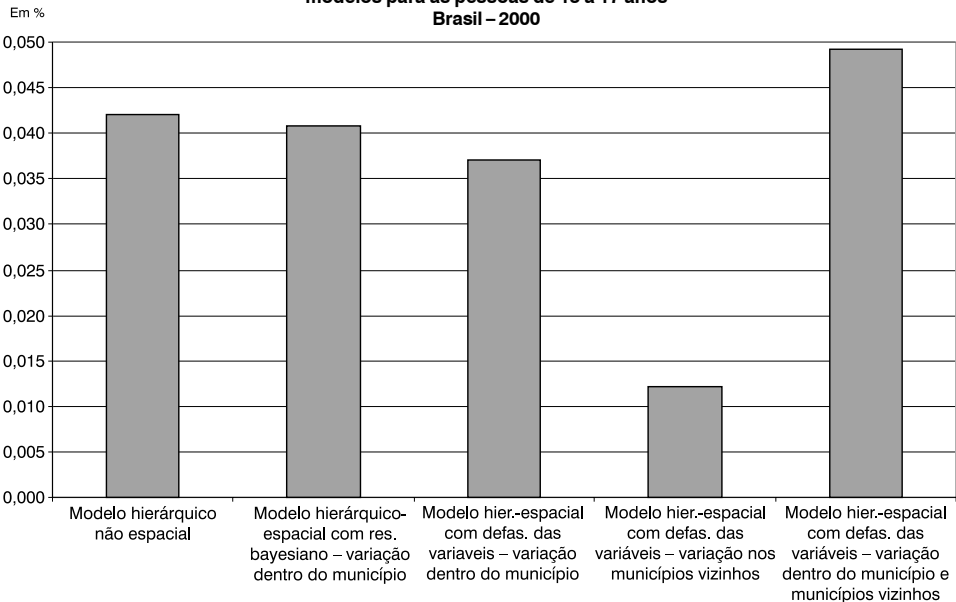
Fonte: IBGE. Censo Demográfico 2000; Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

**GRÁFICO 5**  
**Varição na probabilidade de cursar a escola decorrente da diminuição de um desvio padrão em relação à média do tamanho relativo da coorte em idade escolar – simulações segundo os diferentes modelos para as pessoas de 15 a 17 anos**  
**Brasil – 2000**



Fonte: IBGE. Censo Demográfico 2000. Elaboração própria.

**GRÁFICO 6**  
**Varição na probabilidade de cursar a escola decorrente do aumento de um desvio padrão em relação à média da razão professor/população em idade escolar – simulações segundo diferentes modelos para as pessoas de 15 a 17 anos**  
**Brasil – 2000**



Fonte: Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. Censo Escolar 2000. Elaboração própria.

lação em idade escolar, porém, nesse caso, considerou-se como variação na restrição de oferta o aumento de um desvio-padrão em relação à média. Os Gráficos 5 e 6 são semelhantes aos dois primeiros, mas medem os impactos na probabilidade de os indivíduos de 15 a 17 anos cursarem a escola. A variação negativa nas probabilidades, ou seja, a diminuição em pontos percentuais nas probabilidades observada no Gráfico 4 ocorre pelo fato de o sinal do tamanho relativo da coorte não ser o esperado.

Através dos gráficos percebe-se que, ao se tentar controlar a autocorrelação espacial dos dados, realmente ocorre redução do impacto do tamanho relativo da coorte e da razão professor/população em idade escolar. Porém, quando se acrescenta o efeito indireto dessas variáveis, ou seja, o efeito via seus vizinhos, observa-se que o aumento ou a diminuição são próximos dos encontrados no modelo hierárquico não-espacial.

Esse resultado pode suscitar uma importante questão. Apesar de a estimativa do modelo hierárquico ser enviesada por não controlar a autocorrelação espacial dos dados, a interpretação dos seus resultados pode não estar completamente errada, se considerarmos que ele está captando os efeitos direto e indireto das variáveis contextuais, enquanto o modelo hierárquico-espacial separa esses dois efeitos.

## Conclusão

Esse artigo buscou analisar o dividendo demográfico através de duas metodologias distintas: a hierárquica e a hierárquica-espacial. Com esses modelos foi possível considerar no estudo outras variáveis que não apenas as demográficas. Assim, os fatores familiares e a restrição de oferta escolar dos municípios também entraram na análise, além da correlação espacial advinda do fato de se trabalhar com dados georreferenciados.

Os resultados da primeira metodologia demonstraram que uma menor pressão demográfica tem importante efeito no acesso escolar das crianças de 7 a 14 anos,

constituindo-se no fator contextual mais importante. Menores coortes nessa faixa etária provocam um aumento na probabilidade de freqüentar a escola. Esse resultado corrobora as preposições do estudo de Coale e Hoover (1956).

Para o acesso dos jovens de 15 a 17 anos, o impacto do dividendo demográfico mostrou efeito inverso ao encontrado para a faixa etária anterior. Já a restrição de oferta correspondeu ao principal fator para o maior acesso escolar desses jovens. Dessa forma, fica evidenciada a necessidade de políticas públicas que aumentem a disponibilidade de oferta do nível de ensino médio nos municípios brasileiros.

Quando se incorpora a autocorrelação espacial na análise – modelo hierárquico-espacial –, observam-se dois tipos de impactos das variáveis independentes, que foram chamados de efeito direto e efeito indireto. O primeiro compreende o impacto na freqüência escolar decorrente de mudanças nas variáveis contextuais dentro do próprio município e o segundo corresponde ao impacto relativo a mudanças nas variáveis contextuais nos municípios vizinhos. Observa-se que, quando se controla a dependência espacial, o efeito direto diminui significativamente, tornando a restrição de oferta o fator contextual mais importante para o acesso à escola nas duas faixas etárias. Destaca-se também que, para a freqüência escolar dos jovens de 15 a 17 anos, o coeficiente do tamanho relativo da coorte em idade escolar, no modelo que inclui a defasagem espacial dos resíduos, passa a ser negativo, o que é esperado. Tal fato sugere que o tamanho relativo da coorte de 15 a 17 anos está correlacionado com outros fatores que não foram incluídos na análise, e que esses são correlacionados espacialmente.

Quanto ao impacto indireto do dividendo demográfico, os resultados demonstraram seu importante efeito, ou seja, variações no tamanho relativo da coorte em idade escolar nos municípios vizinhos provocam importante mudança nas probabilidades de as crianças e jovens freqüentarem a escola.

## Referências bibliográficas

- ANSELIN, L. **Spatial econometrics**. Dallas: University of Texas, School of Social Science, 1999.
- \_\_\_\_\_. Spatial externalities, spatial multipliers, and spatial econometrics. **International Regional Science Review**, v.26, n.2, p.153-166, 2003.
- BARROS, R. P. et al. Determinantes do desempenho educacional do Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.31, n.1, p.1-42, abr. 2001.
- BRYK, A. S.; RAUDENBUSH, S. W. **Hierarchical linear models: applications and data analysis methods**. 2.ed. Newbury Park, California: Sage, 2002.
- COALE, A.; HOOVER, E. **Population growth and economic development in low-income countries: a case study of India's prospects**. Princeton, N.J.: Princeton University, 1958.
- HANUSHEK, E. A.; GOMES-NETO, J. B.; HARBISON, R. W. Efficiency-enhancing investments in school quality. In: BIRDSALL, N.; SABOT, R. H. (Eds.). **Opportunity forgone: education in Brazil**. Washington, DC.: Inter-American development Bank, 1996, p.385-424.
- HANUSHEK, E. A. Publicly provided education. In: AUERBACH, A. J.; FELDSTEIN, M. (Eds.). **Handbook of public economics**. Amsterdam: Elsevier, v.4, 2002, p. 2045-2141.
- JONES, G. J. **Population dynamics and education and health planning**. Geneva: International Labour Office, 1990 (Wold Employment Programme. Paper, 8).
- KRUEGER, A. B. Experimental estimates of education production functions. **The Quarterly Journal of Economics**, v.114, n.2, p.497-532, May 1999.
- \_\_\_\_\_. Reassessing the view that American schools are broken. **Economic Research Bank of New York**, v.4, n.1, p.29-46, Mar. 1998.
- LAZEAR, E. P. **Educational production**. Cambridge: National Bureau of Economic Research, 1999 (Working paper, 7349).
- LEE, J.; BARRO, R. J. **Schooling quality in a cross section of countries**. Cambridge: National Bureau of Economic Research, 1997 (Working paper, 6198).
- MORENOFF, J. D. Neighborhood mechanisms and the spatial dynamics of birth weight. **American Journal of Sociology**, v.108, n.5, p.976-1017, Mar. 2003.
- RIANI, J. L. R. Impactos da estrutura etária em indicadores de educação no Brasil, 1991. **Revista Brasileira de Estudos Populacionais**, v.18, n.1/2, p.15-34, jan./dez. 2002.
- \_\_\_\_\_. **Determinantes do resultado educacional no Brasil: família, perfil escolar dos municípios e dividendo demográfico numa abordagem hierárquica e espacial**. Tese de Doutorado. Universidade Federal de Minas Gerais, 2005.
- RIOS-NETO, E. L. G.; CÉSAR, C. C.; RIANI, J. L. R. Estratificação educacional e progressão escolar por série no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.32, n.3, p.395-415, dez. 2002.
- SCHULTZ, T. P. School expenditures and enrollments, 1960-1980: the effects of income, prices and population growth. In: JOHNSON, D. G.; LEE, R. D. (Eds.). **Population growth and economics development: issues and evidence**. Madison, WI: University of Wisconsin, 1987, p.413-478.
- SILVA, N. V. **Atualização da escala socioeconômica de ocupação para 1980**. Laboratório Nacional de Computação Científica, CNPq, 1985. Mimeografado.
- \_\_\_\_\_. **Posição social das ocupações**. Rio de Janeiro: IBGE, Centro de Informática, 1973. Mimeografado.
- SILVA, N. V.; HASENBALG, C. **Recursos familiares e transições educacionais**. Versão preliminar apresentada no Workshop de Demografia da Educação. Associação Brasileira de Estudos Populacionais, Abep, Salvador, Bahia, junho 2001.
- TODD, P. E.; WOLPIN, K. I. On the specification and estimation of the production function for cognitive achievement. **The Economic Journal**, v.113, n.485, p.F3-F33, Feb. 2003.

## Resumen

### *Análisis del dividendo demográfico en la matrícula escolar en Brasil en un abordaje jerárquico y jerárquico-espacial*

Este artículo busca investigar los determinantes de la matrícula escolar en Brasil, en los niveles de enseñanza fundamental y media, combinando dos tradiciones de los estudios educacionales. La primera se refiere al dividendo demográfico y la segunda a los estudios de los determinantes educacionales que utilizan el andamiaje teórico de la Función de Producción Educacional. Sumado a eso, se busca incluir la dinámica espacial en el análisis educacional, a través de la construcción de una metodología alternativa que combina los modelos jerárquico y espacial. Los principales resultados indicaron que una menor presión demográfica constituye uno de los factores contextuales más importantes para la matrícula no fundamental. Por otro lado, para la enseñanza media la disponibilidad de oferta educacional, representada por la razón profesor / población en edad escolar, es el factor municipal más relevante. Cuando se considera el espacio en el análisis a través del modelo jerárquico-espacial, se observa que la razón de dependencia y las demás variables contextuales de las ciudades vecinas afectan la matrícula de determinado municipio.

**Palabras-clave:** Educación. Función de Producción Educacional. Dividendo demográfico.

## Abstract

### *An analysis of the demographic dividend on school enrollment in Brazil a hierarchical and spatial-hierarchical approach*

The aim of this paper is to investigate the determinants of attendance at elementary and secondary schools in Brazil, combining two traditions in educational studies. The first tradition has to do with the demographic dividend and the second with studies on educational determinants based on the theoretical approach of Educational Production Function. This article takes into account spatial dynamics in educational analysis by constructing an alternative methodology that combines spatial and hierarchical models. The main results show that, on the elementary level, lower demographic pressure is among the most important contextual factors that influence enrolment of children in schools. On the other hand, educational supply, represented by the ratio between the number of teachers and the school-age population, is the most important aspect in cities. The hierarchical-spatial model shows that the dependency ratio, together with the additional contextual variables in adjacent cities, has an influence on school enrollment of any given municipality.

**Key words:** Education, educational production function, demographic dividend.

Recebido para publicação em 31/10/2006.

Aceito para publicação em 31/05/2007.

